

MMUSIC WG
Internet-Draft
Intended status: Standards Track
Expires: February 13, 2012

M. Garcia-Martin
Ericsson
S. Veikkolainen
Nokia
R. Gilman
August 12, 2011

Miscellaneous Capabilities Negotiation in the Session Description
Protocol (SDP)
draft-garcia-mmusic-sdp-miscellaneous-caps-00

Abstract

SDP has been extended with a capability negotiation mechanism framework that allows the endpoints to negotiate transport protocols and attributes. This framework has been extended with a media capabilities negotiation mechanism that allows endpoints to negotiate additional media-related capabilities. This negotiation is embedded into the widely-used SDP offer/answer procedures.

This memo extends the SDP capability negotiation framework to allow endpoints to negotiate three additional SDP capabilities. In particular, this memo provides a mechanism to negotiate bandwidth ("b=" line), connection data ("c=" line), and titles ("i=" line for each session or media).

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on February 13, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	4
2. Conventions Used in This Document	4
3. Protocol Description	4
3.1. Extensions to SDP	4
3.1.1. Bandwidth Capability	6
3.1.2. Connection Data Capability	9
3.1.3. Title Capability	11
3.2. Session Level versus Media Level	14
3.3. Offer/Answer model extensions	14
3.3.1. Generating the Initial Offer	14
3.3.2. Generating the Answer	15
3.3.3. Offerer Processing of the Answer	15
3.3.4. Modifying the Session	15
4. Field Replacement Rules	15
5. IANA Considerations	15
5.1. New SDP Attributes	15
5.2. New Option Tags	16
5.3. New SDP Capability Negotiation Configuration Parameters	16
6. Security Considerations	17
7. Acknowledgments	17
8. References	17
8.1. Normative References	17
8.2. Informative References	18
Authors' Addresses	18

1. Introduction

The Session Description Protocol (SDP) [RFC4566] is intended for describing multimedia sessions for the purposes of session announcement, session invitation, and other forms of multimedia session initiation. SDP has been extended with a capability negotiation mechanism framework [RFC5939] which allows the endpoints to negotiate capabilities, such as support for Real-time Transport Protocol (RTP) [RFC3550] and Secure Real-time Transport Protocol (SRTP) [RFC3711]. The SDP media capabilities [I-D.ietf-mmusic-sdp-media-capabilities] provides negotiation capabilities to media lines as well.

The capability negotiation is embedded into the widely used SDP offer/answer procedure [RFC3264]. This memo provides the means to negotiate further capabilities than those specified in the SDP capability negotiation mechanism framework [RFC5939] and the SDP media capabilities negotiation [I-D.ietf-mmusic-sdp-media-capabilities]. In particular, this memo provides a mechanism to negotiate bandwidth ("b="), connection data ("c="), and session or media titles ("i=").

Since the three added capabilities are highly unconnected, it is not expected that implementations will support all of them at the same time. Instead, it is expected that applications will choose their needed capability for their specific purpose. Due to this, we are writing the normative part pertaining to each capability in a self-contained section: Section 3.1.1 describes the bandwidth capability extension, Section 3.1.2 describes the connection data capability extension, and Section 3.1.3 describes the title capability extension. Separate option tags are defined for each capability.

2. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [RFC2119] and indicate requirement levels for compliant implementations.

3. Protocol Description

3.1. Extensions to SDP

The SDP Capability Negotiation Framework [RFC5939] and the SDP media capabilities negotiation [I-D.ietf-mmusic-sdp-media-capabilities]

specify attributes for negotiating SDP capabilities. These documents specify new attributes (e.g., 'acap', 'tcap', 'mcap') for achieving their purpose. In this document we define three new additional capability attributes for SDP lines of the the general form:

type=value

for types "b", "c", and "i". The corresponding capability attributes are defined as "bcap" for bandwidth capability, "ccap" for connection data capability, and "icap" for title capability, respectively.

From the sub-rules of "a=" line in SDP [RFC4566], SDP attributes are of the form:

attribute	= (att-field ":" att-value) / att-field
att-field	= token
att-value	= byte-string

Capability attributes use only the 'att-field:att-value' form.

The new attributes may be referenced in potential configurations ("a=pcfg") or in latent configurations ("a=lcfg"), as productions conforming to the extension-config-list as defined in [RFC5939].

extension-config-list	= ["+"] ext-cap-name "=" ext-cap-list
ext-cap-name	= 1*(ALPHA / DIGIT) ; ALPHA and DIGIT defined in RFC 5234
ext-cap-list	= 1*VCHAR ; VCHAR defined in RFC 5234

The optional "+" is used to indicate that the extension is mandatory and MUST be supported in order to use that potential configuration.

The attributes may be referenced in actual configurations ("a=acfg") as productions conforming to the sel-extension-config defined in [RFC5939].

sel-extension-config = ext-cap-name "=" 1*VCHAR

The specific parameters are defined in the individual description of each capability, below.

The "bcap", "ccap", and "icap" capability attributes can be provided either at the session or media level. According to the SDP Capability Negotiation [RFC5939], each extension capability must specify the implication of making it part of a configuration at the media level.

According to SDP [RFC4566], "b=", "c=", and "i=" lines may appear

either at session or media level. In line with this, the "bcap", "ccap", and "icap" capability attributes, when declared at session level, are to be interpreted as-if that attribute was provided with that value at the session level. The "bcap", "ccap" and "icap" capability attributes declared at media level, are to be interpreted as-if that capability attribute was declared at the media level.

For example, extending the example in [I-D.ietf-mmusic-sdp-media-capabilities] with "icap" and "bcap" capability attributes, we get the following SDP:

```
v=0
a=bcap:1 CT:200
a=icap:1 Video conference
m=audio 54320 RTP/AVP 0
a=mcap:1 L16/8000/1
a=mcap:2 L16/16000/2
a=pcfg:1 m=1|2, pt=1:99,2:98
m=video 66544 RTP/AVP 100
a=mcap:3,4 H263-1998/90000
a=rtpmap:100 H264/90000
a=pcfg:10 m=3 pt=3:101 b=1 i=1
```

Figure 1: Example SDP offer with bcap and icap defined at session level

The above SDP defines one PCMU audio stream and one H.264 video stream. It also defines two Media Format capabilities (numbered 1 and 2), using L16 audio at 8 kbps and 16 kbps, respectively, as well as Media Format capabilities for H.263 video (numbered 3 and 4). The Media Format capabilities all appear at the media level. The example also contains a single bandwidth capability and a single title capability at session level. According to the definition above, when the capabilities defined in "bcap", and "icap" attributes are referenced from the potential configuration, in the resulting SDP they are to be interpreted as session level attributes (but the Media Format capabilities are to be interpreted as media level attributes).

3.1.1. Bandwidth Capability

According to RFC 4566 [RFC4566] the bandwidth field denotes the proposed bandwidth to be used by the session or media. In this memo, we specify the bandwidth capability attribute which can also appear either at session or media level. The bandwidth field is specified in RFC 4566 [RFC4566] with the following syntax:

b=<bwtype>:<bandwidth>

where <bwtype> is an alphanumeric modifier giving the meaning of the <bandwidth> figure.

In this document, we define a new capability attribute: the bandwidth capability attribute "bcap". This attribute lists bandwidth as capabilities according to the following definition:

"a=bcap:" bw-cap-num 1*WSP bwtype ":" bandwidth CRLF

where <bw-cap-num> is a unique integer between 1 and $2^{31}-1$ (both included) user to number the bandwidth capability, and the other elements are as defined for the "b=" field in SDP [RFC4566].

This format satisfies the general attribute production rules in SDP [RFC4566] according to the following Augmented Backus-Naur Form (ABNF) [RFC5234] syntax:

att-field	=	"bcap"
att-value	=	bw-cap-num 1*WSP bwtype ":" bandwidth
bw-cap-num	=	1*10(DIGIT) ; DIGIT defined in RFC 5234

Figure 2: Syntax of the bcap attribute

Negotiation of bandwidth per media stream can be useful when negotiating media encoding capabilities with different bandwidths.

3.1.1.1. Configuration Parameters

The SDP capability negotiation framework [RFC5939] provides for the existence of the "pcfg" and "acfg" attributes. The concept is extended by the SDP media capabilities negotiation [I-D.ietf-mmusic-sdp-media-capabilities] with an "lcfg" attribute that conveys latent configurations.

Extensions to the "pcfg" and "lcfg" attributes are defined through <extension-config-list>, and extensions to the "acfg" attribute are defined through the <sel-extension-config> as defined in the SDP Capability Negotiation [RFC5939].

In this document we extend the <extension-config-list> field to be able to convey lists of bandwidth capabilities in latent or potential configurations, according to the following Augmented Backus-Naur Form (ABNF) [RFC5234] syntax:

```

extension-config-list = bandwidth-config-list
bandwidth-config-list = ["+"] "b=" bw-cap-list *(BAR bw-cap-list)
                        ; BAR defined in RFC 5939
bw-cap-list           = bw-cap-num *("," bw-cap-num)
bw-cap-num            = 1*10(DIGIT)    ; DIGIT defined in RFC 5234

```

Figure 3: Syntax of the bandwidth parameter in lcfg and pcfg attributes

Each bandwidth capability configuration is a comma-separated list of bandwidth capability attribute numbers where 'bw-cap-num' refers to the bw-cap-num bandwidth capability numbers defined explicitly earlier in this document, and hence must be between 1 and $2^{31}-1$ (both included). Alternative bandwidth configurations are separated by a vertical bar ("|").

The above syntax is very flexible, allowing referencing to multiple "b=" lines per configuration, even for the same bwtype. While the need for such definitions is not seen, we have not restricted this, as it is not restricted in SDP [RFC4566] either.

The bandwidth parameter to the actual configuration attribute ("a=acfg") is formulated as a sel-extension-config with

```
ext-cap-name = "b"
```

hence

```

sel-extension-config = sel-bandwidth-config
sel-bandwidth-config = "b=" bw-cap-list ; bw-cap-list as above.

```

Figure 4: Syntax of the bandwidth parameter in acfg attributes

3.1.1.2. Option tag

The SDP Capability Negotiation Framework [RFC5939] allows for capability negotiation extensions to be defined. Associated with each such extension is an option tag that identifies the extension in question. Hereby, we define a new option tag "bcap-v0" that identifies support for the bandwidth capability. The endpoints using the "bcap" capability attribute SHOULD add the option tag to other existing option tags present in the "csup" and "creq" attributes in SDP, according to the procedures defined in the SDP Capability Negotiation Framework [RFC5939].

3.1.2. Connection Data Capability

According to SDP [RFC4566], the connection data field in SDP contains the connection data, and it has the following syntax:

```
c=<nettype> <addrtype> <connection-address>
```

where <nettype> indicates the network type, <addrtype> indicates the address type, and the <connection-address> is the connection address, which is dependent on the address type.

At the moment, network types already defined include "IN", which indicates Internet network type, and ATM (see RFC 3108 [RFC3108]), used for describing ATM bearer connections. The CS descriptions in SDP document [I-D.ietf-mmusic-sdp-cs] adds a "PSTN" network type for expressing a PSTN circuit switch.

SDP [RFC4566] permits specification of connection data at the session or at the media level. In order to permit negotiation of connection data at the media level, we define the connection data capability attribute ("a=ccap") in the form:

```
"a=ccap:" conn-cap-num 1*WSP nettype SP addrtype SP connection-  
address CRLF
```

where <conn-cap-num> is a unique ordinal identifier of the connection data capability, and the other elements are as defined in [RFC4566].

This format corresponds to the [RFC4566] attribute production rules according to the following Augmented Backus-Naur Form (ABNF) [RFC5234] syntax:

```
att-field      = "ccap"  
att-value      = conn-cap-num 1*WSP nettype SP addrtype  
                SP connection-address  
conn-cap-num   = 1*DIGIT ; 1 to 2^31-1, inclusive
```

Figure 5: Syntax of the ccap attribute

The "ccap" capability attribute allows for expressing alternative connection address ("c=") lines in SDP as part of the SDP capability negotiation process. The "ccap" capability attribute is intended to be used only when there is no other mechanism available for negotiating alternative connection address information, such as when the <nettype> is different among the alternative addresses. The "ccap" attribute MUST NOT be used in situations where an existing mechanism (such as Interactive Connectivity Establishment (ICE) [RFC5245]) can be used to select between different connection

addresses.

3.1.2.1. Configuration Parameters

The SDP Capability Negotiation Framework [RFC5939] provides for the existence of the "pcfg" and "acfg" attributes, which can carry one or more potential configurations to be negotiated. The concept is extended by the the Media Capabilities Negotiation [I-D.ietf-mmusic-sdp-media-capabilities] with an "lcfg" attribute that conveys latent configurations.

In this document we define a <connection-config> parameter to be used to specify a connection data capability in a potential or latent configuration attribute. The parameter follows the form of an extension-config-list, with

```
ext-cap-name = "c"

ext-cap-list = conn-cap-list
```

where, according to the following Augmented Backus-Naur Form (ABNF) [RFC5234] syntax:

```
extension-config-list = conn-config-list
conn-config-list      = "c=" conn-cap-list
conn-cap-list         = conn-cap-num *(BAR conn-cap-num)
conn-cap-num          = 1*DIGIT ; 1 to 2^32-1 inclusive
```

Figure 6: Syntax of the connection data parameter in lcfg and pcfg attributes

Each capability configuration alternative contains a single connection data capability attribute number and refers to the conn-cap-num capability number defined explicitly earlier in this document, and hence must be between 1 and $2^{31}-1$ (both included). The connection data capability allows the expression of only a single capability in each alternative, rather than a list of capabilities, since no more than a single connection data field is permitted per media block. Nevertheless, it is still allowed to express alternative potential connection configurations separated by a vertical bar ("|").

The connection data parameter to the actual configuration attribute ("a=acfg") is formulated as a sel-extension-config with

```
ext-cap-name = "c"
```

hence

```
sel-extension-config = sel-connection-config  
sel-connection-config = "c=" conn-cap-num ; as defined above.
```

Figure 7: Syntax of the connection data parameter in acfg attributes

3.1.2.2. Option tag

The SDP Capability Negotiation Framework [RFC5939] solution allows for capability negotiation extensions to be defined. Associated with each such extension is an option tag that identifies the extension in question. Hereby, we define a new option tag of "ccap-v0" that identifies support for the connection data capability. This option tag SHOULD be added to other existing option tags present in the "csup" and "creq" attributes in SDP, according to the procedures defined in the SDP Capability Negotiation Framework [RFC5939].

3.1.3. Title Capability

SDP [RFC4566] provides for the existence of an information field expressed in the format of the "i=" line, which can appear either at the session level or at the media level. An "i=" line that is present at the session level is known as the "session name", and its purpose is to convey a human-readable textual information about the session.

The "i=" line in SDP can also appear at the media level, in which case it is used to provide human-readable information about the media stream to which it is related, e.g., it may indicate the purpose of the media stream. The "i=" line is not to be confused with the label attribute ("a=label:", [RFC4574]) which provides a machine-readable tag. It is foreseen that applications declaring capabilities related to different configurations of a media stream may need to provide different identifying information for each of those configurations. That is, a party might offer alternative media configurations for a stream, each of which represents a different presentation of the same or similar information. For example, an audio stream might offer English or Spanish configurations, or a video stream might offer a choice of video source such as speaker camera, group camera, or document viewer. The title capability is needed to inform the answering user in order to select the proper choice, and the label is used to inform the offering machine which choice the answerer has selected. Hence, there is value in defining a mechanism to provide titles of media streams as capabilities.

According to SDP [RFC4566], the session information ("i=") line has the following syntax:

"i="text

where "text" represents a human-readable text indicating the purpose of the session or media stream.

In this document we define a new capability attribute: the Title capability, "icap". This attribute lists session or media titles as capabilities, according to the following definition:

"a=icap:" title-cap-num 1*WSP text

where <title-cap-num> is a unique integer between 1 and $2^{31}-1$ (both included) user to number the unique ordinal identifier of the particular title capability and <text> is a human-readable text that indicates the purpose of the session or media stream it is supposed to characterize.

As an example, one might use:

a=icap:1 Document Camera

to define a title capability number 1 to identify a particular source of a media stream.

The title capability attribute satisfies the general attribute production rules in SDP [RFC4566] according to the following Augmented Backus-Naur Form (ABNF) [RFC5234] syntax:

```
att-field      = "icap"
att-value      = title-cap-num 1*WSP text
                  ; text defined in RFC 4566
title-cap-num  = 1*10(DIGIT) ; DIGIT defined in RFC 5234
```

Figure 8: Syntax of the icap attribute

3.1.3.1. Configuration Parameters

The SDP Capability Negotiation Framework [RFC5939] provides for the existence of the "pcfg" and "acfg" attributes. The concept is extended by the SDP media capabilities negotiation [I-D.ietf-mmusic-sdp-media-capabilities] with an "lcfg" attribute that conveys latent configurations.

In this document, we define an <title-config-list> parameter to be used to convey title capabilities in a potential or latent configuration. This parameter is defined as an <extension-config-list> with the following associations:

```

ext-cap-name = "i"

ext-cap-list = title-cap-list

```

This leads to the following definition for the title capability parameter:

```

extension-config-list = title-config-list
title-config-list      = ["+"] "i=" title-cap-list
title-cap-list         = title-cap-num *(BAR title-cap-num)
                        ; BAR defined in RFC 5939
title-cap-num          = 1*10(DIGIT) ; DIGIT defined in RFC 5234

```

Figure 9: Syntax of the title capability parameter in lcfg and pcfg attributes

Each potential capability configuration contains a single title capability attribute number where 'title-cap-num' is the title capability number defined explicitly earlier in this document, and hence must be between 1 and $2^{31}-1$ (both included). The title capability allows the expression of only a single capability in each alternative, since no more than a single title field is permitted per block. Nevertheless, it is still allowed to express alternative potential title configurations separated by a vertical bar ("|").

An endpoint includes a plus sign ("+") in this configuration attribute to mandate support for this extension. An endpoint that receives this attribute prefixed with a plus sign and does not support this extension MUST treat that potential configuration as not valid.

3.1.3.2. Option Tag

At present, it is difficult to envision a scenario in which the "icap" attribute must be supported or the offer must be rejected. In most cases, if the icap attribute or its contents were to be ignored, an offered configuration could still be chosen based on other criteria such as configuration numbering. However, one might imagine an SDP offer that contained English and Spanish potential configurations for an audio stream. The session might be unintelligible if the choice is based on configuration numbering, rather than informed user selection. Based on such considerations, it may well prove useful to announce the ability to use the icap attribute and its contents to select media configurations, or to inform the user about the selected configuration(s). Therefore, we define a new option tag of "icap-v0" that identifies support for the title capability. This option tag SHOULD be added to other existing option tags present in the "csup" and/or "creq" attributes in SDP,

according to the procedures defined in the SDP Capability Negotiation Framework [RFC5939]. The discussion above suggests that "icap-v0" will typically appear in a "csup" attribute, but rarely in a "creq" attribute.

3.2. Session Level versus Media Level

The "bcap", "ccap" and "icap" attributes can appear at the session level and/or at the media level. Endpoints MUST interpret capabilities declared at session level as part of the session level in the resulting SDP for that particular configuration. Endpoints MUST interpret capabilities declared at media level as part of the media level in the resulting SDP for that particular configuration.

If a "bcap" capability for the same bwtype is declared at both session and media level, the media level attribute overrides the value of the session level attribute.

To avoid confusion, the <type-attr-num> for each "a=bcap", "a=ccap", and "a=icap" line must be unique across all capability attributes of the same type within the entire session description.

3.3. Offer/Answer model extensions

In this section, we define extensions to the offer/answer model defined in SDP Offer/Answer Model [RFC3264] and extended in the SDP Capability Negotiation [RFC5939] to allow for bandwidth and title capabilities to be used with the SDP Capability Negotiation framework.

3.3.1. Generating the Initial Offer

When an endpoint generates an initial offer and wants to use the functionality described in the current document, it first defines appropriate values for the bandwidth, connection data, and/or title capability attributes according to rules defined in [RFC4566] for "b=", "c=" and "i=" lines. The endpoint then MUST include the respective capability attributes and associated values in the SDP offer. The preferred configurations for each media stream are identified following the media line in a "pcfg" attribute. Bandwidth and title capabilities may also be referenced in latent configurations, defined in [RFC5939].

The offer SHOULD include the level of capability negotiation extensions needed to support this functionality in a "creq" attribute.

3.3.2. Generating the Answer

When the answering party receives the offer, and if it supports the required capability negotiation extensions, it SHOULD select the most preferred configuration it can support for each media stream, and build the answer accordingly, as defined in Section 3.6.2 of the SDP Capability Negotiation [RFC5939].

3.3.3. Offerer Processing of the Answer

When the offerer receives the answer, it MUST process the media lines according to normal SDP processing rules to identify the media stream(s) accepted by the answer, if any. The "acfg" attribute, if present, may be used to verify the proposed configuration used to form the answer, and to infer the lack of acceptability of higher-preference configurations that were not chosen.

3.3.4. Modifying the Session

If, at a later time, one of the parties wishes to modify the operating parameters of a session, e.g. by adding a new media stream, or by changing the properties used on an existing stream, it may do so via the mechanisms defined for SDP offer/answer [RFC3264].

4. Field Replacement Rules

To simplify the construction of SDP records, given the need to include fields within the media description in question for endpoints that do not support capabilities negotiation, we define some simple field-replacement rules for those fields invoked by potential or latent configurations. In particular, any "i=" or "c=" line invoked by a configuration MUST replace the corresponding line, if present within the media description in question. Any "b=" line invoked by a configuration MUST replace any "b=" of the same bandwidth type at the media level.

5. IANA Considerations

5.1. New SDP Attributes

IANA is hereby requested to register new attributes in the "att-field (both session and media level)" of the "Session Description Protocol (SDP) Parameters" registry, according to the following registration form:

Attribute name: bcap
Long form name: Bandwidth Capability
Type of attribute: Both media and session level
Subject to charset: No
Purpose: Negotiate session or media-level bandwidths
Appropriate values: See RFC XXXX
 [Note to the RFC Editor: Please replace the above RFC XXXX
 with the RFC number of this specification.
Contact name: Miguel A. Garcia,
 Miguel.A.Garcia@ericsson.com

Attribute name: ccap
Long form name: Connection Data Capability
Type of attribute: Both media and session level
Subject to charset: No
Purpose: Negotiate media-level connection data
Appropriate values: See RFC XXXX
 [Note to the RFC Editor: Please replace the above RFC XXXX
 with the RFC number of this specification.
Contact name: Miguel A. Garcia,
 Miguel.A.Garcia@ericsson.com

Attribute name: icap
Long form name: Title Capability
Type of attribute: Both media and session level
Subject to charset: Yes
Purpose: Negotiate human-readable information
 describing the session or media
Appropriate values: See RFC XXXX
 [Note to the RFC Editor: Please replace the above RFC XXXX
 with the RFC number of this specification.
Contact name: Miguel A. Garcia,
 Miguel.A.Garcia@ericsson.com

5.2. New Option Tags

IANA is hereby requested to add the new option tags "bcap-v0", "ccap-v0", and "icap-v0", defined herein, to the "SDP Capability Negotiation Option Tag subregistry" of the "Session Description Protocol (SDP) Parameters" registry.

5.3. New SDP Capability Negotiation Configuration Parameters

IANA is hereby requested to add the new parameter identifiers "b" for "bandwidth", "c" for "connection data", and "i" for "title" to the "SDP Capability Negotiation Potential Configuration Parameters"

subregistry of the "Session Description Protocol (SDP) Parameters" registry. These parameters are permitted in 'lcfg', 'acfg', and 'pcfg' attributes.

6. Security Considerations

This document provides an extension on top of RFC 4566 [RFC4566], RFC 3264 [RFC3264], SDP Capability Negotiation Framework [RFC5939], and SDP media capabilities negotiation [I-D.ietf-mmusic-sdp-media-capabilities]. As such, the security considerations of those documents apply.

The bandwidth capability attribute may be used for reserving resources at endpoints and intermediaries which inspect the SDP. Modification of the bandwidth value by an attacker can lead to the network being underutilized (too high bandwidth value) or congested (too low bandwidth value). In case it is essential to protect the bandwidth value, one of the security mechanisms proposed in [RFC5939] should be used.

The "i=" line and thus the value carried in the title capability attribute is intended for human-readable description only. It should not be parsed programmatically.

7. Acknowledgments

Thanks to Christer Holmberg, Alf Heidermark, and Ingemar Johansson for arguing for the existence of this document and early reviewing it. Thanks to Flemming Andreassen, Andrew Allen, and Jonathan Lennox for a detailed review and many improvement suggestions.

8. References

8.1. Normative References

- [I-D.ietf-mmusic-sdp-media-capabilities]
Gilman, R., Even, R., and F. Andreassen, "SDP Media Capabilities Negotiation",
draft-ietf-mmusic-sdp-media-capabilities-11 (work in progress), February 2011.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model

with Session Description Protocol (SDP)", RFC 3264, June 2002.

- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.
- [RFC5939] Andreassen, F., "Session Description Protocol (SDP) Capability Negotiation", RFC 5939, September 2010.

8.2. Informative References

- [I-D.ietf-mmusic-sdp-cs]
Garcia, M. and S. Veikkolainen, "Session Description Protocol (SDP) Extension For Setting Up Audio and Video Media Streams Over Circuit-Switched Bearers In The Public Switched Telephone Network (PSTN)", draft-ietf-mmusic-sdp-cs-06 (work in progress), February 2011.
- [RFC3108] Kumar, R. and M. Mostafa, "Conventions for the use of the Session Description Protocol (SDP) for ATM Bearer Connections", RFC 3108, May 2001.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.
- [RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.
- [RFC4574] Levin, O. and G. Camarillo, "The Session Description Protocol (SDP) Label Attribute", RFC 4574, August 2006.
- [RFC5245] Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols", RFC 5245, April 2010.

Authors' Addresses

Miguel A. Garcia-Martin
Ericsson
Calle Via de los Poblados 13
Madrid, 28033
Spain

Phone: +34 91 339 1000
Email: miguel.a.garcia@ericsson.com

Simo Veikkolainen
Nokia
P.O. Box 407
NOKIA GROUP, FI 00045
Finland

Phone: +358 50 486 4463
Email: simo.veikkolainen@nokia.com

Robert R. Gilman
3243 W. 11th Ave. Dr.
Broomfield, Colorado 80020
U.S.A.

Phone: +1 303 898 9780
Email: bob_gilman@comcast.net

This Internet-Draft, draft-garcia-mmusic-sdp-miscellaneous-caps-00.txt, has expired, and has been deleted from the Internet-Drafts directory. An Internet-Draft expires 185 days from the date that it is posted unless it is replaced by an updated version, or the Secretariat has been notified that the document is under official review by the IESG or has been passed to the RFC Editor for review and/or publication as an RFC. This Internet-Draft was not published as an RFC.

Internet-Drafts are not archival documents, and copies of Internet-Drafts that have been deleted from the directory are not available. The Secretariat does not have any information regarding the future plans of the authors or working group, if applicable, with respect to this deleted Internet-Draft. For more information, or to request a copy of the document, please contact the authors directly.

Draft Authors:

Miguel Garcia-Martin<miguel.a.garcia@ericsson.com>
Simo Veikkolainen<simo.veikkolainen@nokia.com>
Robert Gilman<bob_gilman@comcast.net>

mmusic
Internet-Draft
Intended status: Standards Track
Expires: May 3, 2012

B. Greevenbosch
Y. Hui
Huawei
October 31, 2011

Signal 3D format
draft-greevenbosch-mmusic-signal-3d-format-02

Abstract

This document introduces the SDP attribute "3dFormat", which provides format description of stereoscopic 3D video. In addition, the grouping mechanism for SDP is extended to cater for stereoscopic 3D video.

Note

Discussion and suggestions for improvement are requested, and should be sent to mmusic@ietf.org.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 3, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	4
2. Requirements notation	5
3. Definitions	6
4. The "3dFormat" attribute	8
5. Grouping	11
6. Combinations of attribute values and group usage	12
7. SDP offer/answer with 3D support	14
8. SDP offer/answer without 3D support	16
8.1. Frame packing	16
8.2. 2D and auxiliary as a single stream	16
8.3. 2D and auxiliary as two separate streams	16
8.4. Simulcast of L- and R-views	17
9. Examples	18
9.1. One single frame compatible stream	18
9.2. Two separate streams	18
9.3. C-stream and depth map stream	18
9.4. Stereoscopic 3D video with two different formats	19
10. Formal ABNF grammar of the "3dFormat" attribute	21
11. Security Considerations	22
12. IANA Considerations	23
12.1. "3dFormat" attribute	23
12.2. "3DS" value for "group" semantics	24
13. Acknowledgements	25
14. Normative References	26
Authors' Addresses	27

1. Introduction

In stereoscopic 3D multimedia applications, two views are displayed, one for the left eye and one for the right eye.

There are various ways of formatting the views of Stereoscopic 3D video. Examples of 3D formats are frame packing (see [HDMIV1.4a] and [ISO/IEC 14496-10]) and the combination of 2D video and auxiliary data such as depth maps or parallax maps (for both, see [ISO/IEC 23002-3]). Stereoscopic 3D video may be carried over a single stream or over several streams, depending on its 3D format.

In multimedia streaming applications, the Session Description Protocol (SDP) [RFC4566] can be used to provide to the receiver sufficient information about the media streams, and to enable the receiver to join and participate in the session.

This document defines an extension to SDP that provides sufficient information about the format of stereoscopic 3D video carried in the media stream(s). Before accessing the stream(s), the receiver can use the 3D format description from SDP to determine whether it has the capability to receive and render the stereoscopic 3D video content, and whether it can participate in the session.

The mentioned SDP extension is a new SDP attribute "3dFormat", which provides the format description of stereoscopic 3D video. The design of the attribute is based on the following requirements, which are listed only for informational purposes:

- o It MUST be possible to signal that the left and right views are carried in a single stream, by the use of frame packing.
- o It MUST be possible to signal that 2D video and auxiliary video (such as depth maps) are carried in a single stream.
- o It MUST be possible to signal that the left and right views are carried in two separate streams.
- o It MUST be possible to signal that 2D video and auxiliary video (such as depth maps) are carried in separate streams.

To bind multiple video streams that carry a single stereoscopic 3D video, this document also extends the SDP grouping mechanism from [RFC5888].

2. Requirements notation

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Definitions

2D video

Video that does not in itself contain depth or parallax information.

auxiliary video

A sequence of depth or parallax maps, which are used to add depth to 2D video.

C-view

The centre view: a visual entity as seen from a viewpoint between the left and right eyes. The C-view can be used to calculate the L- and R-views.

C-stream

A 2D video stream consisting of a sequence of C-views.

depth map

A two dimensional map, each pixel of which defines the depth of one or more pixels in an associated 2D video frame.

depth map stream

An auxiliary stream, which contains a sequence of depth maps. The depth map stream is synchronised with the associated 2D video stream.

frame packing

A format that packs the L- and R-views into a single 2D video stream. The packing may be done spatially, where each video frame is divided into sub-frames, one containing the L-view and one containing the R-view. The packing can also be done sequentially, where alternating video frames represent L- and R-views.

legacy answerer

An answerer (in the SDP offer/answer model [RFC3264]) that does not support the "3dFormat" attribute. The legacy answerer can be the streaming server or the streaming client, but is not compliant to this document.

L-view

A visual entity that is to be projected to the left eye.

L-stream

A 2D video stream consisting of a sequence of L-views.

parallax map

A two dimensional map, each pixel of which defines the parallax of one or more pixels in an associated 2D video frame.

parallax map stream

An auxiliary stream, which contains a sequence of parallax maps. The parallax map stream is synchronised with the associated 2D video stream.

R-view

A visual entity that is to be projected to the right eye.

R-stream

A 2D video stream consisting of a sequence of R-views.

stereoscopic 3D video

The L- and R-streams, ready to be projected to the viewer's left and right eyes.

sub-frame

A part of a video frame.

4. The "3dFormat" attribute

The media-level SDP attribute "3dFormat" signals the format of stereoscopic 3D video. The attribute transfers this information through two parameters: one indicating the format type of the stereoscopic 3D video carried in the media stream(s), and the other indicating the type of the video component, which is a constituent element of the stereoscopic 3D video. The video component type depends on the format type of the stereoscopic 3D video. The syntax of the attribute is defined as follows:

a=3dFormat:<Format Type> <Component Type>

The <Format Type> can have the following values (as indicated between the quotes):

"FP" Frame Packing

The L- and R-views are packed into a single stream. The packing may use a side-by-side, top-and-bottom, interleaved, checkerboard or frame sequential format.

"SC" Simulcast

The L- and R-streams are transmitted separately.

"2DA" 2D + auxiliary

2D video and auxiliary data (such as depth maps or parallax maps) are transmitted. These can be transmitted in a single stream, as well as in two separate streams.

The <Component Type> can have the following values (as indicated between the quotes):

"C" Centre view

The associated stream is a C-stream.

"CD" centre view and depth map

The associated stream contains both the C-view and depth map sequences.

"ChB" Checkerboard

The video frame consists of alternating pixels from the corresponding L- and R-views, as illustrated by Figure 1.

"CP" Centre view and parallax map

The associated stream contains both the C-view and parallax map sequences.

- "D" Depth map
The associated stream is a sequence of depth maps.
- "L" Left view
The associated stream is the L-stream.
- "LD" Left view and depth map
The associated stream contains both the L-view and depth map sequences.
- "LIL" Line Interleaved
Each video frame consists of alternating scan lines from the L- and R-views.
- "LP" Left view and parallax map
The associated stream contains both the L-view and parallax map sequences.
- "P" Parallax map
The associated stream is a sequence of parallax maps.
- "R" Right view
The associated stream is the R-stream.
- "SbS" Side by Side
Each video frame is divided in two equally sized sub-frames, spatially positioned side by side of each other. One sub-frame contains the L-view, whereas the other contains the R-view.
- "Seq" Frame Sequential
The single video stream consists of alternating frames from the L- and R-streams. Additional signalling, e.g. AVC SEI messages [ISO/IEC 14496-10], is needed to signal which frames contain L- and which contain R-views.
- "TaB" Top and Bottom
Each video frame is divided in two equally sized sub-frames, spatially positioned above each other. One sub-frame contains the L-view, whereas the other contains the R-view.

```
+--+--+--+--+--+
|L|R|L|R|L|R|
+--+--+--+--+--+
|R|L|R|L|R|L|
+--+--+--+--+--+
|L|R|L|R|L|R|
+--+--+--+--+--+
```

The checkerboard pattern. The transmitted video frame is composed of pixels from the L- and R-views. Samples from the L-view are indicated with "L", whereas samples from the R-view are indicated with "R".

Figure 1

5. Grouping

When multiple streams carry a single stereoscopic 3D video, (e.g. C-stream and parallax map, or separately transmitted L- and R-streams), the grouping mechanism from [RFC5888] MUST be used.

However, to cater for the special requirements of 3D signalling, the semantics are expanded:

```
group-attribute      = "a=group:" semantics *(SP identification-tag)
semantics             = "LS" / "FID" / "3DS" / semantics-extension
semantics-extension = token
```

The grouping is needed when multiple streams carry a single stereoscopic 3D video. This is the case when the <format type> is "SC", or the <format type> is "2DA" and the 2D video and auxiliary data are transmitted as multiple streams. A group with the "3DS" semantics is called a "3DS group".

A 3DS group MUST NOT contain data that is (potentially) inconsistent with other data in the 3DS group:

- o A 3DS group MUST NOT contain both a parallax map stream and a depth map stream.
- o A 3DS group MUST NOT contain more than one parallax map stream.
- o A 3DS group MUST NOT contain more than one depth map stream.
- o A 3DS group MUST contain at least one 2D video stream.
- o If a 3GS group contains an L- and an R-stream, it MUST NOT contain a depth map or a parallax map.
- o If a 3DS group contains only one 2D video stream, it MUST also contain a parallax map stream or a depth map stream.
- o If a 3DS group contains a parallax map stream or a depth map stream, it MUST also contain a 2D video stream.

6. Combinations of attribute values and group usage

The following table summarises the possible combinations of attribute values and grouping:

	FP	SC	2DA
C			D/P, 3DS
CD			T
ChB	T		
CP			T
D			C/L, 3DS
L		R, 3DS	D/P, 3DS
LD			T
LIL	T		
LP			T
P			C/L, 3DS
R		L, 3DS	
SbS	T		
Seq	T		
TaB	T		

The table is to be read as follows:

- o The columns indicate <Format Type> values, whereas the rows indicate <Component Type> values.
- o For one particular column, we denote the <Format Type> value by "FT" and the <Component Type> value by "CT".
- o When an entry in the table is empty, it means that the corresponding combination of FT and CT is not allowed.

- o When an entry in the table contains a single <Component Type> value CTsec, it means that another stream with the <Component Type> value CTsec and the same <Format Type> value FT is needed.
- o When multiple <Component Type> values are listed, separated by a "/" symbol, only one secondary stream is needed, which must have one of the listed <Component Type> values, and the same <Format Type> value FT.
- o When an entry contains "3DS", it means that a 3DS group is needed.
- o When an entry in the table contains the letter "T" (true), it means that the corresponding combination FT and CT is allowed, that there is no required secondary stream, and that a 3DS group is not needed.

7. SDP offer/answer with 3D support

This section describes how the SDP offer/answer model (see [RFC3264]) can be used to negotiate the 3D format. It is assumed that both offerer and answerer are compliant to this document. The case where the answerer is a legacy answerer is described in Section 8.

An example where the SDP offer/answer model can be used to negotiate the 3D format, is the case where the offerer offers two representations of the same stereoscopic 3D video: one frame packed and one as L/R simulcast. In this case, the answerer can select the format of its preference, according to its capabilities or as a trade-off between bandwidth and video quality.

There may also be cases where the answerer prefers to receive a 2D version, even when it supports stereoscopic 3D video and the "3dFormat" attribute. For example, this might happen when the user prefers to watch without glasses this time.

The following statements apply for the answerer:

- o The answerer **MUST NOT** omit the "3dFormat" attribute for the accepted streams. The answerer **MAY** omit the "3dFormat" attribute for the rejected streams.
- o The answerer **MUST NOT** change the value of the "3dFormat" attribute. This means, that the answerer can only choose between the 3D formats advertised in the offer.
- o In case the offer contains simulcast of the L- and R-view, the answerer **MAY** choose just one view. In this case, it **MUST** select only that view. This means that the port number of the other view **MUST** be set to zero in the answer.
- o In case the offer contains a 2D stream and an auxiliary stream as separate streams, the answerer **MAY** choose only the 2D stream. In this case, it **MUST** select the 2D stream, and **MUST NOT** select the auxiliary stream. This means that the port number of the auxiliary stream **MUST** be set to zero in the answer.
- o In case the offer contains a 2D stream and an auxiliary stream as a single stream, the answerer **MAY** choose to reject the stream by setting the port number in the answer to zero.
- o In case of frame packing, if the answerer prefers not to have frame packing, it **MUST** reject the stream by setting the port number in the answer to zero.

- o If the answerer selects multiple 3D formats, it MUST be prepared to send/receive (depending on whether it is a streaming server or client or both) associated streams simultaneously.

The following statements apply for the offerer:

- o The offerer MUST check if the "3dFormat" attribute is included in the answer. If it is not, it SHOULD handle the answer as described in Section 8.
- o The offerer SHOULD list the 3D formats in order of preference.
- o When multiple 3D formats are selected, the offerer MAY initiate all associated streams. Alternatively, it MAY update its offer with a reduced number of 3D formats.
- o If all 3D formats have been rejected, the offerer MAY issue a new offer with 2D video instead.
- o If only an auxiliary stream is selected in the answer, the offerer SHOULD update its offer with only the associated 2D video stream. Alternatively, it MAY update its offer advertising another 3D format.

8. SDP offer/answer without 3D support

Since a legacy answerer does not support the "3dFormat" attribute, it might reject the offer. In this case the offerer MAY send a new offer with only a 2D video stream.

On the other hand, it is also possible that the legacy answerer accepts the offer but omits the "3dFormat" attribute in the answer. In this case the offerer is able to deduct that the answerer is a legacy answerer without 3D support. In the following subsections, we describe what the offerer still can do to provide a good user experience with a legacy answerer, for each of the 3D format styles. We assume that the offer was accepted, but a legacy answerer was detected.

8.1. Frame packing

In case the original offer contains frame packing, and the answer does not contain the "3dFormat" attribute, the offerer SHOULD treat that media stream as a 2D stream.

Note: in some cases, the answerer may be a legacy device that is capable of rendering a frame packed 3D stream, but does not understand the "3dFormat" attribute. For example, the user may be able to switch manually to 3D. Therefore, the server MAY stream the frame packed video as it is.

8.2. 2D and auxiliary as a single stream

If the original offer contains a 2D video and an auxiliary video in a single stream, and the answer does not contain the "3dFormat" attribute, the offerer SHOULD treat that media stream as a 2D stream.

8.3. 2D and auxiliary as two separate streams

When the offerer sends an offer to a legacy answerer, and the offer contains a 2D video and an auxiliary video in two separate streams, there are the following possibilities:

- o If the answerer selects only the 2D video stream then 2D video streaming can be done as agreed.
- o If the answerer selects only the auxiliary video, the offerer MAY treat that stream as a 2D video stream. If it does not, the offerer SHOULD update its offer without the auxiliary video.
- o If the answerer selects both video streams, but omits the "3dFormat" attribute, the offerer MAY update its offer without the

auxiliary video.

In case the offerer updates its offer by setting the port for auxiliary video to zero, it MUST NOT include the "3dFormat" attribute or use "3DS" grouping for the 2D stream.

8.4. Simulcast of L- and R-views

When the offerer sends an offer to simulcast the L- and R-view to the legacy answerer, we have the following possibilities:

- o If the answerer selects only one video stream, the offerer MAY stream the 2D video as agreed.
- o If the answerer selects both video streams, but omits the "3dFormat" attribute, the offerer MAY update its offer with only the L- or the R-stream.

In case the offerer updates its offer with only the L- or R-stream by setting one of the ports to zero, it MUST NOT include the "3dFormat" attribute or use "3DS" grouping for the offered stream.

9. Examples

9.1. One single frame compatible stream

The following is an example of an SDP description of a session which contains a single stream, in which the L- and R-streams are packed, in side by side fashion.

```
v=0
o=Alice 2890844526 2890842807 IN IP4 131.163.72.4
s=The technology of 3D-TV
c=IN IP4 131.164.74.2
t=0 0
m=video 49170 RTP/AVP 99
a=rtpmap:99 H264/90000
a=3dFormat:FP Sbs
m=audio 52890 RTP/AVP 10
a=rtpmap:10 L16/16000/2
```

9.2. Two separate streams

The following is an example of an SDP description of a session with an audio stream, an L-stream and an R-stream.

```
v=0
o=Alice 2890844526 2890842807 IN IP4 131.163.72.4
s=The technology of 3D-TV
c=IN IP4 131.164.74.2
t=0 0
a=group:3DS 1 2
m=video 49170 RTP/AVP 99
a=rtpmap:99 H264/90000
a=3dFormat:SC L
a=mid:1
m=video 49172 RTP/AVP 101
a=rtpmap:101 H264/90000
a=3dFormat:SC R
a=mid:2
m=audio 52890 RTP/AVP 10
a=rtpmap:10 L16/16000/2
```

9.3. C-stream and depth map stream

The following is an example of an SDP description of a session with an audio stream, a C-stream and a depth map stream.

```
v=0
o=Alice 2890844526 2890842807 IN IP4 131.163.72.4
s=The technology of 3D-TV
c=IN IP4 131.164.74.2
t=0 0
a=group:3DS 1 2
m=video 49170 RTP/AVP 99
a=rtpmap:99 H264/90000
a=3dFormat:2DA C
a=mid:1
m=video 49172 RTP/AVP 101
a=rtpmap:101 H264/90000
a=3dFormat:2DA D
a=mid:2
m=audio 52890 RTP/AVP 10
a=rtpmap:10 L16/16000/2
```

9.4. Stereoscopic 3D video with two different formats

In the following example, there are two different formats for stereoscopic 3D video. One consists of stream 1 (C-stream) and stream 2 (parallax map stream), whereas the other consists of stream 3 (L-stream) and stream 4 (R-stream). There also is an audio stream, which can be used with both formats.


```
v=0
o=Alice 2890844526 2890842807 IN IP4 131.163.72.4
s=The technology of 3D-TV
c=IN IP4 131.164.74.2
t=0 0
a=group:3DS 1 2
a=group:3DS 3 4
m=video 49170 RTP/AVP 99
a=rtpmap:99 H264/90000
a=3dFormat:2DA C
a=mid:1
m=video 49172 RTP/AVP 101
a=rtpmap:101 H264/90000
a=3dFormat:2DA P
a=mid:2
m=video 49174 RTP/AVP 103
a=rtpmap:103 H264/90000
a=3dFormat:SC L
a=mid:3
m=video 49176 RTP/AVP 105
a=rtpmap:105 H264/90000
a=3dFormat:SC R
a=mid:4
m=audio 52890 RTP/AVP 10
a=rtpmap:10 L16/16000/2
```

10. Formal ABNF grammar of the "3dFormat" attribute

This section contains the formal ABNF grammar of the "3dFormat" attribute.

```
3dFormat-attribute      = "a=3dFormat:" formatType componentType
formatType              = "FP"/"SC"/"2DA"/formatType-extension
formatType-extension    = token
componentType           = "C"/"CD"/"ChB"/"CP"/"D"/"L"/"LD"/
                        "LIL"/"LP"/"P"/"R"/"SbS"/"Seq"/"TaB"/
                        componentType-extension
componentType-extension = token
```

11. Security Considerations

The authors foresee no security issues in addition to those already listed in [RFC4566].

12. IANA Considerations

12.1. "3dFormat" attribute

Following the guidelines in [RFC4566], the SDP attribute has to be registered at IANA:

- o Contact name/email: authors of this RFC
- o Attribute name: 3dFormat
- o Long-form attribute name: Attribute for signalling the format of a stereoscopic 3D video carried in the media stream(s).
- o Type of attribute: media level
- o Subject to charset: no

The "3dFormat" SDP media-level attribute is used to signal the format of stereoscopic 3D video, carried in one or more media stream(s).

The attribute has the following syntax:

a=3dFormat:<Format Type> <Component Type>

The <Format Type> indicates the format type of the stereoscopic 3D video carried in the media stream(s). It indicates whether the stereoscopic 3D video is frame packed, simulcast or consists of a 2D video stream and an auxiliary stream. The <Format Type> can have the following values (as indicated between the quotes):

"FP"	frame packed
"SC"	simulcast
"2DA"	2D + auxiliary

The <Component Type> indicates the type of the video component, which is a constituent element of the stereoscopic 3D video. It can have the following values:

"C"	centre view
"CD"	centre view and depth map
"ChB"	checkerboard
"CP"	centre view and parallax map
"D"	depth map
"L"	left view
"LD"	left view and depth map
"LIL"	line interleaved
"LP"	left view and parallax map
"P"	parallax map
"R"	right view
"SbS"	side by side
"Seq"	frame sequential
"TaB"	top and bottom

12.2. "3DS" value for "group" semantics

Following the standards action policy from [RFC5226], the following semantics have to be registered with IANA in the "Semantics for the "group" SDP Attribute" registry under "SDP Parameters":

-----+-----+-----+
Semantics Token Reference
-----+-----+-----+
3D synchronised 3DS this RFC
-----+-----+-----+

13. Acknowledgements

The authors would like to thank Stephen Botzko, Imed Bouazizi, Pedro Capelastegui, Roni Even, Miguel Garcia, Ted Hardie, Jonathan Lennox, Yue Peiyu and Tian Linyi for their review comments.

14. Normative References

- [HDMIv1.4a]
HDMI, "HDMI Specification Version 1.4a", March 2010.
- [ISO/IEC 23002-3]
MPEG, "MPEG video technologies part 3: Representation of auxiliary video and supplemental information", ISO/IEC FDIS 23002-3:2007(E), December 2002.
- [ISO/IEC 14496-10]
MPEG, "H.264/MPEG-4 Part 10: Advanced video coding for generic audiovisual services", ISO/IEC FDIS 14496-10:2010, March 2010.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.
- [RFC5888] Camarillo, G. and H. Schulzrinne, "The Session Description Protocol (SDP) Grouping Framework", RFC 5888, June 2010.

Authors' Addresses

Bert Greevenbosch
Huawei Technologies Co., Ltd.
Huawei Industrial Base
Bantian, Longgang District
Shenzhen 518129
P.R. China

Phone: +86-755-28978088
Email: bert.greevenbosch@huawei.com

Hui Yu
Huawei Technologies Co., Ltd.
Huawei Nanjing R&D Center
101 Software Avenue
Yuhuatai District
Nanjing 210012
P.R. China

Phone: +86-25-56620323
Email: huiyu@huawei.com

MMUSIC Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 21, 2012

C. Holmberg
Ericsson
H. Alvestrand
Google
October 19, 2011

Multiplexing Negotiation Using Session Description Protocol (SDP) Port
Numbers
draft-holmberg-mmusic-sdp-bundle-negotiation-00.txt

Abstract

This specification defines a new SDP Grouping Framework SDP grouping framework extension, "BUNDLE", that can be used with the Session Description Protocol (SDP) Offer/Answer mechanism to negotiate the usage of bundled media, which refers to the usage of a single 5-tuple for media associated with multiple SDP media descriptions ("m=" lines).

Status of this Memo

This Internet-Draft is submitted to IETF in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 21, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must

include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Terminology	3
3. Conventions	4
4. Applicability Statement	4
5. SDP Grouping Framework BUNDLE Extension Semantics	4
6. SDP Offer/Answer Procedures	4
6.1. General	4
6.2. SDP Offerer Procedures	5
6.3. SDP Answerer Procedures	6
6.4. Bundled SDP Information	6
6.4.1. General	6
6.4.2. Bandwidth (b=)	6
7. Single vs Multiple RTP Sessions	6
7.1. General	6
7.2. Single RTP Session	6
8. Usage With ICE	7
8.1. General	7
8.2. Candidates	7
9. Security Considerations	8
10. Example	8
11. IANA Considerations	10
12. Acknowledgements	10
13. Change Log	10
14. References	10
14.1. Normative References	10
14.2. Informative References	11
Authors' Addresses	11

1. Introduction

In the IETF RTCWEB WG, a need to use a single 5-tuple for sending and receiving media associated with multiple SDP media descriptions ("m=" lines) has been identified. This would e.g. allow the usage of a single set of Interactive Connectivity Establishment (ICE) [RFC5245] candidates for multiple media descriptions. Normally different media types (audio, video etc) will be described using different media descriptions.

This specification defines a new SDP Grouping Framework SDP grouping framework [RFC5888] extension, "BUNDLE", that can be used with the Session Description Protocol (SDP) Offer/Answer mechanism [RFC3264] to negotiate the usage of bundled media, which refers to the usage of a single 5-tuple for media associated with multiple SDP media descriptions ("m=" lines).

When an endpoint generates an SDP Offer or SDP Answer [RFC3264], which includes a "BUNDLE" group, each "m=" line associated with the group will share a single port number value.

As defined in RFC 4566 [RFC4566], the semantics of multiple "m=" lines using the same port number value are undefined, and there is no grouping defined by such means. Instead, an explicit grouping mechanism needs to be used to express the intended semantics. This specification provides such extension.

When media is transported using the Real-Time Protocol (RTP) [RFC3550], the default assumption of the mechanism is that all media associated with a "BUNDLE" group will form a single RTP Session [RFC3550]. However, future specifications can extend the mechanism, in order to negotiate RTP Session multiplexing, i.e. "BUNDLE" groups where media associated with a group form multiple RTP Sessions.

The mechanism is backward compatible. Entities that do not support the "BUNDLE" grouping extension, or do not want to enable the mechanism for a given session, are expected to generate a "normal" SDP Answer, using different port number values for each "m=" line, to the SDP Offer. The SDP Offerer [RFC3264] will still use a single port number value for each media, but as the SDP Answerer [RFC3264] will use separate ports a single 5-tuple will not be used for media associated with multiple "m=" lines between the endpoints.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this

document are to be interpreted as described in RFC 2119 [RFC2119].

5-tuple: A collection of the following values: source address, source port, destination address, destination port and protocol.

Bundled media: Two or more RTP streams using a single 5-tuple. The RTCP streams associated with the RTP streams also use a single 5-tuple, which might be the same, but can also be different, as the one used by the RTP streams.

3. Conventions

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [RFC2119].

4. Applicability Statement

The mechanism in this specification only applies to the Session Description Protocol (SDP) [RFC4566], when used together with the SDP Offer/Answer mechanism [RFC3264].

5. SDP Grouping Framework BUNDLE Extension Semantics

This section defines a new SDP Grouping Framework extension, "BUNDLE".

The "BUNDLE" extension can be indicated using an SDP session-level 'group' attribute. Each SDP media description ("m=" line) that is grouped together, using an SDP media-level 'mid' attribute, is part of a specific "BUNDLE" group.

6. SDP Offer/Answer Procedures

6.1. General

When an SDP Offerer or SDP Answerer generates an SDP Offer or SDP Answer, that describes bundled media, it MUST insert an SDP session-level 'group' attribute, with a "BUNDLE" value, and assign SDP media-level 'mid' attribute values to each "m=" line associated with the "BUNDLE" group.

In addition, the entity that generates the SDP Offer or SDP Answer

MUST, for each "m=" line that is part of the "BUNDLE" group:

- o 1. Use the same port number value.
- o 2. Use the same connection data ("c=" line) value.
- o 3. Use the same SDP 'rtcp' attribute value, when used.
- o 4. Use the same ICE candidate values, when used.
- o 5. Insert an SDP 'rtcp-mux' attribute.

NOTE: If an entity wants to disable specific media ("m=" line) associated with a "BUNDLE" group, it will use a zero port number value for the "m=" line associated with the media.

6.2. SDP Offerer Procedures

When an SDP Offerer creates an SDP Offer, that offers bundled media, it MUST create the SDP Offer according to the procedures in Section 6.1.

If the associated SDP Answer contains an SDP session-level 'group' attribute, with a "BUNDLE" value, and the SDP Answer is created according to the procedures in Section 6.1 (the same port number value is used for each "m=" line associated with the "BUNDLE" group, etc), the SDP Offerer can start using the same 5-tuple for sending and receiving media, associated with the group, between the entities.

If the SDP Answer does not include a session-level SDP 'group' attribute, with a "BUNDLE" value, the SDP Offerer cannot use the same 5-tuple for media associated with multiple "m=" lines.

If the SDP Answerer indicates that it will not use bundled media, the SDP Offerer will still use the single port number value for each "m=" line associated with the offered "BUNDLE" group, and it will normally be able to separate each individual media. The default mechanism for separating media received on a single IP address and port doing this is by using a 5-tuple based mapping for each individual media. If the SDP Offerer is aware of the Synchronization Source (SSRC) [RFC3550] values that the SDP Answerer will use in the media it sends, and the SSRC values will be unique for each media, the SDP Offerer can separate media based on the SSRC values.

NOTE: Assuming symmetric media is used, the SDP Offerer can use the port information from the SDP Answer in order to create the 5-tuple mapping for each media.

If the SDP Offerer is not able to separate multiple media received on a single port, it MUST send a new SDP Offer, without offering bundled media, where a separate port number value is provided for each "m=" line of the SDP Offer.

If an SDP Offer, offering a "BUNDLE" group, and the SDP Offerer has reasons to believe that the rejection is due to the usage of a single port number value for multiple "m=" lines, the SDP Offerer SHOULD send a new SDP Offer, without a "BUNDLE" group, where a separate port number value is provide for each "m=" line of the SDP offer.

6.3. SDP Answerer Procedures

When an SDP Answerer receives an SDP Offer, which offers bundled media, and the SDP Answerer accepts the offered bundle group, the SDP Answerer MUST create an SDP Answer according to the procedures in Section 6.1.

If the SDP Answerer does not accept the "BUNDLE" group in the SDP Offer, it MUST NOT include a session-level 'group' attribute, with a "BUNDLE" value, in the associated SDP Answer. In addition, the SDP Answerer MUST provide separate port number values for each "m=" line of the SDP Answer.

6.4. Bundled SDP Information

6.4.1. General

This section describes how SDP information, given for each media description, is calculated into a single value for a "BUNDLE" group.

6.4.2. Bandwidth (b=)

The total proposed bandwidth is the sum of the proposed bandwidth for each "m=" line associated with a negotiated BUNDLE group.

7. Single vs Multiple RTP Sessions

7.1. General

When entities negotiate the usage of bundled media, the default assumption is that all media associated with the bundled media will form a single RTP session.

The usage of multiple RTP Sessions within a "BUNDLE" group is outside the scope of this specification. Other specification needs to extend the mechanism in order to allow negotiation of such bundle groups.

7.2. Single RTP Session

When a single RTP Session is used, media associated with all "m=" lines part of a bundle group share a single SSRC [RFC3550] numbering

space.

In addition, the following rules and restrictions apply for a single RTP Session:

- o - The dynamic payload type values used in the "m=" lines MUST NOT overlap.
- o - The "proto" value in each "m=" line MUST be identical (e.g. RTP/AVPF).
- o - A given SSRC SHOULD NOT transmit RTP packets using payload types that originates from different "m=" lines.

NOTE: The last bullet above is to avoid sending multiple media types from the same SSRC. If transmission of multiple media types are done with time overlap RTP and RTCP fails to function. Even if done in proper sequence this causes RTP Timestamp rate switching issues [ref to draft-ietf-avtext-multiple-clock-rates].

8. Usage With ICE

8.1. General

This section describes how to use the "BUNDLE" grouping mechanism together with the Interactive Connectivity Establishment (ICE) mechanism [RFC5245].

8.2. Candidates

When an ICE-enabled SDP Offerer sends an SDP offer, it MUST include ICE candidates for each "m=" line associated with a "BUNDLE" group. The candidate values MUST be identical for each "m=" line associated with the group. This rule applies also to subsequent SDP Offers, when the usage of bundled media has already been negotiated.

When an ICE-enabled SDP Answerer receives an SDP Offer, offering a "BUNDLE" group and ICE, if the SDP Answerer enables ICE, MUST include ICE candidates for each "m=" line of the SDP Answer. This also applies for "m=" lines that are part of a "BUNDLE" group, in which case the candidate values MUST be identical for each "m=" line associated with the group. This rule applies also to subsequent SDP Answers, when the usage of bundled media has already been negotiated.

Once the usage of bundled media has been negotiated, ICE connectivity checks and keep-alives only needs to be performed for the whole "BUNDLE" group, instead of for each individual m= line associated with the group.

9. Security Considerations

TBA

10. Example

The example below shows an SDP Offer, where bundled media is offered. The example also shows two SDP Answer alternatives: one where bundled media is accepted, and one where bundled media is rejected (or, not even supported) by the SDP Answerer.

SDP Offer (Bundled media offered)

```
v=0
o=alice 2890844526 2890844526 IN IP4 host.atlanta.com
s=
c=IN IP4 host.atlanta.com
t=0 0
a=group:BUNDLE foo bar
m=audio 10000 RTP/AVP 0 8 97
a=mid:foo
b=AS:200
a=rtpmap:0 PCMU/8000
a=rtpmap:8 PCMA/8000
a=rtpmap:97 iLBC/8000
m=video 10000 RTP/AVP 31 32
a=mid:bar
b=AS:1000
a=rtpmap:31 H261/90000
a=rtpmap:32 MPV/90000
```

SDP Answer (Bundled media accepted)

```
v=0
o=bob 2808844564 2808844564 IN IP4 host.biloxi.com
s=
c=IN IP4 host.biloxi.com
t=0 0
a=group:BUNDLE foo bar
m=audio 20000 RTP/AVP 0
a=mid:foo
b=AS:200
a=rtpmap:0 PCMU/8000
m=video 20000 RTP/AVP 32
a=mid:bar
b=AS:1000
```

a=rtpmap:32 MPV/90000

SDP Answer (Bundled media not accepted)

v=0
o=bob 2808844564 2808844564 IN IP4 host.biloxi.com
s=
c=IN IP4 host.biloxi.com
t=0 0
m=audio 20000 RTP/AVP 0
b=AS:200
a=rtpmap:0 PCMU/8000
m=video 30000 RTP/AVP 32
b=AS:1000
a=rtpmap:32 MPV/90000

SDP Offer with ICE (Bundled media offered)

v=0
o=alice 2890844526 2890844526 IN IP4 host.atlanta.com
s=
c=IN IP4 host.atlanta.com
t=0 0
a=group:BUNDLE foo bar
m=audio 10000 RTP/AVP 0 8 97
a=mid:foo
b=AS:200
a=rtpmap:0 PCMU/8000
a=rtpmap:8 PCMA/8000
a=rtpmap:97 iLBC/8000
a=candidate:1 1 UDP 1694498815 host.atlanta.com 10000 typ host
m=video 10000 RTP/AVP 31 32
a=mid:bar
b=AS:1000
a=rtpmap:31 H261/90000
a=rtpmap:32 MPV/90000
a=candidate:1 1 UDP 1694498815 host.atlanta.com 10000 typ host

11. IANA Considerations

This document requests IANA to register the new SDP Grouping semantic extension called BUNDLE.

12. Acknowledgements

The usage of the SDP grouping mechanism is based on a similar alternative proposed by Harald Alvestrand. The SDP examples are also modified versions from the ones in the Alvestrand proposal.

Thanks to the nice flight crew on AY 021 for providing good sparkling wine, and a nice working atmosphere, for working on this draft.

13. Change Log

[RFC EDITOR NOTE: Please remove this section when publishing]

Changes from draft-holmberg-mmusic-sdp-multiplex-negotiation-00

- o Draft name changed.
- o Harald Alvestrand added as co-author.
- o "Multiplex" terminology changed to "bundle".
- o Added text about single versus multiple RTP Sessions.
- o Added reference to RFC 3550.

14. References

14.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC5888] Camarillo, G. and H. Schulzrinne, "The Session Description Protocol (SDP) Grouping Framework", RFC 5888, June 2010.

14.2. Informative References

- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.
- [RFC5245] Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols", RFC 5245, April 2010.

Authors' Addresses

Christer Holmberg
Ericsson
Hirsalantie 11
Jorvas 02420
Finland

Email: christer.holmberg@ericsson.com

Harald Tveit Alvestrand
Google
Kungsbron 2
Stockholm 11122
Sweden

Email: harald@alvestrand.no

MMUSIC Working Group
Internet-Draft
Obsoletes: 2326 (if approved)
Intended status: Standards Track
Expires: April 30, 2012

H. Schulzrinne
Columbia University
A. Rao
Cisco
R. Lanphier

M. Westerlund
Ericsson AB
M. Stiemerling (Ed.)
NEC
October 28, 2011

Real Time Streaming Protocol 2.0 (RTSP)
draft-ietf-mmusic-rfc2326bis-28

Abstract

This memorandum defines RTSP version 2.0 which obsoletes RTSP version 1.0 which is defined in RFC 2326.

The Real Time Streaming Protocol, or RTSP, is an application-level protocol for setup and control of the delivery of data with real-time properties. RTSP provides an extensible framework to enable controlled, on-demand delivery of real-time data, such as audio and video. Sources of data can include both live data feeds and stored clips. This protocol is intended to control multiple data delivery sessions, provide a means for choosing delivery channels such as UDP, multicast UDP and TCP, and provide a means for choosing delivery mechanisms based upon RTP (RFC 3550).

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 30, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1.	Introduction	11
2.	Protocol Overview	13
2.1.	Presentation Description	13
2.2.	Session Establishment	14
2.3.	Media Delivery Control	15
2.4.	Session Parameter Manipulations	17
2.5.	Media Delivery	17
2.5.1.	Media Delivery Manipulations	18
2.6.	Session Maintenance and Termination	20
2.7.	Extending RTSP	21
3.	Document Conventions	23
3.1.	Notational Conventions	23
3.2.	Terminology	23
4.	Protocol Parameters	27
4.1.	RTSP Version	27
4.2.	RTSP IRI and URI	27
4.3.	Session Identifiers	29
4.4.	SMPTE Relative Timestamps	29
4.5.	Normal Play Time	30
4.6.	Absolute Time	31
4.7.	Feature-Tags	31
4.8.	Message Body Tags	31
4.9.	Media Properties	32
4.9.1.	Random Access and Seeking	33
4.9.2.	Retention	33
4.9.3.	Content Modifications	34
4.9.4.	Supported Scale Factors	34
4.9.5.	Mapping to the Attributes	34
5.	RTSP Message	35
5.1.	Message Types	35
5.2.	Message Headers	35
5.3.	Message Body	36
5.4.	Message Length	36
6.	General Header Fields	38
7.	Request	39
7.1.	Request Line	39
7.2.	Request Header Fields	41
8.	Response	43
8.1.	Status-Line	43
8.1.1.	Status Code and Reason Phrase	43
8.2.	Response Headers	46
9.	Message Body	48
9.1.	Message-Body Header Fields	48
9.2.	Message Body	49
10.	Connections	50
10.1.	Reliability and Acknowledgements	50

10.2.	Using Connections	51
10.3.	Closing Connections	53
10.4.	Timing Out Connections and RTSP Messages	54
10.5.	Showing Liveness	55
10.6.	Use of IPv6	56
10.7.	Overload Control	56
11.	Capability Handling	58
12.	Pipelining Support	60
13.	Method Definitions	61
13.1.	OPTIONS	62
13.2.	DESCRIBE	63
13.3.	SETUP	65
13.3.1.	Changing Transport Parameters	68
13.4.	PLAY	69
13.4.1.	General Usage	69
13.4.2.	Aggregated Sessions	74
13.4.3.	Updating current PLAY Requests	75
13.4.4.	Playing On-Demand Media	77
13.4.5.	Playing Dynamic On-Demand Media	78
13.4.6.	Playing Live Media	78
13.4.7.	Playing Live with Recording	79
13.4.8.	Playing Live with Time-Shift	79
13.5.	PLAY_NOTIFY	80
13.5.1.	End-of-Stream	81
13.5.2.	Media-Properties-Update	82
13.5.3.	Scale-Change	83
13.6.	PAUSE	84
13.7.	TEARDOWN	87
13.7.1.	Client to Server	87
13.7.2.	Server to Client	88
13.8.	GET_PARAMETER	89
13.9.	SET_PARAMETER	91
13.10.	REDIRECT	92
14.	Embedded (Interleaved) Binary Data	95
15.	Status Code Definitions	97
15.1.	Success 1xx	97
15.1.1.	100 Continue	97
15.2.	Success 2xx	97
15.2.1.	200 OK	97
15.3.	Redirection 3xx	97
15.3.1.	301 Moved Permanently	98
15.3.2.	302 Found	98
15.3.3.	303 See Other	98
15.3.4.	304 Not Modified	98
15.3.5.	305 Use Proxy	99
15.4.	Client Error 4xx	99
15.4.1.	400 Bad Request	99
15.4.2.	401 Unauthorized	99

15.4.3.	402 Payment Required	100
15.4.4.	403 Forbidden	100
15.4.5.	404 Not Found	100
15.4.6.	405 Method Not Allowed	100
15.4.7.	406 Not Acceptable	100
15.4.8.	407 Proxy Authentication Required	101
15.4.9.	408 Request Timeout	101
15.4.10.	410 Gone	101
15.4.11.	411 Length Required	101
15.4.12.	412 Precondition Failed	102
15.4.13.	413 Request Message Body Too Large	102
15.4.14.	414 Request-URI Too Long	102
15.4.15.	415 Unsupported Media Type	102
15.4.16.	451 Parameter Not Understood	102
15.4.17.	452 reserved	102
15.4.18.	453 Not Enough Bandwidth	103
15.4.19.	454 Session Not Found	103
15.4.20.	455 Method Not Valid in This State	103
15.4.21.	456 Header Field Not Valid for Resource	103
15.4.22.	457 Invalid Range	103
15.4.23.	458 Parameter Is Read-Only	103
15.4.24.	459 Aggregate Operation Not Allowed	103
15.4.25.	460 Only Aggregate Operation Allowed	103
15.4.26.	461 Unsupported Transport	104
15.4.27.	462 Destination Unreachable	104
15.4.28.	463 Destination Prohibited	104
15.4.29.	464 Data Transport Not Ready Yet	104
15.4.30.	465 Notification Reason Unknown	104
15.4.31.	466 Key Management Error	104
15.4.32.	470 Connection Authorization Required	105
15.4.33.	471 Connection Credentials not accepted	105
15.4.34.	472 Failure to establish secure connection	105
15.5.	Server Error 5xx	105
15.5.1.	500 Internal Server Error	105
15.5.2.	501 Not Implemented	105
15.5.3.	502 Bad Gateway	105
15.5.4.	503 Service Unavailable	106
15.5.5.	504 Gateway Timeout	106
15.5.6.	505 RTSP Version Not Supported	106
15.5.7.	551 Option not supported	106
16.	Header Field Definitions	107
16.1.	Accept	117
16.2.	Accept-Credentials	117
16.3.	Accept-Encoding	118
16.4.	Accept-Language	119
16.5.	Accept-Ranges	120
16.6.	Allow	120
16.7.	Authorization	120

16.8.	Bandwidth	121
16.9.	Blocksize	122
16.10.	Cache-Control	122
16.11.	Connection	125
16.12.	Connection-Credentials	125
16.13.	Content-Base	126
16.14.	Content-Encoding	126
16.15.	Content-Language	127
16.16.	Content-Length	128
16.17.	Content-Location	128
16.18.	Content-Type	129
16.19.	CSeq	129
16.20.	Date	130
16.21.	Expires	131
16.22.	From	131
16.23.	If-Match	132
16.24.	If-Modified-Since	132
16.25.	If-None-Match	133
16.26.	Last-Modified	134
16.27.	Location	134
16.28.	Media-Properties	134
16.29.	Media-Range	136
16.30.	MTag	137
16.31.	Notify-Reason	137
16.32.	Pipelined-Requests	137
16.33.	Proxy-Authenticate	139
16.34.	Proxy-Authorization	139
16.35.	Proxy-Require	139
16.36.	Proxy-Supported	140
16.37.	Public	141
16.38.	Range	141
16.39.	Referrer	143
16.40.	Request-Status	144
16.41.	Require	144
16.42.	Retry-After	145
16.43.	RTP-Info	145
16.44.	Scale	148
16.45.	Seek-Style	149
16.46.	Server	150
16.47.	Session	151
16.48.	Speed	152
16.49.	Supported	153
16.50.	Terminate-Reason	153
16.51.	Timestamp	154
16.52.	Transport	154
16.53.	Unsupported	161
16.54.	User-Agent	161
16.55.	Vary	162

16.56. Via	162
16.57. WWW-Authenticate	163
17. Proxies	164
17.1. Proxies and Protocol Extensions	165
17.2. Multiplexing and Demultiplexing of Messages	166
18. Caching	167
18.1. Validation Model	167
18.1.1. Last-Modified Dates	169
18.1.2. Message Body Tag Cache Validators	169
18.1.3. Weak and Strong Validators	169
18.1.4. Rules for When to Use Message Body Tags and Last-Modified Dates	171
18.1.5. Non-validating Conditionals	173
18.2. Invalidation After Updates or Deletions	173
19. Security Framework	175
19.1. RTSP and HTTP Authentication	175
19.2. RTSP over TLS	175
19.3. Security and Proxies	176
19.3.1. Accept-Credentials	177
19.3.2. User approved TLS procedure	178
20. Syntax	181
20.1. Base Syntax	181
20.2. RTSP Protocol Definition	183
20.2.1. Generic Protocol elements	183
20.2.2. Message Syntax	186
20.2.3. Header Syntax	190
20.3. SDP extension Syntax	199
21. Security Considerations	200
21.1. Remote denial of Service Attack	202
22. IANA Considerations	204
22.1. Feature-tags	204
22.1.1. Description	205
22.1.2. Registering New Feature-tags with IANA	205
22.1.3. Registered entries	205
22.2. RTSP Methods	206
22.2.1. Description	206
22.2.2. Registering New Methods with IANA	206
22.2.3. Registered Entries	206
22.3. RTSP Status Codes	207
22.3.1. Description	207
22.3.2. Registering New Status Codes with IANA	207
22.3.3. Registered Entries	207
22.4. RTSP Headers	207
22.4.1. Description	207
22.4.2. Registering New Headers with IANA	208
22.4.3. Registered entries	208
22.5. Accept-Credentials	209
22.5.1. Accept-Credentials policies	209

22.5.2.	Accept-Credentials hash algorithms	209
22.6.	Cache-Control Cache Directive Extensions	210
22.7.	Media Properties	211
22.7.1.	Description	211
22.7.2.	Registration Rules	211
22.7.3.	Registered Values	211
22.8.	Notify-Reason header	211
22.8.1.	Description	211
22.8.2.	Registration Rules	212
22.8.3.	Registered Values	212
22.9.	Range header formats	212
22.9.1.	Description	212
22.9.2.	Registration Rules	212
22.9.3.	Registered Values	213
22.10.	Terminate-Reason Header	213
22.10.1.	Redirect Reasons	213
22.10.2.	Terminate-Reason Header Parameters	213
22.11.	RTP-Info header parameters	214
22.11.1.	Description	214
22.11.2.	Registration Rules	214
22.11.3.	Registered Values	214
22.12.	Seek-Style Policies	215
22.12.1.	Description	215
22.12.2.	Registration Rules	215
22.12.3.	Registered Values	215
22.13.	Transport Header Registries	215
22.13.1.	Transport Protocol Specification	216
22.13.2.	Transport modes	217
22.13.3.	Transport Parameters	217
22.14.	URI Schemes	218
22.14.1.	The rtsp URI Scheme	218
22.14.2.	The rtspS URI Scheme	219
22.14.3.	The rtspU URI Scheme	220
22.15.	SDP attributes	221
22.16.	Media Type Registration for text/parameters	222
23.	References	224
23.1.	Normative References	224
23.2.	Informative References	226
Appendix A.	Examples	229
A.1.	Media on Demand (Unicast)	229
A.2.	Media on Demand using Pipelining	233
A.3.	Media on Demand (Unicast)	235
A.4.	Single Stream Container Files	239
A.5.	Live Media Presentation Using Multicast	241
A.6.	Capability Negotiation	242
Appendix B.	RTSP Protocol State Machine	244
B.1.	States	244
B.2.	State variables	244

B.3.	Abbreviations	244
B.4.	State Tables	245
Appendix C.	Media Transport Alternatives	251
C.1.	RTP	251
C.1.1.	AVP	251
C.1.2.	AVP/UDP	251
C.1.3.	AVPF/UDP	252
C.1.4.	SAVP/UDP	253
C.1.5.	SAVPF/UDP	255
C.1.6.	RTCP usage with RTSP	255
C.2.	RTP over TCP	257
C.2.1.	Interleaved RTP over TCP	257
C.2.2.	RTP over independent TCP	257
C.3.	Handling Media Clock Time Jumps in the RTP Media Layer	261
C.4.	Handling RTP Timestamps after PAUSE	265
C.5.	RTSP / RTP Integration	267
C.6.	Scaling with RTP	267
C.7.	Maintaining NPT synchronization with RTP timestamps	267
C.8.	Continuous Audio	267
C.9.	Multiple Sources in an RTP Session	267
C.10.	Usage of SSRCS and the RTCP BYE Message During an RTSP Session	267
C.11.	Future Additions	268
Appendix D.	Use of SDP for RTSP Session Descriptions	269
D.1.	Definitions	269
D.1.1.	Control URI	269
D.1.2.	Media Streams	270
D.1.3.	Payload Type(s)	271
D.1.4.	Format-Specific Parameters	271
D.1.5.	Directionality of media stream	271
D.1.6.	Range of Presentation	272
D.1.7.	Time of Availability	273
D.1.8.	Connection Information	273
D.1.9.	Message Body Tag	273
D.2.	Aggregate Control Not Available	274
D.3.	Aggregate Control Available	275
D.4.	Grouping of Media Lines in SDP	276
D.5.	RTSP external SDP delivery	276
Appendix E.	RTSP Use Cases	277
E.1.	On-demand Playback of Stored Content	277
E.2.	Unicast Distribution of Live Content	278
E.3.	On-demand Playback using Multicast	279
E.4.	Inviting an RTSP server into a conference	279
E.5.	Live Content using Multicast	280
Appendix F.	Text format for Parameters	282
Appendix G.	Requirements for Unreliable Transport of RTSP	283
Appendix H.	Backwards Compatibility Considerations	285
H.1.	Play Request in Play State	285

H.2.	Using Persistent Connections	285
Appendix I.	Changes	286
I.1.	Brief Overview	286
I.2.	Detailed List of Changes	287
Appendix J.	Acknowledgements	294
J.1.	Contributors	294
Appendix K.	RFC Editor Consideration	296
Authors' Addresses	297

1. Introduction

This memo defines version 2.0 of the Real Time Streaming Protocol (RTSP 2.0). RTSP 2.0 is an application-level protocol for setup and control over the delivery of data with real-time properties, typically streaming media. Streaming media is, for instance, video on demand or audio live streaming. Put simply, RTSP acts as a "network remote control" for multimedia servers, similar to the remote control for a DVD player.

The protocol operates between RTSP 2.0 clients and servers, but also supports the usage of proxies placed between clients and servers. Clients can request information about streaming media from servers by asking for a description of the media or use media description provided externally. The media delivery protocol is used to establish the media streams described by the media description. Clients can then request to play out the media, pause it, or stop it completely, as known from DVD players remote control or media players. The requested media can consist of multiple audio and video streams that are delivered as a time-synchronized streams from servers to clients.

RTSP 2.0 is a replacement of RTSP 1.0 [RFC2326] and obsoletes that specification. This protocol is based on RTSP 1.0 but is not backwards compatible other than in the basic version negotiation mechanism. The changes are documented in Appendix I. There are many

reasons why RTSP 2.0 can't be backwards compatible with RTSP 1.0 but some of the main ones are:

- o Most headers that needed to be extensible did not define the allowed syntax, preventing safe deployment of extensions;
- o The changed behavior of the PLAY method when received in Play state;
- o Changed behavior of the extensibility model and its mechanism;
- o The change of syntax for some headers.

In summary, there are so many small details that changing version became necessary to enable clarification and consistent behavior.

This document is structured as follows. It begins with an overview of the protocol operations and its functions in an informal way. Then a set of definitions of used terms and document conventions is introduced. It is followed by the actual RTSP 2.0 core protocol specification. The appendixes describe and define some

functionalities that are not part of the core RTSP specification, but which are still important to enable some usages. Among them, the RTP usage is defined in Appendix C and the SDP usage with RTSP is defined in Appendix D, which are two mandatory appendixes. While others include a number of informational parts discussing the changes, use cases, different considerations or motivations.

2. Protocol Overview

This section provides an informative overview of the different mechanisms in the RTSP 2.0 protocol, to give the reader a high level understanding before getting into all the different details. In case of conflict with this description and the later sections, the later sections take precedence. For more information about considered use cases for RTSP see Appendix E.

RTSP 2.0 is a bi-directional request and response protocol that first establishes a context including content resources (the media) and then controls the delivery of these content resources from the provider to the consumer. RTSP has three fundamental parts: Session Establishment, Media Delivery Control, and an extensibility model described below. The protocol is based on some assumptions about existing functionality to provide a complete solution for client controlled real-time media delivery.

RTSP uses text-based messages, requests and responses, that may contain a binary message body. An RTSP request starts with a method line that identifies the method, the protocol and version and the resource to act on. Following the method line are a number of RTSP headers. This part is ended by two consecutive carriage return line feed (CRLF) character pairs. The message body if present follows the two CRLF and the body's length is described by a message header. RTSP responses are similar, but start with a response line with the protocol and version, followed by a status code and a reason phrase. RTSP messages are sent over a reliable transport protocol between the client and server. RTSP 2.0 requires clients and servers to implement TCP, and TLS over TCP, as mandatory transports for RTSP messages.

2.1. Presentation Description

RTSP exists to provide access to multi-media presentations and content, but tries to be agnostic about the media type or the actual media delivery protocol that is used. To enable a client to implement a complete system, an RTSP-external mechanism for describing the presentation and the delivery protocol(s) is used. RTSP assumes that this description is either delivered completely out of bands or as a data object in the response to a client's request using the DESCRIBE method (Section 13.2).

Parameters that commonly have to be included in the Presentation Description are the following:

- o Number of media streams;

- o The resource identifier for each media stream/resource that is to be controlled by RTSP;
- o The protocol that each media stream is to be delivered over;
- o Transport protocol parameters that are not negotiated or vary with each client;
- o Media encoding information enabling a client to correctly decode the media upon reception;
- o An aggregate control resource identifier.

RTSP uses its own URI schemes ("rtsp" and "rtsps") to reference media resources and aggregates under common control.

This specification describes in Appendix D how one uses SDP [RFC4566] for Presentation Description

2.2. Session Establishment

The RTSP client can request the establishment of an RTSP session after having used the presentation description to determine which media streams are available, and also which media delivery protocol is used and their particular resource identifiers. The RTSP session is a common context between the client and the server that consists of one or more media resources that are to be under common media delivery control.

The client creates an RTSP session by sending a request using the SETUP method (Section 13.3) to the server. In the SETUP request the client also includes all the transport parameters necessary to enable the media delivery protocol to function in the "Transport" header (Section 16.52). This includes parameters that are pre-established by the presentation description but necessary for any middlebox to correctly handle the media delivery protocols. The Transport header in a request may contain multiple alternatives for media delivery in a prioritized list, which the server can select from. These alternatives are typically based on information in the presentation description.

The server determines if the media resource is available upon receiving a SETUP request and if any of the transport parameter specifications are acceptable. If that is successful, an RTSP session context is created and the relevant parameters and state is stored. An identifier is created for the RTSP session and included in the response in the Session header (Section 16.47). The SETUP response includes a Transport header that specifies which of the

alternatives has been selected and relevant parameters.

A SETUP request that references an existing RTSP session but identifies a new media resource is a request to add that media resource under common control with the already present media resources in an aggregated session. A client can expect this to work for all media resources under RTSP control within a multi-media content. However, aggregating resources from different content are likely to be refused by the server. The RTSP session as aggregate is referenced by the aggregate control URI, even if the RTSP session only contains a single media.

To avoid an extra round trip in the session establishment of aggregated RTSP sessions, RTSP 2.0 supports pipelined requests; i.e., the client can send multiple requests back-to-back without waiting first for the completion of any of them. The client uses client-selected identifier in the Pipelined-Requests header to instruct the server to bind multiple requests together as if they included the session identifier.

The SETUP response also provides additional information about the established sessions in a couple of different headers. The Media-Properties header includes a number of properties that apply for the aggregate that is valuable when doing media delivery control and configuring user interface. The Accept-Ranges header informs the client about which range formats that the server supports with these media resources. The Media-Range header inform the client about the time range of the media currently available.

2.3. Media Delivery Control

After having established an RTSP session, the client can start controlling the media delivery. The basic operations are Start by using the PLAY method (Section 13.4) and Halt by using the PAUSE method (Section 13.6). PLAY also allows for choosing the starting media position from which the server should deliver the media. The positioning is done by using the Range header (Section 16.38) that supports several different time formats: Normal Play Time (NPT) (Section 4.5), SMPTE Timestamps (Section 4.4) and absolute time (Section 4.6). The Range header does further allow the client to specify a position where delivery should end, thus allowing a specific interval to be delivered.

The support for positioning/searching within a content depends on the content's media properties. Content exists in a number of different types, such as: on-demand, live, and live with simultaneous recording. Even within these categories there are differences in how the content is generated and distributed, which affect how it can be

accessed for playback. The properties applicable for the RTSP session are provided by the server in the SETUP response using the Media-Properties header (Section 16.28). These are expressed using one or several independent attributes. A first attribute is Random Access, which expresses if positioning can be done, and with what granularity. Another aspect is whether the content will change during the lifetime of the session. While on-demand content will be provided in full from the beginning, a live stream being recorded results in the length of the accessible content growing as the session goes on. There also exist content that is dynamically built by another protocol than RTSP and thus also changes in steps during the session, but maybe not continuously. Furthermore, when content is recorded, there are cases where not the complete content is maintained, but, for example, only the last hour. All these properties result in the need for mechanisms that will be discussed below.

When the client accesses on-demand content that allows random access, the client can issue the PLAY request for any point in the content between the start and the end. The server will deliver media from the closest random access point prior to the requested point and indicate that in its PLAY response. If the client issues a PAUSE, the delivery will be halted and the point at which the server stopped will be reported back in the response. The client can later resume by sending a PLAY request without a range header. When the server is about to complete the PLAY request by delivering the end of the content or the requested range, the server will send a PLAY_NOTIFY request indicating this.

When playing live content with no extra functions, such as recording, the client will receive the live media from the server after having sent a PLAY request. Seeking in such content is not possible as the server does not store it, but only forwards it from the source of the session. Thus delivery continues until the client sends a PAUSE request, tears down the session, or the content ends.

For live sessions that are being recorded the client will need to keep track of how the recording progresses. Upon session establishment the client will learn the current duration of the recording from the Media-Range header. As the recording is ongoing the content grows in direct relation to the passed time. Therefore, each server's response to a PLAY request will contain the current Media-Range header. The server should also regularly send every 5 minutes the current media range in a PLAY_NOTIFY request. If the live transmission ends, the server must send a PLAY_NOTIFY request with the updated Media-Properties indicating that the content stopped being a recorded live session and instead became on-demand content; the request also contains the final media range. While the live

delivery continues the client can request to play the current live point by using the NPT timescale symbol "now", or it can request a specific point in the available content by an explicit range request for that point. If the requested point is outside of the available interval the server will adjust the position to the closest available point, i.e., either at the beginning or the end.

A special case of recording is that where the recording is not retained longer than a specific time period, thus as the live delivery continues the client can access any media within a moving window that covers, for example, "now" to "now" minus 1 hour. A client that pauses on a specific point within the content may not be able to retrieve the content anymore. If the client waits too long before resuming the pause point, the content may no longer be available. In this case the pause point will be adjusted to the end of the available media.

2.4. Session Parameter Manipulations

A session may have additional state or functionality that effects how the server or client treats the session, content, how it functions, or feedback on how well the session works. Such extensions are not defined in this specification, but may be done in various extensions. RTSP has two methods for retrieving and setting parameter values on either the client or the server: GET_PARAMETER (Section 13.8) and SET_PARAMETER (Section 13.9). These methods carry the parameters in a message body of the appropriate format. One can also use headers to query state with the GET_PARAMETER method. As an example, clients needing to know the current media-range for a time-progressing session can use the GET_PARAMETER method and include the media-range. Furthermore, synchronization information can be requested by using a combination of RTP-Info and Range.

RTSP 2.0 does not have a strong mechanism for providing negotiation of which headers, or parameters and their formats, that can be used. However, responses will indicate request headers or parameters that are not supported. A priori determination of what features are available needs to be done through out-of-band mechanisms, like the session description, or through the usage of feature tags (Section 4.7).

2.5. Media Delivery

The delivery of media to the RTSP client is done with a protocol outside of RTSP and this protocol is determined during the session establishment. This document specifies how media is delivered with RTP over UDP, TCP or the RTSP control connection. Additional protocols may be specified in the future based on demand.

The usage of RTP as media delivery protocol requires some additional information to function well. The PLAY response contains information to enable reliable and timely delivery of how a client should synchronize different sources in the different RTP sessions. It also provides a mapping between RTP timestamps and the content time scale. When the server wants to notify the client about the completion of the media delivery, it sends a PLAY_NOTIFY request to the client. The PLAY_NOTIFY request includes information about the stream end, including the last RTP sequence number for each stream, thus enabling the client to empty the buffer smoothly.

2.5.1. Media Delivery Manipulations

The basic playback functionality of RTSP enables delivery of a range of requested content to the client at the pace intended by the content's creator. However, RTSP can also manipulate the delivery to the client in two ways.

Scale: The ratio of media content time delivered per unit playback time.

Speed: The ratio of playback time delivered per unit of wallclock time.

Both affect the media delivery per time unit. However, they manipulate two independent time scales and the effects are possible to combine.

Scale is used for fast forward or slow motion control as it changes the amount of content timescale that should be played back per time unit. Scale > 1.0, means fast forward, e.g. Scale=2.0 results in that 2 seconds of content is played back every second of playback. Scale = 1.0 is the default value that is used if no Scale is specified, i.e., playback at the content's original rate. Scale values between 0 and 1.0 is providing for slow motion. Scale can be negative to allow for reverse playback in either regular pace (Scale = -1.0) or fast backwards (Scale < -1.0) or slow motion backwards (-1.0 < Scale < 0). Scale = 0 is equal to pause and is not allowed.

In most cases the realization of scale means server side manipulation of the media to ensure that the client can actually play it back. These media manipulation and when they are needed are highly media-type dependent. Let's consider an example with two common media types audio and video.

It is very difficult to modify the playback rate of audio. A maximum of 10-30% is possible by changing the pitch-rate of speech. Music goes out of tune if one tries to manipulate the playback rate by

resampling it. This is a well known problem and audio is commonly muted or played back in short segments with skips to keep up with the current playback point.

For video it is possible to manipulate the frame rate, although the rendering capabilities are often limited to certain frame rates. Also the allowed bitrates in decoding, the structure used in the encoding and the dependency between frames and other capabilities of the rendering device limits the possible manipulations. Therefore, the basic fast forward capabilities often are implemented by selecting certain subsets of frames.

Due to the media restrictions, the possible scale values are commonly restricted to the set of realizable scale ratios. To enable the clients to select from the possible scale values, RTSP can signal the supported Scale ratios for the content. To support aggregated or dynamic content, where this may change during the ongoing session and dependent on the location within the content, a mechanism for updating the media properties and the currently used scale factor exist.

Speed affects how much of the playback timeline is delivered in a given wallclock period. The default is Speed = 1 which means to deliver at the same rate the media is consumed. Speed > 1 means that the receiver will get content faster than it regularly would consume it. Speed < 1 means that delivery is slower than the regular media rate. Speed values of 0 or lower have no meaning and are not allowed. This mechanism enables two general functionalities. One is client side scale operations, i.e. the client receives all the frames and makes the adjustment to the playback locally. The second is delivery control for buffering of media. By specifying a speed over 1.0 the client can build up the amount of playback time it has present in its buffers to a level that is sufficient for its needs.

A naive implementation of Speed would only affect the transmission schedule of the media and has a clear impact on the needed bandwidth. This would result in the data rate being proportional to the speed factor. Speed = 1.5, i.e., 50% faster than normal delivery, would result in a 50% increase in the data transport rate. If that can be supported or not depends solely on the underlying network path. Scale may also have some impact on the required bandwidth due to the manipulation of the content in the new playback schedule. An example is fast forward where only the independently decodable intra frames are included in the media stream. This usage of solely intra frames increases the data rate significantly compared to a normal sequence with the same number of frames, where most frames are encoded using prediction.

This potential increase of the data rate needs to be handled by the media sender. The client has requested that the media will be delivered in a specific way, which should be honored. However, the media sender cannot ignore if the network path between the sender and the receiver can't handle the resulting media stream. In that case the media stream needs to be adapted to fit the available resources of the path. This can result in a reduced media quality.

The need for bitrate adaptation becomes especially problematic in connection with the Speed semantics. If the goal is to fill up the buffer, the client may not want to do that at the cost of reduced quality. If the client wants to make local playout changes then it may actually require that the requested speed be honored. To resolve this issue, Speed uses a range so that both cases can be supported. The server is requested to use the highest possible speed value within the range which is compatible with the available bandwidth. As long as the server can maintain a speed value within the range it shall not change the media quality, but instead modify the actual delivery rate in response to available bandwidth and reflect this in the Speed value in the response. However, if this is not possible, the server should instead modify the media quality to respect the lowest speed value and the available bandwidth.

This functionality enables the local scaling implementation to use a tight range, or even a range where the lower bound equals the upper bound, to identify that it requires the server to deliver the requested amount of media time per delivery time independent of how much it needs to adapt the media quality to fit within the available path bandwidth. For buffer filling, it is suitable to use a range with a reasonable span and with a lower bound at the nominal media rate 1.0, such as 1.0 - 2.5. If the client wants to reduce the buffer, it can specify an upper bound that is below 1.0 to force the server to deliver slower than the nominal media rate.

2.6. Session Maintenance and Termination

The session context that has been established is kept alive by having the client show liveness. This is done in two main ways:

- o Media transport protocol keep-alive. RTCP may be used when using RTP.
- o Any RTSP request referencing the session context.

Section 10.5 discusses the methods for showing liveness in more depth. If the client fails to show liveness for more than the established session timeout value (normally 60 seconds), the server may terminate the context. Other values may be selected by the

server through the inclusion of the timeout parameter in the session header.

The session context is normally terminated by the client sending a TEARDOWN request to the server referencing the aggregated control URI. An individual media resource can be removed from a session context by a TEARDOWN request referencing that particular media resource. If all media resources are removed from a session context, the session context is terminated.

A client may keep the session alive indefinitely if allowed by the server; however, it is recommended to release the session context when an extended period of time without media delivery activity has passed. The client can re-establish the session context if required later. What constitutes an extended period of time is dependent on the server and its usage. It is recommended that the client terminates the session before 10*times the session timeout value has passed. A server may terminate the session after one session timeout period without any client activity beyond keep-alive. When a server terminates the session context, it does that by sending a TEARDOWN request indicating the reason.

A server can also request that the client tear down the session and re-establish it at an alternative server, as may be needed for maintenance. This is done by using the REDIRECT method. The Terminate-Reason header is used to indicate when and why. The Location header indicates where it should connect if there is an alternative server available. When the deadline expires, the server simply stops providing the service. To achieve a clean closure, the client needs to initiate session termination prior to the deadline. In case the server has no other server to redirect to, and wants to close the session for maintenance, it shall use the TEARDOWN method with a Terminate-Reason header.

2.7. Extending RTSP

RTSP is quite a versatile protocol which supports extensions in many different directions. Even this core specification contains several blocks of functionality that are optional to implement. The use case and need for the protocol deployment should determine what parts are implemented. Allowing for extensions makes it possible for RTSP to reach out to additional use cases. However, extensions will affect the interoperability of the protocol and therefore it is important that they can be added in a structured way.

The client can learn the capability of a server by using the OPTIONS method (Section 13.1) and the Supported header (Section 16.49). It can also try and possibly fail using new methods, or require that

particular features are supported using the Require or Proxy-Require header.

The RTSP protocol in itself can be extended in three ways, listed here in order of the magnitude of changes supported:

- o Existing methods can be extended with new parameters, for example, headers, as long as these parameters can be safely ignored by the recipient. If the client needs negative acknowledgment when a method extension is not supported, a tag corresponding to the extension may be added in the field of the Require or Proxy-Require headers (see Section 16.35).
- o New methods can be added. If the recipient of the message does not understand the request, it must respond with error code 501 (Not Implemented) so that the sender can avoid using this method again. A client may also use the OPTIONS method to inquire about methods supported by the server. The server must list the methods it supports using the Public response header.
- o A new version of the protocol can be defined, allowing almost all aspects (except the position of the protocol version number) to change. A new version of the protocol must be registered through an IETF standard track document.

The basic capability discovery mechanism can be used to both discover support for a certain feature and to ensure that a feature is available when performing a request. For a detailed explanation of this see Section 11.

New media delivery protocols may be added and negotiated at session establishment, in addition to extensions to the core protocol. Certain types of protocol manipulations can be done through parameter formats using SET_PARAMETER and GET_PARAMETER.

3. Document Conventions

3.1. Notational Conventions

Since a few of the definitions are identical to HTTP/1.1, this specification only points to the section where they are defined rather than copying it. For brevity, [HX.Y] is to be taken to refer to Section X.Y of the current HTTP/1.1 specification ([RFC2616]).

All the mechanisms specified in this document are described in both prose and the Augmented Backus-Naur form (ABNF) described in detail in [RFC5234].

Indented and smaller-type paragraphs are used to provide informative background and motivation. This is intended to give readers who were not involved with the formulation of the specification an understanding of why things are the way they are in RTSP.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The word, "unspecified" is used to indicate functionality or features that are not defined in this specification. Such functionality cannot be used in a standardized manner without further definition in an extension specification to RTSP.

3.2. Terminology

Aggregate control: The concept of controlling multiple streams using a single timeline, generally maintained by the server. A client, for example, uses aggregate control when it issues a single play or pause message to simultaneously control both the audio and video in a movie. A session which is under aggregate control is referred to as an aggregated session.

Aggregate control URI: The URI used in an RTSP request to refer to and control an aggregated session. It normally, but not always, corresponds to the presentation URI specified in the session description. See Section 13.3 for more information.

Client: The client requests media service from the media server.

Connection: A transport layer virtual circuit established between two programs for the purpose of communication.

Container file: A file which may contain multiple media streams which often constitutes a presentation when played together. The concept of a container file is not embedded in the protocol. However, RTSP servers may offer aggregate control on the media streams within these files.

Continuous media: Data where there is a timing relationship between source and sink; that is, the sink needs to reproduce the timing relationship that existed at the source. The most common examples of continuous media are audio and motion video. Continuous media can be real-time (interactive or conversational), where there is a "tight" timing relationship between source and sink, or streaming where the relationship is less strict.

Feature-tag: A tag representing a certain set of functionality, i.e. a feature.

IRI: Internationalized Resource Identifier, is the same as an URI, with the exception that it allows characters from the whole Universal Character Set (Unicode/ISO 10646), rather than the US-ASCII only. See [RFC3987] for more information.

Live: Normally used to describe a presentation or session with media coming from an ongoing event. This generally results in the session having an unbound or only loosely defined duration, and sometimes no seek operations are possible.

Media initialization: Datatype/codecs specific initialization. This includes such things as clock rates, color tables, etc. Any transport-independent information which is required by a client for playback of a media stream occurs in the media initialization phase of stream setup.

Media parameter: Parameter specific to a media type that may be changed before or during stream delivery.

Media server: The server providing media delivery services for one or more media streams. Different media streams within a presentation may originate from different media servers. A media server may reside on the same host or on a different host from which the presentation is invoked.

(Media) stream: A single media instance, e.g., an audio stream or a video stream as well as a single whiteboard or shared application group. When using RTP, a stream consists of all RTP and RTCP packets created by a source within an RTP session.

Message: The basic unit of RTSP communication, consisting of a structured sequence of octets matching the syntax defined in Section 20 and transmitted over a connection or a connectionless transport. A message is either a Request or a Response.

Message Body: The information transferred as the payload of a message (Request and response). A message body consists of meta-information in the form of message-body headers and content in the form of a message-body, as described in Section 9.

Non-Aggregated Control: Control of a single media stream.

Presentation: A set of one or more streams presented to the client as a complete media feed and described by a presentation description as defined below. Presentations with more than one media stream are often handled in RTSP under aggregate control.

Presentation description: A presentation description contains information about one or more media streams within a presentation, such as the set of encodings, network addresses and information about the content. Other IETF protocols such as SDP ([RFC4566]) use the term "session" for a presentation. The presentation description may take several different formats, including but not limited to the session description protocol format, SDP.

Response: An RTSP response to a Request. One type of RTSP message. If an HTTP response is meant, it is indicated explicitly.

Request: An RTSP request. One type of RTSP message. If an HTTP request is meant, it is indicated explicitly.

Request-URI: The URI used in a request to indicate the resource on which the request is to be performed.

RTSP agent: Refers to either an RTSP client, an RTSP server, or an RTSP proxy. In this specification, there are many capabilities that are common to these three entities such as the capability to send requests or receive responses. This term will be used when describing functionality that is applicable to all three of these entities.

RTSP session: A stateful abstraction upon which the main control methods of RTSP operate. An RTSP session is a common context; it is created and maintained on client's request and can be destroyed by either the client or server. It is established by an RTSP server upon the completion of a successful SETUP request (when a 200 OK response is sent) and is labeled with a session identifier at that time. The session exists until timed out by the server or

explicitly removed by a TEARDOWN request. An RTSP session is a stateful entity; an RTSP server maintains an explicit session state machine (see Appendix B) where most state transitions are triggered by client requests. The existence of a session implies the existence of state about the session's media streams and their respective transport mechanisms. A given session can have one or more media streams associated with it. An RTSP server uses the session to aggregate control over multiple media streams.

Origin Server: The server on which a given resource resides.

Transport initialization: The negotiation of transport information (e.g., port numbers, transport protocols) between the client and the server.

URI: Universal Resource Identifier, see [RFC3986]. The URIs used in RTSP are generally URLs as they give a location for the resource. As URLs are a subset of URIs, they will be referred to as URIs to cover also the cases when an RTSP URI would not be an URL.

URL: Universal Resource Locator, is an URI which identifies the resource through its primary access mechanism, rather than identifying the resource by name or by some other attribute(s) of that resource.

4. Protocol Parameters

4.1. RTSP Version

This specification defines version 2.0 of RTSP.

RTSP uses a "<major>.<minor>" numbering scheme to indicate versions of the protocol. The protocol versioning policy is intended to allow the sender to indicate the format of a message and its capacity for understanding further RTSP communication, rather than the features obtained via that communication. No change is made to the version number for the addition of message components which do not affect communication behavior or which only add to extensible field values.

The <minor> number is incremented when the changes made to the protocol add features which do not change the general message parsing algorithm, but which may add to the message semantics and imply additional capabilities of the sender. The <major> number is incremented when the format of a message within the protocol is changed. The version of an RTSP message is indicated by an RTSP-Version field in the first line of the message. Note that the major and minor numbers MUST be treated as separate integers and that each MAY be incremented higher than a single digit. Thus, RTSP/2.4 is a lower version than RTSP/2.13, which in turn is lower than RTSP/12.3. Leading zeros MUST be ignored by recipients and MUST NOT be sent.

4.2. RTSP IRI and URI

RTSP 2.0 defines and registers three URI schemes "rtsp", "rtsp" and "rtspu". The usage of the last, "rtspu", is unspecified in RTSP 2.0, and is defined here to register and reserve the URI scheme that is defined in RTSP 1.0. The "rtspu" scheme indicates unspecified transport of the RTSP messages over unreliable transport (UDP in RTSP 1.0). An RTSP server MUST response with an error code indicating the "rtspu" scheme is not implemented (501) to a request that carries a "rtspu" URI scheme. The details of the syntax of "rtsp" and "rtsp" URIs has been changed from RTSP 1.0.

This specification also defines the format of the RTSP IRI [RFC3987] that can be used as RTSP resource identifiers and locators, in web pages, user interfaces, on paper, etc. However, the RTSP request message format only allows usage of the absolute URI format. The RTSP IRI format MUST use the rules and transformation for IRIs to URIs, as defined in [RFC3987]. This way RTSP 2.0 URIs for request can be produced from an RTSP IRI.

The RTSP IRI and URI are both syntax restricted compared to the generic syntax defined in [RFC3986] and [RFC3987]:

- o An absolute URI requires the authority part; i.e., a host identity must be provided.
- o Parameters in the path element are prefixed with the reserved separator ";".

The RTSP URI and IRI are case sensitive, with the exception of those parts that [RFC3986] and [RFC3987] defines as case-insensitive; for example, the scheme and host part.

The fragment identifier is used as defined in sections 3.5 and 4.3 of [RFC3986], i.e. the fragment is to be stripped from the IRI by the requester and not included in the request URI. The user agent needs to interpret the value of the fragment based on the media type the request relates to; i.e., the media type indicated in Content-Type header in the response to DESCRIBE.

The syntax of any URI query string is unspecified and responder (usually the server) specific. The query is, from the requester's perspective, an opaque string and needs to be handled as such. Please note that relative URI with queries are difficult to handle due to the RFC 3986 relative URI handling rules. Any change of the path element using a relative URI results in the stripping of the query, which means the relative part needs to contain the query.

The URI scheme "rtsp" requires that commands are issued via a reliable protocol (within the Internet, TCP), while the scheme "rtsp" identifies a reliable transport using secure transport (TLS [RFC5246], see (Section 19).

For the scheme "rtsp", if no port number is provided in the authority part of the URI port number 554 MUST be used. For the scheme "rtsp", the TCP port 322 is registered and MUST be assumed.

A presentation or a stream is identified by a textual media identifier, using the character set and escape conventions of URIs [RFC3986]. URIs may refer to a stream or an aggregate of streams; i.e., a presentation. Accordingly, requests described in (Section 13) can apply to either the whole presentation or an individual stream within the presentation. Note that some request methods can only be applied to streams, not presentations, and vice versa.

For example, the RTSP URI:

```
rtsp://media.example.com:554/twister/audiotrack
```

may identify the audio stream within the presentation "twister", which can be controlled via RTSP requests issued over a TCP connection to port 554 of host media.example.com.

Also, the RTSP URI:

```
rtsp://media.example.com:554/twister
```

identifies the presentation "twister", which may be composed of audio and video streams, but could also be something else like a random media redirector.

This does not imply a standard way to reference streams in URIs. The presentation description defines the hierarchical relationships in the presentation and the URIs for the individual streams. A presentation description may name a stream "a.mov" and the whole presentation "b.mov".

The path components of the RTSP URI are opaque to the client and do not imply any particular file system structure for the server.

This decoupling also allows presentation descriptions to be used with non-RTSP media control protocols simply by replacing the scheme in the URI.

4.3. Session Identifiers

Session identifiers are strings of length 8-128 characters. A session identifier **MUST** be chosen cryptographically random (see [RFC4086]). It is **RECOMMENDED** that it contains 128 bits of entropy, i.e. approximately 22 characters from a high quality generator (see Section 21). However, note that the session identifier does not provide any security against session hijacking unless it is kept confidential by the client, server and trusted proxies.

4.4. SMPTE Relative Timestamps

A SMPTE relative timestamp expresses time relative to the start of the clip. Relative timestamps are expressed as SMPTE time codes for frame-level access accuracy. The time code has the format

```
hours:minutes:seconds:frames.subframes,
```

with the origin at the start of the clip. The default SMPTE format is "SMPTE 30 drop" format, with frame rate is 29.97 frames per second. Other SMPTE codes **MAY** be supported (such as "SMPTE 25")

through the use of "smpte-type". For SMPTE 30, the "frames" field in the time value can assume the values 0 through 29. The difference between 30 and 29.97 frames per second is handled by dropping the first two frame indices (values 00 and 01) of every minute, except every tenth minute. If the frame and the subframe values are zero, they may be omitted. Subframes are measured in one-hundredth of a frame.

Examples:

```
smpte=10:12:33:20-
smpte=10:07:33-
smpte=10:07:00-10:07:33:05.01
smpte-25=10:07:00-10:07:33:05.01
```

4.5. Normal Play Time

Normal play time (NPT) indicates the stream absolute position relative to the beginning of the presentation, not to be confused with the Network Time Protocol (NTP) [RFC5905]. The timestamp consists of two parts: the mandatory first part may be expressed in either seconds or hours, minutes, and seconds. The optional second part consists of a decimal point and decimal figures and indicates fractions of a second.

The beginning of a presentation corresponds to 0.0 seconds. Negative values are not defined.

The special constant "now" is defined as the current instant of a live event. It MAY only be used for live events, and MUST NOT be used for on-demand (i.e., non-live) content.

NPT is defined as in DSM-CC [ISO.13818-6.1995]: "Intuitively, NPT is the clock the viewer associates with a program. It is often digitally displayed on a VCR. NPT advances normally when in normal play mode (scale = 1), advances at a faster rate when in fast scan forward (high positive scale ratio), decrements when in scan reverse (negative scale ratio) and is fixed in pause mode. NPT is (logically) equivalent to SMPTE time codes."

Examples:

```
npt=123.45-125
npt=12:05:35.3-
npt=now-
```

The syntax conforms to ISO 8601 [ISO.8601.2000]. The npt-sec notation is optimized for automatic generation, the npt-hhmmss notation for consumption by human readers. The "now" constant allows clients to request to receive the live feed rather than the stored or time-delayed version. This is needed since neither absolute time nor zero time are appropriate for this case.

4.6. Absolute Time

Absolute time is expressed as ISO 8601 [ISO.8601.2000] timestamps, using UTC (GMT). Fractions of a second may be indicated.

Example for November 8, 1996 at 14h 37 min and 20 and a quarter seconds UTC:

19961108T143720.25Z

4.7. Feature-Tags

Feature-tags are unique identifiers used to designate features in RTSP. These tags are used in Require (Section 16.41), Proxy-Require (Section 16.35), Proxy-Supported (Section 16.36), Supported (Section 16.49) and Unsupported (Section 16.53) header fields.

A feature-tag definition **MUST** indicate which combination of clients, servers or proxies it applies to.

The creator of a new RTSP feature-tag should either prefix the feature-tag with a reverse domain name (e.g., "com.example.mynewfeature" is an apt name for a feature whose inventor can be reached at "example.com"), or register the new feature-tag with the Internet Assigned Numbers Authority (IANA) (see IANA Section 22).

The usage of feature-tags is further described in Section 11 that deals with capability handling.

4.8. Message Body Tags

Message body tags are opaque strings that are used to compare two message bodies from the same resource, for example in caches or to optimize setup after a redirect. Message body tags can be carried in the MTag header (see Section 16.30) or in SDP (see Appendix D.1.9). MTag is similar to ETag in HTTP/1.1.

A message body tag **MUST** be unique across all versions of all message bodies associated with a particular resource. A given message body tag value **MAY** be used for message bodies obtained by requests on

different URIs. The use of the same message body tag value in conjunction with message bodies obtained by requests on different URIs does not imply the equivalence of those message bodies

Message body tags are used in RTSP to make some methods conditional. The methods are made conditional through the inclusion of headers; see "If-Match" (Section 16.23) and "If-None-Match" (Section 16.25). Note that RTSP message body tags apply to the complete presentation; i.e., both the presentation description and the individual media streams. Thus message body tags can be used to verify at setup time after a redirect that the same session description applies to the media at the new location using the If-Match header.

4.9. Media Properties

When an RTSP server handles media, it is important to consider the different properties a media instance for delivery and playback can have. This specification considers the below listed media properties in its protocol operations. They are derived from the differences between a number of supported usages.

On-demand: Media that has a fixed (given) duration that doesn't change during the life time of the RTSP session and is known at the time of the creation of the session. It is expected that the content of the media will not change, even if the representation, i.e encoding, quality, etc, may change. Generally one can seek, i.e. request any range, within the media.

Dynamic On-demand: This is a variation of the on-demand case where external methods are used to manipulate the actual content of the media setup for the RTSP session. The main example is a content defined by a playlist.

Live: Live media represents a progressing content stream (such as broadcast TV) where the duration may or may not be known. It is not seekable, only the content presently being delivered can be accessed.

Live with Recording: A Live stream that is combined with a server-side capability to store and retain the content of the live session, and allow for random access delivery within the part of the already recorded content. The actual behavior of the media stream is very much dependent on the retention policy for the media stream; either the server will be able to capture the complete media stream, or it will have a limitation in how much will be retained. The media range will dynamically change as the session progress. For servers with a limited amount of storage available for recording, there will typically be a sliding window

that moves forwards while new data is made available and older data is discarded.

To cover the above usages, the following media properties with appropriate values are specified:

4.9.1. Random Access and Seeking

Random Access is the ability to specify and get media delivered from any point inside the content, an operation called seeking. This possibility is signaled using the Seek-Style header (see Section 16.45) which can take the following different values:

Random Access: The media is seekable to any out of a large number of points within the media. Due to media encoding limitations, a particular point may not be reachable, but seeking to a point close by is enabled. A floating point number of seconds may be provided to express the worst case distance between random access points.

Conditional Random Access: Based on the above Random Access but intended to handle a case where the distance in the media between random access points is large, and where small seek forward using Random Access would move the client further away than the current point.

Return To Start: Seeking is only possible to the beginning of the content.

No seeking: Seeking is not possible at all.

4.9.2. Retention

Media may have different retention policies in place that affect the operation on media. The following different media retention policies are envisioned and taken into consideration where applicable:

Unlimited: The media will not be removed as long as the RTSP session is in existence.

Time Limited: The media will not be removed before given wallclock time. After that time it may or may not be available any more.

Duration limited: Each individual unit of the media will be retained for the specified duration.

4.9.3. Content Modifications

There is also the question of how the content may change during time for a given media resource:

Immutable: The content of the media will not change, even if the representation, i.e., encoding, quality, etc., may change.

Dynamic: Between explicit updates the media content will not change, but the content may change due to external methods or triggers, such as playlists.

Time Progressing: As times progresses new content will become available. If the content also is retained it will become longer as everything between the start point and the point currently being made available can be accessed. If the media server uses a sliding window policy for retention, the start point will also change as time progresses.

4.9.4. Supported Scale Factors

Content often supports only a limited set or range of scales when delivering the media.. To enable the client to know what values or ranges of scale operations that the whole content or the current position supports, a media properties attribute for this is defined which contains a list with the values and/or ranges that are supported. The attribute is named "Scales". It may be updated at any point in the content due to content consisting of spliced pieces or content being dynamically updated by out-of-band mechanisms.

4.9.5. Mapping to the Attributes

This section shows examples of how one would map the above usages to the properties and their values.

On-demand: Random Access: Random Access=5s, Content Modifications: Immutable, Retention: unlimited or time limited.

Dynamic On-demand: Random Access: Random Access=3s, Content Modifications: Dynamic, Retention: unlimited or time limited.

Live: Random Access: No seeking, Content Modifications: Time Progressing, Retention: Duration limited=0.0s

Live with Recording: Random Access: Random Access=3s, Content Modifications: Time Progressing, Retention: Duration limited=2H

5. RTSP Message

RTSP is a text-based protocol and uses the ISO 10646 character set in UTF-8 encoding RFC 3629 [RFC3629]. Lines MUST be terminated by CRLF.

Text-based protocols make it easier to add optional parameters in a self-describing manner. Since the number of parameters and the frequency of commands is low, processing efficiency is not a concern. Text-based protocols, if done carefully, also allow easy implementation of research prototypes in scripting languages such as TCL, Visual Basic and Perl.

The ISO 10646 character set avoids tricky character set switching, but is invisible to the application as long as US-ASCII is being used. This is also the encoding used for RTCP [RFC3550].

Requests contain methods, the object the method is operating upon and parameters to further describe the method. Methods are idempotent unless otherwise noted. Methods are also designed to require little or no state maintenance at the media server.

5.1. Message Types

RTSP messages consist of requests from client to server, or server to client, and responses in the reverse direction. Request (Section 7) and Response (Section 8) messages use a format based on the generic message format of RFC 2822 [RFC2822] for transferring bodies (the payload of the message). Both types of messages consist of a start-line, zero or more header fields (also known as "headers"), an empty line (i.e., a line with nothing preceding the CRLF) indicating the end of the header, and possibly the data of the message body.

```
generic-message = start-line
                  *(message-header CRLF)
                  CRLF
                  [ message-body-data ]
start-line = Request-Line | Status-Line
```

In the interest of robustness, agents MUST ignore any empty line(s) received where a Request-Line or Response-Line is expected. In other words, if the agent is reading the protocol stream at the beginning of a message and receives a CRLF first, it MUST ignore the CRLF.

5.2. Message Headers

RTSP header fields (see Section 16) include general-header, request-header, response-header, and message-body header fields.

The order in which header fields with differing field names are received is not significant. However, it is "good practice" to send general-header fields first, followed by request-header or response-header fields, and ending with the Message-body header fields.

Multiple message-header fields with the same field-name MAY be present in a message if and only if the entire field-value for that header field is defined as a comma-separated list. It MUST be possible to combine the multiple header fields into one "field-name: field-value" pair, without changing the semantics of the message, by appending each subsequent field-value to the first, each separated by a comma. The order in which header fields with the same field-name are received is therefore significant to the interpretation of the combined field value, and thus a proxy MUST NOT change the order of these field values when a message is forwarded.

Unknown message headers MUST be ignored (skipping over the header to the next protocol element, and not causing an error) by a RTSP server or client. An RTSP Proxy MUST forward unknown message headers. Message headers defined outside of this specification that are required to be interpreted by the RTSP agent will need to use feature tags (Section 4.7) and include them in the appropriate Require (Section 16.41) or Proxy-Require (Section 16.35) header.

5.3. Message Body

The message body (if any) of an RTSP message is used to carry further information for a particular resource associated with the request or response. An example of a message body is the Session Description Protocol (SDP).

The presence of a message body in either a request or a response MUST be signaled by the inclusion of a Content-Length header (see Section 16.16). A message body MUST NOT be included in a request or response if the specification of the particular method (see Method Definitions (Section 13)) does not allow sending a message body.

5.4. Message Length

When a message body is included in a message, the length of that body is determined by one of the following (in order of precedence):

1. Any response message which MUST NOT include a message body (such as the 1xx, 204, and 304 responses) is always terminated by the first empty line after the header fields, regardless of the message-header fields present in the message. (Note: An empty line is a line with nothing preceding the CRLF.)

2. If a Content-Length header(Section 16.16) is present, its value in bytes represents the length of the message-body. If this header field is not present, a value of zero is assumed.

Unlike an HTTP message, an RTSP message MUST contain a Content-Length header whenever it contains a message body. Note that RTSP does not support the HTTP/1.1 "chunked" transfer coding (see [H3.6.1]).

Given the moderate length of presentation descriptions returned, the server should always be able to determine its length, even if it is generated dynamically, making the chunked transfer encoding unnecessary.

6. General Header Fields

General headers are headers that may be used in both requests and responses. The general headers are listed in Table 1:

Header Name	Defined in Section
Accept-Ranges	Section 16.5
Cache-Control	Section 16.10
Connection	Section 16.11
CSeq	Section 16.19
Date	Section 16.20
Media-Properties	Section 16.28
Media-Range	Section 16.29
Pipelined-Requests	Section 16.32
Proxy-Supported	Section 16.36
RTP-Info	Section 16.43
Seek-Style	Section 16.45
Supported	Section 16.49
Timestamp	Section 16.51
Via	Section 16.56

Table 1: The general headers used in RTSP

7. Request

A request message uses the format outlined below regardless of the direction of a request, client to server or server to client:

- o Request line, containing the method to be applied to the resource, the identifier of the resource, and the protocol version in use;
- o Zero or more Header lines, that can be of the following types: general headers (Section 6), request headers (Section 7.2), or message body headers (Section 9.1);
- o One empty line (CRLF) to indicate the end of the header section;
- o Optionally a message-body, consisting of one or more lines. The length of the message body in bytes is indicated by the Content-Length message header.

7.1. Request Line

The request line provides the key information about the request: what method, on what resources and using which RTSP version. The methods that are defined by this specification are listed in Table 2.

Method	Defined in Section
DESCRIBE	Section 13.2
GET_PARAMETER	Section 13.8
OPTIONS	Section 13.1
PAUSE	Section 13.6
PLAY	Section 13.4
PLAY_NOTIFY	Section 13.5
REDIRECT	Section 13.10
SETUP	Section 13.3
SET_PARAMETER	Section 13.9
TEARDOWN	Section 13.7

Table 2: The RTSP Methods

The syntax of the RTSP request line is the following:

```
<Method> SP <Request-URI> SP <RTSP-Version> CRLF
```

Note: This syntax cannot be freely changed in future versions of RTSP. This line needs to remain parsable by older RTSP implementations since it indicates the RTSP version of the message.

In contrast to HTTP/1.1 [RFC2616], RTSP requests identify the resource through an absolute RTSP URI (including scheme, host, and port) (see Section 4.2) rather than just the absolute path.

HTTP/1.1 requires servers to understand the absolute URI, but clients are supposed to use the Host request header. This is purely needed for backward-compatibility with HTTP/1.0 servers, a consideration that does not apply to RTSP.

An asterisk "*" can be used instead of an absolute URI in the Request-URI part to indicate that the request does not apply to a particular resource, but to the server or proxy itself, and is only allowed when the request method does not necessarily apply to a resource.

For example:

```
OPTIONS * RTSP/2.0
```

An OPTIONS in this form will determine the capabilities of the server or the proxy that first receives the request. If the capability of the specific server needs to be determined, without regard to the capability of an intervening proxy, the server should be addressed explicitly with an absolute URI that contains the server's address.

For example:

```
OPTIONS rtsp://example.com RTSP/2.0
```

7.2. Request Header Fields

The RTSP headers in Table 3 can be included in a request, as request headers, to modify the specifics of the request. Some of these headers may also be used in the response to a request, as response headers, to modify the specifics of a response (Section 8.2).

Header	Defined in Section
Accept	Section 16.1
Accept-Credentials	Section 16.2
Accept-Encoding	Section 16.3
Accept-Language	Section 16.4
Authorization	Section 16.7
Bandwidth	Section 16.8
Blocksize	Section 16.9
From	Section 16.22
If-Match	Section 16.23
If-Modified-Since	Section 16.24
If-None-Match	Section 16.25
Notify-Reason	Section 16.31

Proxy-Require	Section 16.35
Range	Section 16.38
Referrer	Section 16.39
Request-Status	Section 16.40
Require	Section 16.41
Scale	Section 16.44
Session	Section 16.47
Speed	Section 16.48
Supported	Section 16.49
Terminate-Reason	Section 16.50
Transport	Section 16.52
User-Agent	Section 16.54

Table 3: The RTSP request headers

Detailed header definition are provided in Section 16.

New request headers may be defined. If the receiver of the request is required to understand the request header, the request **MUST** include a corresponding feature tag in a Require or Proxy-Require header to ensure the processing of the header.

8. Response

After receiving and interpreting a request message, the recipient responds with an RTSP response message. Normally, there is only one, final, response. Only responses using the response code class lxx, are allowed to send one or more lxx response messages prior to the final response message.

The valid response codes and the methods they can be used with are listed in Table 4.

8.1. Status-Line

The first line of a Response message is the Status-Line, consisting of the protocol version followed by a numeric status code and the textual phrase associated with the status code, with each element separated by SP characters. No CR or LF is allowed except in the final CRLF sequence.

<RTSP-Version> SP <Status-Code> SP <Reason-Phrase> CRLF

8.1.1. Status Code and Reason Phrase

The Status-Code element is a 3-digit integer result code of the attempt to understand and satisfy the request. These codes are fully defined in Section 15. The Reason-Phrase is intended to give a short textual description of the Status-Code. The Status-Code is intended for use by automata and the Reason-Phrase is intended for the human user. The client is not required to examine or display the Reason-Phrase.

The first digit of the Status-Code defines the class of response. The last two digits do not have any categorization role. There are 5 values for the first digit:

1xx: Informational - Request received, continuing process

2xx: Success - The action was successfully received, understood, and accepted

3xx: Redirection - Further action needs to be taken in order to complete the request

4xx: Client Error - The request contains bad syntax or cannot be fulfilled

5xx: Server Error - The server failed to fulfill an apparently valid request

The individual values of the numeric status codes defined for RTSP/2.0, and an example set of corresponding Reason-Phrases, are presented in Table 4. The reason phrases listed here are only recommended; they may be replaced by local equivalents without affecting the protocol. Note that RTSP adopts most HTTP/1.1 [RFC2616] status codes and adds RTSP-specific status codes starting at x50 to avoid conflicts with future HTTP status codes that are desirable to import into RTSP.

RTSP status codes are extensible. RTSP applications are not required to understand the meaning of all registered status codes, though such understanding is obviously desirable. However, applications **MUST** understand the class of any status code, as indicated by the first digit, and treat any unrecognized response as being equivalent to the x00 status code of that class, with the exception that an unrecognized response **MUST NOT** be cached. For example, if an unrecognized status code of 431 is received by the client, it can safely assume that there was something wrong with its request and treat the response as if it had received a 400 status code. In such cases, user agents **SHOULD** present to the user the message body returned with the response, since that message body is likely to include human-readable information which will explain the unusual status.

Code	Reason	Method
100	Continue	all
200	OK	all
301	Moved Permanently	all
302	Found	all
303	reserved	n/a
304	Not Modified	all
305	Use Proxy	all
400	Bad Request	all

401	Unauthorized	all
402	Payment Required	all
403	Forbidden	all
404	Not Found	all
405	Method Not Allowed	all
406	Not Acceptable	all
407	Proxy Authentication Required	all
408	Request Timeout	all
410	Gone	all
411	Length Required	all
412	Precondition Failed	DESCRIBE, SETUP
413	Request Message Body Too Large	all
414	Request-URI Too Long	all
415	Unsupported Media Type	all
451	Parameter Not Understood	SET_PARAMETER, GET_PARAMETER
452	reserved	n/a
453	Not Enough Bandwidth	SETUP
454	Session Not Found	all
455	Method Not Valid In This State	all
456	Header Field Not Valid	all
457	Invalid Range	PLAY, PAUSE
458	Parameter Is Read-Only	SET_PARAMETER
459	Aggregate Operation Not Allowed	all

460	Only Aggregate Operation Allowed	all
461	Unsupported Transport	all
462	Destination Unreachable	all
463	Destination Prohibited	SETUP
464	Data Transport Not Ready Yet	PLAY
465	Notification Reason Unknown	PLAY_NOTIFY
466	Key Management Error	all
470	Connection Authorization Required	all
471	Connection Credentials not accepted	all
472	Failure to establish secure connection	all
500	Internal Server Error	all
501	Not Implemented	all
502	Bad Gateway	all
503	Service Unavailable	all
504	Gateway Timeout	all
505	RTSP Version Not Supported	all
551	Option Not Support	all

Table 4: Status codes and their usage with RTSP methods

8.2. Response Headers

The response-header allows the request recipient to pass additional information about the response which cannot be placed in the Status-Line. This header gives information about the server and about further access to the resource identified by the Request-URI. All

headers currently classified as response headers are listed in Table 5.

Header	Defined in Section
Connection-Credentials	Section 16.12
Location	Section 16.27
MTag	Section 16.30
Proxy-Authenticate	Section 16.33
Public	Section 16.37
Range	Section 16.38
Retry-After	Section 16.42
Scale	Section 16.44
Session	Section 16.47
Server	Section 16.46
Speed	Section 16.48
Transport	Section 16.52
Unsupported	Section 16.53
Vary	Section 16.55
WWW-Authenticate	Section 16.57

Table 5: The RTSP response headers

Response-header names can be extended reliably only in combination with a change in the protocol version. However, the usage of feature-tags in the request allows the responding party to learn the capability of the receiver of the response. A new or experimental header MAY be given the semantics of response-header if all parties in the communication recognize them to be response-header. Unrecognized headers in responses are treated as message-headers and hence MUST be ignored.

9. Message Body

Request and Response messages MAY transfer a message body, if not otherwise restricted by the request method or response status code. The message body consists of the content data itself (see also Section 5.2).

The SET_PARAMETER and GET_PARAMETER request and response, and DESCRIBE response MAY have a message body. All 4xx and 5xx responses MAY also have a message body.

In this section, both sender and recipient refer to either the client or the server, depending on who sends and who receives the message body.

9.1. Message-Body Header Fields

Message-body header fields define meta-information about the content data in the message body. The message-body header fields are listed in Table 6.

Header	Defined in Section
Allow	Section 16.6
Content-Base	Section 16.13
Content-Encoding	Section 16.14
Content-Language	Section 16.15
Content-Length	Section 16.16
Content-Location	Section 16.17
Content-Type	Section 16.18
Expires	Section 16.21
Last-Modified	Section 16.26

Table 6: The RTSP message-body headers

The extension-header mechanism allows additional message-body header fields to be defined without changing the protocol, but these fields cannot be assumed to be recognizable by the recipient. Unrecognized

header fields MUST be ignored by the recipient and forwarded by proxies.

9.2. Message Body

An RTSP message with a message body MUST include the Content-Type and Content-Length headers. When a message body is included with a message, the data type of that content data is determined via the header fields Content-Type and Content-Encoding.

Content-Type specifies the media type of the underlying data. Content-Encoding may be used to indicate any additional content codings applied to the data, usually for the purpose of data compression, that are a property of the requested resource. There is no default encoding.

The Content-Length of a message is the length of the content, measured in bytes.

10. Connections

RTSP requests can be transmitted using the two different connection scenarios listed below:

- o persistent - a transport connection is used for several request/response transactions;
- o transient - a transport connection is used for a single request/response transaction.

RFC 2326 attempted to specify an optional mechanism for transmitting RTSP messages in connectionless mode over a transport protocol such as UDP. However, it was not specified in sufficient detail to allow for interoperable implementations. In an attempt to reduce complexity and scope, and due to lack of interest, RTSP 2.0 does not attempt to define a mechanism for supporting RTSP over UDP or other connectionless transport protocols. A side-effect of this is that RTSP requests **MUST NOT** be sent to multicast groups since no connection can be established with a specific receiver in multicast environments.

Certain RTSP headers, such as the CSeq header (Section 16.19), which may appear to be relevant only to connectionless transport scenarios are still retained and **MUST** be implemented according to the specification. In the case of CSeq, it is quite useful for matching responses to requests if the requests are pipelined (see Section 12). It is also useful in proxies for keeping track of the different requests when aggregating several client requests on a single TCP connection.

10.1. Reliability and Acknowledgements

Since RTSP messages are transmitted using reliable transport protocols, they **MUST NOT** be retransmitted at the RTSP protocol level. Instead, the implementation must rely on the underlying transport to provide reliability. The RTSP implementation may use any indication of reception acknowledgment of the message from the underlying transport protocols to optimize the RTSP behavior.

If both the underlying reliable transport such as TCP and the RTSP application retransmit requests, each packet loss or message loss may result in two retransmissions. The receiver typically cannot take advantage of the application-layer retransmission since the transport stack will not deliver the application-layer retransmission before the first attempt has reached the receiver. If the packet loss is caused by congestion, multiple retransmissions at different layers will exacerbate the

congestion.

Lack of acknowledgment of an RTSP request should be handled within the constraints of the connection timeout considerations described below (Section 10.4).

10.2. Using Connections

A TCP transport can be used for both persistent connections (for several message exchanges) and transient connections (for a single message exchange). Implementations of this specification **MUST** support RTSP over TCP. The scheme of the RTSP URI (Section 4.2) indicates the default port that the server will listen on if the port is not explicitly given.

A server **MUST** handle both persistent and transient connections.

Transient connections facilitate mechanisms for fault tolerance. They also allow for application layer mobility. A server and client pair that support transient connections can survive the loss of a TCP connection; e.g., due to a NAT timeout. When the client has discovered that the TCP connection has been lost, it can set up a new one when there is need to communicate again.

A persistent connection is **RECOMMENDED** to be used for all transactions between the server and client, including messages for multiple RTSP sessions. However, a persistent connection **MAY** be closed after a few message exchanges. For example, a client may use a persistent connection for the initial **SETUP** and **PLAY** message exchanges in a session and then close the connection. Later, when the client wishes to send a new request, such as a **PAUSE** for the session, a new connection would be opened. This connection may either be transient or persistent.

An RTSP agent **SHOULD NOT** have more than one connection to the server at any given point. If a client or proxy handles multiple RTSP sessions on the same server, it **SHOULD** use only one connection for managing those sessions.

This saves connection resources on the server. It also reduces complexity by enabling the server to maintain less state about its sessions and connections.

RTSP allows a server to send requests to a client. However, this can be supported only if a client establishes a persistent connection with the server. In cases where a persistent connection does not exist between a server and its client, due to the lack of a signaling channel the server may be forced to silently discard RTSP messages,

and may even drop an RTSP session without notifying the client. An example of such a case is when the server desires to send a REDIRECT request for an RTSP session to the client but is not able to do so because it cannot reach the client. A server that attempts to send a request to a client that has no connection currently to the server SHOULD discard the request directly, but it MAY queue it for later delivery. However, if the server queues the request it SHOULD when adding additional requests to the queue ensure to remove older requests that are now redundant.

Without a persistent connection between the client and the server, the media server has no reliable way of reaching the client. Because the likely failure of server to client established connections the server will not even attempt establishing any connection.

The sending of client and server requests can be asynchronous events. To avoid deadlock situations both client and server MUST be able to send and receive requests simultaneously. As an RTSP response may be queued up for transmission, reception or processing behind the peer RTSP agent's own requests, all RTSP agents are required to have a certain capability of handling outstanding messages. A potential issue is that outstanding requests may timeout despite them being processed by the peer due to the response is caught in the queue behind a number of request that the RTSP agent is processing but that take some time to complete. To avoid this problem an RTSP agent is recommended to buffer incoming messages locally so that any response messages can be processed immediately upon reception. If responses are separated from requests and directly forwarded for processing, not only can the result be used immediately, the state associated with that outstanding request can also be released. However, buffering a number of requests on the receiving RTSP agent consumes resources and enables a resource exhaustion attack on the agent. Therefore this buffer should be limited so that an unreasonable number of requests or total message size is not allowed to consume the receiving agent's resources. In most APIs, having the receiving agent stop reading from the TCP socket will result in TCP's window being clamped. Thus forcing the buffering onto the sending agent when the load is larger than expected. However, as both RTSP message sizes and frequency may be changed in the future by protocol extensions, an agent should be careful against taking harsher measurements against a potential attack. When under attack an RTSP agent can close TCP connections and release state associated with that TCP connection.

To provide some guidance on what is reasonable the following guidelines are given. It is RECOMMENDED that:

- o an RTSP agent should not have more than 10 outstanding requests per RTSP session;
- o an RTSP agent should not have more than 10 outstanding requests that are not related to an RTSP session or that are requesting to create an RTSP session.

In light of the above, it is RECOMMENDED that clients use persistent connections whenever possible. A client that supports persistent connections MAY "pipeline" its requests (see Section 12).

10.3. Closing Connections

The client MAY close a connection at any point when no outstanding request/response transactions exist for any RTSP session being managed through the connection. The server, however, SHOULD NOT close a connection until all RTSP sessions being managed through the connection have been timed out (Section 16.47). A server SHOULD NOT close a connection immediately after responding to a session-level TEARDOWN request for the last RTSP session being controlled through the connection. Instead, it should wait for a reasonable amount of time for the client to receive the TEARDOWN response, take appropriate action, and initiate the connection closing. The server SHOULD wait at least 10 seconds after sending the TEARDOWN response before closing the connection.

This is to ensure that the client has time to issue a SETUP for a new session on the existing connection after having torn the last one down. 10 seconds should give the client ample opportunity to get its message to the server.

A server SHOULD NOT close the connection directly as a result of responding to a request with an error code.

Certain error responses such as "460 Only Aggregate Operation Allowed" (Section 15.4.25) are used for negotiating capabilities of a server with respect to content or other factors. In such cases, it is inefficient for the server to close a connection on an error response. Also, such behavior would prevent implementation of advanced/special types of requests or result in extra overhead for the client when testing for new features. On the flip side, keeping connections open after sending an error response poses a Denial of Service security risk (Section 21).

The server MAY close a connection if it receives an incomplete message and if the message is not completed within a reasonable amount of time. It is RECOMMENDED that the server waits at least 10 seconds for the completion of a message or for the next part of the

message to arrive (which is an indication that the transport and the client are still alive). Servers believing they are under attack or otherwise starved for resources during that event MAY consider using a shorter timeout.

If a server closes a connection while the client is attempting to send a new request, the client will have to close its current connection, establish a new connection and send its request over the new connection.

An RTSP message SHOULD NOT be terminated by closing the connection. Such a message MAY be considered to be incomplete by the receiver and discarded. An RTSP message is properly terminated as defined in Section 5.

10.4. Timing Out Connections and RTSP Messages

Receivers of a request (responder) SHOULD respond to requests in a timely manner even when a reliable transport such as TCP is used. Similarly, the sender of a request (requester) SHOULD wait for a sufficient time for a response before concluding that the responder will not be acting upon its request.

A responder SHOULD respond to all requests within 5 seconds. If the responder recognizes that processing of a request will take longer than 5 seconds, it SHOULD send a 100 (Continue) response as soon as possible. It SHOULD continue sending a 100 response every 5 seconds thereafter until it is ready to send the final response to the requester. After sending a 100 response, the receiver MUST send a final response indicating the success or failure of the request.

A requester SHOULD wait at least 10 seconds for a response before concluding that the responder will not be responding to its request. After receiving a 100 response, the requester SHOULD continue waiting for further responses. If more than 10 seconds elapses without receiving any response, the requester MAY assume that the responder is unresponsive and abort the connection.

A requester SHOULD wait longer than 10 seconds for a response if it is experiencing significant transport delays on its connection to the responder. The requester is capable of determining the RTT of the request/response cycle using the Timestamp header (Section 16.51) in any RTSP request.

10 seconds was chosen for the following reasons. It gives TCP time to perform a couple of retransmissions, even if operating on default values. It is short enough that users may not abandon the process themselves. However, it should be noted that 10 seconds

can be aggressive on certain type of networks. The 5 seconds value for lxx messages is half the timeout giving a reasonable change of successful delivery before timeout happens on the requester side.

10.5. Showing Liveness

The mechanisms for showing liveness of the client is, any RTSP request with a Session header, if RTP & RTCP is used an RTCP message, or through any other used media protocol capable of indicating liveness of the RTSP client. It is RECOMMENDED that a client does not wait to the last second of the timeout before trying to send a liveness message. The RTSP message may be lost or when using reliable protocols, such as TCP, the message may take some time to arrive safely at the receiver. To show liveness between RTSP request issued to accomplish other things, the following mechanisms can be used, in descending order of preference:

RTCP: If RTP is used for media transport RTCP SHOULD be used. If RTCP is used to report transport statistics, it MUST also work as keep alive. The server can determine the client by network address and port together with the fact that the client is reporting on the servers SSRC(s). A downside of using RTCP is that it only gives statistical guarantees to reach the server. However, the probability of a false client timeout is so low that it can be ignored in most cases. For example, assume a session with 60 seconds timeout and enough bitrate assigned to RTCP messages to send a message from client to server on average every 5 seconds. That client has, for a network with 5 % packet loss, the probability to fail showing liveness sign in that session within the timeout interval of 2.4×10^{-16} . Sessions with shorter timeouts, or much higher packet loss, or small RTCP bandwidths SHOULD also use any of the mechanisms below.

SET_PARAMETER: When using SET_PARAMETER for keep alive, no body SHOULD be included. This method is the RECOMMENDED RTSP method to use for a request intended only to perform keep-alive.

GET_PARAMETER: When using GET_PARAMETER for keep alive, no body SHOULD be included.

OPTIONS: This method is also usable, but it causes the server to perform more unnecessary processing and results in bigger responses than necessary for the task. The reason is that the server needs to determine the capabilities associated with the media resource to correctly populate the Public and Allow headers.

The timeout parameter MAY be included in a SETUP response, and MUST NOT be included in requests. The server uses it to indicate to the client how long the server is prepared to wait between RTSP commands or other signs of life before closing the session due to lack of activity (see Appendix B). The timeout is measured in seconds, with a default of 60 seconds. The length of the session timeout MUST NOT be changed in an established session.

10.6. Use of IPv6

Explicit IPv6 support was not present in RTSP 1.0 (RFC 2326). RTSP 2.0 has been updated for explicit IPv6 support. Implementations of RTSP 2.0 MUST understand literal IPv6 addresses in URIs and headers.

10.7. Overload Control

Overload in RTSP can occur when servers and proxies have insufficient resources to complete the processing of a request. An improper handling of such an overload situation at proxies and servers can impact the operation of the RTSP deployment, and probably worsen the situation. RTSP defines the 503 (Service Unavailable) response (Section 15.5.4) to let servers and proxies notify requesting proxies and RTSP clients about an overload situation. In conjunction with the Retry-After header (Section 16.42) the server or proxy can indicate the time after the requesting entity can send another request to the proxy or server.

Simply implementing and using the 503 (Service Unavailable) is not sufficient for properly handling overload situations. For instance, a simplistic approach would be to send the 503 response with a Retry-After header set to a fixed value. However, this can cause the situation where multiple RTSP clients again send requests to a proxy or server at roughly the same time which may again cause an overload situation, or if the "old" overload situation is not yet solved, i.e., the length indicated in the Retry-After header was too short.

An RTSP server or proxy in an overload situation must select the value of the Retry-After header carefully and in dependency of its current load situation. It is RECOMMENDED to increase the length proportional with the current load of the server, i.e., an increasing workload should result in an increased length of the indicated unavailability. It is RECOMMENDED to not send the same value in the Retry-After header to all requesting proxies and clients, but to add a variation to the mean value of the Retry-After header.

Another issue are load balancing RTSP proxies, i.e., where an RTSP proxy is used to select amongst a set of RTSP servers to handle the requests, or when multiple server addresses are available for a given

server name. The proxy or client may receive a 503 (Service Unavailable) from one of its RTSP servers or a TCP timeout (if the server is even unable to handle the request message). The proxy or client simply retries the other addresses, but may also receive a 503 (Service Unavailable) response or TCP timeouts from those addresses. In such a situation, where none of the RTSP servers/addresses can handle the request, the RTSP agent has to wait before it can send any new requests to the RTSP server. Any additional request to a specific address **MUST** be delayed according to the Retry-After headers received. For addresses where no response was received or TCP timeout occurred, an initial wait timer **SHOULD** be set to 5 seconds. That timer **MUST** be doubled for each additional failure to connect or receive response. It is **RECOMMENDED** to not set the same value in the timer, but to add a variation to the mean value.

11. Capability Handling

This section describes the available capability handling mechanism which allows RTSP to be extended. Extensions to this version of the protocol are basically done in two ways. First, new headers can be added. Secondly, new methods can be added. The capability handling mechanism is designed to handle both cases.

When a method is added, the involved parties can use the OPTIONS method to discover whether it is supported. This is done by issuing an OPTIONS request to the other party. Depending on the URI it will either apply in regards to a certain media resource, the whole server in general, or simply the next hop. The OPTIONS response **MUST** contain a Public header which declares all methods supported for the indicated resource.

It is not necessary to use OPTIONS to discover support of a method, as the client could simply try the method. If the receiver of the request does not support the method it will respond with an error code indicating the method is either not implemented (501) or does not apply for the resource (405). The choice between the two discovery methods depends on the requirements of the service.

Feature-Tags are defined to handle functionality additions that are not new methods. Each feature-tag represents a certain block of functionality. The amount of functionality that a feature-tag represents can vary significantly. A feature-tag can for example represent the functionality a single RTSP header provides. Another feature-tag can represent much more functionality, such as the "play.basic" feature-tag which represents the minimal media delivery for playback implementation.

Feature-tags are used to determine whether the client, server or proxy supports the functionality that is necessary to achieve the desired service. To determine support of a feature-tag, several different headers can be used, each explained below:

Supported: This header is used to determine the complete set of functionality that both client and server have. The intended usage is to determine before one needs to use a functionality that it is supported. It can be used in any method, but OPTIONS is the most suitable one as it at the same time determines all methods that are implemented. When sending a request the requester declares all its capabilities by including all supported feature-tags. This results in the receiver learns the requester's feature support. The receiver then includes its set of features in the response.

Proxy-Supported: This header is used similarly to the Supported header, but instead of giving the supported functionality of the client or server it provides both the requester and the responder a view of what functionality the proxy chain between the two supports. Proxies are required to add this header whenever the Supported header is present, but proxies may also add it independently of the requester.

Require: This header can be included in any request where the end-point, i.e. the client or server, is required to understand the feature to correctly perform the request. This can, for example, be a SETUP request where the server is required to understand a certain parameter to be able to set up the media delivery correctly. Ignoring this parameter would not have the desired effect and is not acceptable. Therefore the end-point receiving a request containing a Require MUST negatively acknowledge any feature that it does not understand and not perform the request. The response in cases where features are not supported are 551 (Option Not Supported). Also the features that are not supported are given in the Unsupported header in the response.

Proxy-Require: This header has the same purpose and workings as Require except that it only applies to proxies and not the end-point. Features that need to be supported by both proxies and end-points need to be included in both the Require and Proxy-Require header.

Unsupported: This header is used in a 551 error response, to indicate which features were not supported. Such a response is only the result of the usage of the Require and/or Proxy-Require header where one or more feature were not supported. This information allows the requester to make the best of situations as it knows which features are not supported.

12. Pipelining Support

Pipelining is a general method to improve performance of request response protocols by allowing the requesting agent to have more than one request outstanding and send them over the same persistent connection. For RTSP, where the relative order of requests will matter, it is important to maintain the order of the requests. Because of this, the responding agent **MUST** process the incoming requests in their sending order. The sending order can be determined by the CSeq header and its sequence number. For TCP the delivery order will be the same as the sending order. The processing of the request **MUST** also have been finished before processing the next request from the same agent. The responses **MUST** be sent in the order the requests were processed.

RTSP 2.0 has extended support for pipelining compared to RTSP 1.0. The major improvement is to allow all requests to setup and initiate media delivery to be pipelined after each other. This is accomplished by the utilization of the Pipelined-Requests header (see Section 16.32). This header allows a client to request that two or more requests are processed in the same RTSP session context which the first request creates. In other words, a client can request that two or more media streams are set-up and then played without needing to wait for a single response. This speeds up the initial startup time for an RTSP session with at least one RTT.

If a pipelined request builds on the successful completion of one or more prior requests the requester must verify that all requests were executed as expected. A common example will be two SETUP requests and a PLAY request. In case one of the SETUP fails unexpectedly, the PLAY request can still be successfully executed. However, the resulting presentation will not be as expected by the requesting client, as only a single media instead of two will be played. In this case the client can send a PAUSE request, correct the failing SETUP request and then request it to be played.

13. Method Definitions

The method indicates what is to be performed on the resource identified by the Request-URI. The method name is case-sensitive. New methods may be defined in the future. Method names MUST NOT start with a \$ character (decimal 36) and MUST be a token as defined by the ABNF [RFC5234] in the syntax chapter Section 20. The methods are summarized in Table 7.

method	direction	object	Server req.	Client req.
DESCRIBE	C -> S	P,S	recommended	recommended
GET_PARAMETER	C -> S	P,S	optional	optional
	S -> C	P,S	optional	optional
OPTIONS	C -> S	P,S	required	required
	S -> C	P,S	optional	optional
PAUSE	C -> S	P,S	required	required
PLAY	C -> S	P,S	required	required
PLAY_NOTIFY	S -> C	P,S	required	required
REDIRECT	S -> C	P,S	optional	required
SETUP	C -> S	S	required	required
SET_PARAMETER	C -> S	P,S	required	optional
	S -> C	P,S	optional	optional
TEARDOWN	C -> S	P,S	required	required
	S -> C	P	required	required

Table 7: Overview of RTSP methods, their direction, and what objects (P: presentation, S: stream) they operate on.

Note on Table 7: GET_PARAMETER is optional. For example, a fully functional server can be built to deliver media without any parameters. SET_PARAMETER is required, however, due to its usage for keep-alive. PAUSE is now required because it is the only way of leaving the Play state without terminating the whole session.

If an RTSP agent does not support a particular method, it MUST return 501 (Not Implemented) and the requesting RTSP agent, in turn, SHOULD NOT try this method again for the given agent / resource combination. An RTSP proxy whose main function is to log or audit and not modify transport or media handling in any way MAY forward RTSP messages with unknown methods. Note that the proxy still needs to perform the minimal required processing, like adding the Via header.

13.1. OPTIONS

The semantics of the RTSP OPTIONS method is similar to that of the HTTP OPTIONS method described in [H9.2]. In RTSP however, OPTIONS is bi-directional, in that a client can request it to a server and vice versa. A client MUST implement the capability to send an OPTIONS request and a server or a proxy MUST implement the capability to respond to an OPTIONS request. The client, server or proxy MAY also implement the converse of their required capability, but still retain the prior mentioned about what is a "MUST implement".

An OPTIONS request may be issued at any time. Such a request does not modify the session state. However, it may prolong the session lifespan (see below). The URI in an OPTIONS request determines the scope of the request and the corresponding response. If the Request-URI refers to a specific media resource on a given host, the scope is limited to the set of methods supported for that media resource by the indicated RTSP agent. A Request-URI with only the host address limits the scope to the specified RTSP agent's general capabilities without regard to any specific media. If the Request-URI is an asterisk ("*"), the scope is limited to the general capabilities of the next hop (i.e. the RTSP agent in direct communication with the request sender).

Regardless of scope of the request, the Public header MUST always be included in the OPTIONS response listing the methods that are supported by the responding RTSP agent. In addition, if the scope of the request is limited to a media resource, the Allow header MUST be included in the response to enumerate the set of methods that are allowed for that resource unless the set of methods completely matches the set in the Public header. If the given resource is not available, the RTSP agent SHOULD return an appropriate response code such as 3rr or 4xx. The Supported header MAY be included in the request to query the set of features that are supported by the

responding RTSP agent.

The OPTIONS method can be used to keep an RTSP session alive. However, this is not the preferred way of session keep-alive signaling, see Section 16.47. An OPTIONS request intended for keeping alive an RTSP session MUST include the Session header with the associated session ID. Such a request SHOULD also use the media or the aggregated control URI as the Request-URI.

Example:

```
C->S: OPTIONS rtsp://server.example.com RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2
      Proxy-Require: gzipped-messages
      Supported: play.basic

S->C: RTSP/2.0 200 OK
      CSeq: 1
      Public: DESCRIBE, SETUP, TEARDOWN, PLAY, PAUSE, OPTIONS
      Supported: play.basic, setup.rtp.rtcp.mux, play.scale
      Server: PhonyServer/1.1
```

Note that some of the feature-tags in Supported and Proxy-Require are fictional features.

13.2. DESCRIBE

The DESCRIBE method is used to retrieve the description of a presentation or media object from a server. The Request-URI of the DESCRIBE request identifies the media resource of interest. The client MAY include the Accept header in the request to list the description formats that it understands. The server MUST respond with a description of the requested resource and return the description in the message body of the response, if the DESCRIBE method request can be successfully fulfilled. The DESCRIBE reply-response pair constitutes the media initialization phase of RTSP.

The DESCRIBE response SHOULD contain all media initialization information for the resource(s) that it describes. Servers SHOULD NOT use the DESCRIBE response as a means of media indirection by having the description point at another server; instead, using the 3rr responses is RECOMMENDED.

By forcing a DESCRIBE response to contain all media initialization information for the set of streams that it describes, and discouraging the use of DESCRIBE for media indirection, any looping problems can be avoided that might have resulted from other approaches.

Example:

```
C->S: DESCRIBE rtsp://server.example.com/fizzle/foo RTSP/2.0
      CSeq: 312
      User-Agent: PhonyClient/1.2
      Accept: application/sdp, application/example
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 312
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Server: PhonyServer/1.1
      Content-Base: rtsp://server.example.com/fizzle/foo/
      Content-Type: application/sdp
      Content-Length: 358
```

```
v=0
o=mhandley 2890844526 2890842807 IN IP4 192.0.2.46
s=SDP Seminar
i=A Seminar on the session description protocol
u=http://www.example.com/lectures/sdp.ps
e=seminar@example.com (Seminar Management)
c=IN IP4 0.0.0.0
a=control:*
t=2873397496 2873404696
m=audio 3456 RTP/AVP 0
a=control:audio
m=video 2232 RTP/AVP 31
a=control:video
```

Media initialization is a requirement for any RTSP-based system, but the RTSP specification does not dictate that this is required to be done via the DESCRIBE method. There are three ways that an RTSP client may receive initialization information:

- o via an RTSP DESCRIBE request
- o via some other protocol (HTTP, email attachment, etc.)
- o via some form of user interface

If a client obtains a valid description from an alternate source, the client MAY use this description for initialization purposes without

issuing a DESCRIBE request for the same media. The client should use any MTag to either validate the presentation description or make the session establishment conditional on being valid.

It is RECOMMENDED that minimal servers support the DESCRIBE method, and highly recommended that minimal clients support the ability to act as "helper applications" that accept a media initialization file from a user interface, and/or other means that are appropriate to the operating environment of the clients.

13.3. SETUP

Note: The states described in this section and the following are described in detail in Appendix B.

The SETUP request for an URI specifies the transport mechanism to be used for the streamed media. The SETUP method may be used in two different cases; Create an RTSP session and change the transport parameters of already set up media stream. SETUP can be used in all three states; Init, and Ready, for both purposes and in PLAY to change the transport parameters. There is also a third possible usage for the SETUP method which is not specified in this memo: adding a media to a session. Using SETUP to add media to an existing session, when the session is in Play state, is unspecified.

The Transport header, see Section 16.52, specifies the media transport parameters acceptable to the client for data transmission; the response will contain the transport parameters selected by the server. This allows the client to enumerate in descending order of preference the transport mechanisms and parameters acceptable to it, while the server can select the most appropriate. It is expected that the session description format used will enable the client to select a limited number of possible configurations that are offered to the server to choose from. All transport related parameters SHALL be included in the Transport header; the use of other headers for this purpose is NOT RECOMMENDED due to middleboxes, such as firewalls or NATs.

For the benefit of any intervening firewalls, a client MUST indicate the known transport parameters, even if it has no influence over these parameters, for example, where the server advertises a fixed multicast address as destination.

Since SETUP includes all transport initialization information, firewalls and other intermediate network devices (which need this information) are spared the more arduous task of parsing the DESCRIBE response, which has been reserved for media initialization.

The client MUST include the Accept-Ranges header in the request indicating all supported unit formats in the Range header. This allows the server to know which formats it may use in future session related responses, such as a PLAY response without any range in the request. If the client does not support a time format necessary for the presentation the server MUST respond using 456 (Header Field Not Valid for Resource) and include the Accept-Ranges header with the range unit formats supported for the resource.

In a SETUP response the server MUST include the Accept-Ranges header (see Section 16.5) to indicate which time formats are acceptable to use for this media resource.

The SETUP response 200 OK MUST include the Media-Properties header (see Section 16.28). The combination of the parameters of the Media-Properties header indicates the nature of the content present in the session (see also Section 4.9). For example, a live stream with time shifting is indicated by

- o Random Access set to Random-Access,
- o Content Modifications set to Time Progressing,
- o Retention set to Time-Duration (with specific recording window time value).

The SETUP response 200 OK MUST include the Media-Range header (see Section 16.29) if the media is Time-Progressing.

A basic example for SETUP:

```

C->S: SETUP rtsp://example.com/foo/bar/baz.rm RTSP/2.0
      CSeq: 302
      Transport: RTP/AVP;unicast;dest_addr=":4588"/":4589",
                RTP/AVP/TCP;unicast;interleaved=0-1
      Accept-Ranges: NPT, UTC
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 302
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Server: PhonyServer/1.1
      Session: 47112344;timeout=60
      Transport: RTP/AVP;unicast;dest_addr="192.0.2.53:4588"/
                "192.0.2.53:4589"; src_addr="198.51.100.241:6256"/
                "198.51.100.241:6257"; ssrc=2A3F93ED
      Accept-Ranges: NPT
      Media-Properties: Random-Access=3.2, Time-Progressing,
                      Time-Duration=3600.0
      Media-Range: npt=0-2893.23

```

In the above example the client wants to create an RTSP session containing the media resource "rtsp://example.com/foo/bar/baz.rm". The transport parameters acceptable to the client are either RTP/AVP/UDP (UDP per default) to be received on client port 4588 and 4589 at the address the RTSP setup connection comes from or RTP/AVP interleaved on the RTSP control channel. The server selects the RTP/AVP/UDP transport and adds the address and ports it will send and receive RTP and RTCP from, and the RTP SSRC that will be used by the server.

The server MUST generate a session identifier in response to a successful SETUP request, unless a SETUP request to a server includes a session identifier or a Pipelined-Requests header referencing an existing session context, in which case the server MUST bundle this setup request into the existing session (aggregated session) or return error 459 (Aggregate Operation Not Allowed) (see Section 15.4.24). An Aggregate control URI MUST be used to control an aggregated session. This URI MUST be different from the stream control URIs of the individual media streams included in the aggregate (see Section 13.4.2 for aggregated sessions and for the particular URIs see Appendix D.1.1). The Aggregate control URI is to be specified by the session description if the server supports aggregated control and aggregated control is desired for the session. However, even if aggregated control is offered the client MAY chose to not set up the session in aggregated control. If an Aggregate control URI is not specified in the session description, it is normally an indication that non-aggregated control should be used. The SETUP of media streams in an aggregate which has not been given

an aggregated control URI is unspecified.

While the session ID sometimes carries enough information for aggregate control of a session, the Aggregate control URI is still important for some methods such as SET_PARAMETER where the control URI enables the resource in question to be easily identified. The Aggregate control URI is also useful for proxies, enabling them to route the request to the appropriate server, and for logging, where it is useful to note the actual resource that a request was operating on.

A session will exist until it is either removed by a TEARDOWN request or is timed-out by the server. The server MAY remove a session that has not demonstrated liveness signs from the client(s) within a certain timeout period. The default timeout value is 60 seconds; the server MAY set this to a different value and indicate so in the timeout field of the Session header in the SETUP response. For further discussion see Section 16.47. Signs of liveness for an RTSP session are:

- o Any RTSP request from a client which includes a Session header with that session's ID.
- o If RTP is used as a transport for the underlying media streams, an RTCP sender or receiver report from the client(s) for any of the media streams in that RTSP session. RTCP Sender Reports may for example be received in sessions where the server is invited into a conference session and is valid for keep-alive.

If a SETUP request on a session fails for any reason, the session state, as well as transport and other parameters for associated streams MUST remain unchanged from their values as if the SETUP request had never been received by the server.

13.3.1. Changing Transport Parameters

A client MAY issue a SETUP request for a stream that is already set up or playing in the session to change transport parameters, which a server MAY allow. If it does not allow changing of parameters, it MUST respond with error 455 (Method Not Valid In This State). The reasons to support changing transport parameters include allowing application layer mobility and flexibility to utilize the best available transport as it becomes available. If a client receives a 455 when trying to change transport parameters while the server is in Play state, it MAY try to put the server in ready state using PAUSE, before issuing the SETUP request again. If that also fails the changing of transport parameters will require that the client performs a TEARDOWN of the affected media and then to set it up

again. For an aggregated session avoiding tearing down all the media at the same time will avoid the creation of a new session.

All transport parameters MAY be changed. However, the primary usage expected is to either change the transport protocol completely, like switching from Interleaved TCP mode to UDP or vice versa, or to change the delivery address.

In a SETUP response for a request to change the transport parameters while in Play state, the server MUST include the Range to indicate at what point the new transport parameters will be used. Further, if RTP is used for delivery, the server MUST also include the RTP-Info header to indicate at what timestamp and RTP sequence number the change will take place. If both RTP-Info and Range are included in the response the "rtp_time" parameter and start point in the Range header MUST be for the corresponding time, i.e. be used in the same way as for PLAY to ensure the correct synchronization information is available.

If the transport parameters change while in Play state results in a change of synchronization related information, for example changing RTP SSRC, the server MUST provide in the SETUP response the necessary synchronization information. However, the server is RECOMMENDED to avoid changing the synchronization information if possible.

13.4. PLAY

This section describes the usage of the PLAY method in general, for aggregated sessions, and in different usage scenarios.

13.4.1. General Usage

The PLAY method tells the server to start sending data via the mechanism specified in SETUP and which part of the media should be played out. PLAY requests are valid when the session is in Ready or Play states. A PLAY request MUST include a Session header to indicate which session the request applies to.

Upon receipt of the PLAY request, the server MUST position the normal play time to the beginning of the range specified in the received Range header and deliver stream data until the end of the range if given, until a new PLAY request is received, or until the end of the media is reached. If no Range header is present in the PLAY request the server SHALL play from current pause point until the end of media. The pause point defaults at session start to the beginning of the media. For media that is time-progressing and has no retention, the pause point will always be set equal to NPT "now", i.e., the current delivery point. The pause point may also be set to a

particular point in the media by the PAUSE method, see Section 13.6. The pause point for media that is currently playing is equal to the current media position. For time-progressing media with time-limited retention, if the pause point represents a position that is older than what is retained by the server, the pause point will be moved to the oldest retained.

What range values are valid depends on the type of content. For content that isn't time progressing the range value is valid if the given range is part of any media within the aggregate. In other words the valid media range for the aggregate is the union of all of the media components in the aggregate. If a given range value points outside of the media, the response MUST be the 457 (Invalid Range) error code and include the Media-Range header (Section 16.29) with the valid range for the media. Except for time progressing content where the client requests a start point prior to what is retained, the start point is adjusted to the oldest retained content. For a start point that is beyond the media front edge, i.e. beyond the current value for "now", the server SHALL adjust the start value to the current front edge. The Range header's stop point value may point beyond the current media edge. In that case, the server SHALL deliver media from the requested (and possibly adjusted) start point until the provided stop point, or the end of the media is reached prior to the specified stop point. Please note that if one simply wants to play from a particular start point until the end of media using a Range header with an implicit stop point is RECOMMENDED.

If a client requests to start playing at the end of media, either explicitly with a Range header or implicitly with a pause point that is at the end of media, a 457 (Invalid Range) error MUST be sent and include the Media-Range header (Section 16.29). It is specified below that the Range header also must be included in the response and that it will carry the pause point in the media, in the case of the session being in Ready State. Note that this also applies if the pause point or requested start point is at the beginning of the media and a Scale header (Section 16.44) is included with a negative value (playing backwards).

For media with random access properties a client may express its preference on which policy for start point selection the server shall use. This is done by including the Seek-Style header (Section 16.45) in the PLAY request. The Seek-Style applied will effect the content of the Range header as it will be adjusted to indicate from what point the media actually is delivered.

A client desiring to play the media from the beginning MUST send a PLAY request with a Range header pointing at the beginning, e.g. npt=0-. If a PLAY request is received without a Range header and

media delivery has stopped at the end, the server SHOULD respond with a 457 "Invalid Range" error response. In that response, the current pause point MUST be included in a Range header.

All range specifiers in this specification allow for ranges with an implicit start point (e.g. "npt=-30"). When used in a PLAY request, the server treats this as a request to start or resume delivery from the current pause point, ending at the end time specified in the Range header. If the pause point is located later than the given end value, a 457 (Invalid Range) response MUST be given.

The example below will play seconds 10 through 25. It also requests the server to deliver media from the first Random Access Point prior to the indicated start point.

```
C->S: PLAY rtsp://audio.example.com/audio RTSP/2.0
      CSeq: 835
      Session: 12345678
      Range: npt=10-25
      Seek-Style: RAP
      User-Agent: PhonyClient/1.2
```

Servers MUST include a "Range" header in any PLAY response, even if no Range header was present in the request. The response MUST use the same format as the request's range header contained. If no Range header was in the request, the format used in any previous PLAY request within the session SHOULD be used. If no format has been indicated in a previous request the server MAY use any time format supported by the media and indicated in the Accept-Ranges header in the SETUP request. It is RECOMMENDED that NPT is used if supported by the media.

For any error response to a PLAY request, the server's response depends on the current session state. If the session is in Ready state, the current pause-point is returned using Range header with the pause point as the explicit start-point and an implicit stop-point. For time-progressing content where the pause-point moves with real-time due to limited retention, the current pause point is returned. For sessions in Play state, the current playout point and the remaining parts of the range request is returned. For any media with retention longer than 0 seconds the currently valid Media-Range header SHALL also be included in the response.

A PLAY response MAY include a header carrying synchronization information. As the information necessary is dependent on the media transport format, further rules specifying the header and its usage are needed. For RTP the RTP-Info header is specified, see Section 16.43, and used in the following example.

Here is a simple example for a single audio stream where the client requests the media starting from 3.52 seconds and to the end. The server sends a 200 OK response with the actual play time which is 10 ms prior (3.51) and the RTP-Info header that contains the necessary parameters for the RTP stack.

```
C->S: PLAY rtsp://example.com/audio RTSP/2.0
      CSeq: 836
      Session: 12345678
      Range: npt=3.52-
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 836
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Server: PhonyServer/1.0
      Range: npt=3.51-324.39
      Seek-Style: First-Prior
      RTP-Info:url="rtsp://example.com/audio"
                ssrc=0D12F123;seq=14783;rtptime=2345962545

S->C: RTP Packet TS=2345962545 => NPT=3.51
      Media duration=0.16 seconds
```

The server replies with the actual start point that will be delivered. This may differ from the requested range if alignment of the requested range to valid frame boundaries is required for the media source. Note that some media streams in an aggregate may need to be delivered from even earlier points. Also, some media formats have a very long duration per individual data unit, therefore it might be necessary for the client to parse the data unit, and select where to start. The server SHALL also indicate which policy it uses for selecting the actual start point by including a Seek-Style header.

In the following example the client receives the first media packet that stretches all the way up and past the requested playtime. Thus, it is the client's decision whether to render to the user the time between 3.52 and 7.05, or to skip it. In most cases it is probably most suitable not to render that time period.

```
C->S: PLAY rtsp://example.com/audio RTSP/2.0
      CSeq: 836
      Session: 12345678
      Range: npt=7.05-
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 836
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Server: PhonyServer/1.0
      Range: npt=3.52-
      Seek-Style: First-Prior
      RTP-Info:url="rtsp://example.com/audio"
               ssrc=0D12F123:seq=14783;rtptime=2345962545

S->C: RTP Packet TS=2345962545 => NPT=3.52
      Duration=4.15 seconds
```

After playing the desired range, the presentation does NOT change to the Ready state, media delivery simply stops. A PAUSE request MUST be issued to make the stream enter the Ready state. A PLAY request while the stream is still in the Play state is legal, and can be issued without an intervening PAUSE request. Such a request MUST replace the current PLAY action with the new one requested, i.e. being handled the same as the request was received in Ready state. In the case the range in Range header has an implicit start time (-endtime), the server MUST continue to play from where it currently was until the specified end point. This is useful to change end at another point than in the previous request.

The following example plays the whole presentation starting at SMPTE time code 0:10:20 until the end of the clip. Note: The RTP-Info headers has been broken into several lines, where following lines start with whitespace as allowed by the syntax.

```
C->S: PLAY rtsp://audio.example.com/twister.en RTSP/2.0
      CSeq: 833
      Session: 12345678
      Range: smpte=0:10:20-
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 833
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: smpte=0:10:22-0:15:45
      Seek-Style: Next
      RTP-Info:url="rtsp://example.com/twister.en"
               ssrc=0D12F123;seq=14783;rtptime=2345962545
```

For playing back a recording of a live presentation, it may be desirable to use clock units:

```
C->S: PLAY rtsp://audio.example.com/meeting.en RTSP/2.0
      CSeq: 835
      Session: 12345678
      Range: clock=19961108T142300Z-19961108T143520Z
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 835
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: clock=19961108T142300Z-19961108T143520Z
      Seek-Style: Next
      RTP-Info:url="rtsp://example.com/meeting.en"
               ssrc=0D12F123;seq=53745;rtptime=484589019
```

13.4.2. Aggregated Sessions

PLAY requests can operate on sessions controlling a single media and on aggregated sessions controlling multiple media.

In an aggregated session the PLAY request MUST contain an aggregated control URI. A server MUST respond with error 460 (Only Aggregate Operation Allowed) if the client PLAY Request-URI is for a single media. The media in an aggregate MUST be played in sync. If a client wants individual control of the media, it needs to use separate RTSP sessions for each media.

For aggregated sessions where the initial SETUP request (creating a

session) is followed by one or more additional SETUP requests, a PLAY request MAY be pipelined after those additional SETUP requests without awaiting their responses. This procedure can reduce the delay from start of session establishment until media play-out has started with one round trip time. However, a client needs to be aware that using this procedure will result in the playout of the server state established at the time of processing the PLAY, i.e., after the processing of all the requests prior to the PLAY request in the pipeline. This state may not be the intended one due to failure of any of the prior requests. A client can easily determine this based on the responses from those requests. In case of failure, the client can halt the media playout using PAUSE and try to establish the intended state again before issuing another PLAY request.

13.4.3. Updating current PLAY Requests

Clients can issue PLAY requests while the stream is in Play state and thus updating their request.

The important difference compared to a PLAY request in Ready state is the handling of the current play point and how the Range header in request is constructed. The session is actively playing media and the play point will be moving, making the exact time a request will take action hard to predict. Depending on how the PLAY header appears two different cases exist: total replacement or continuation. A total replacement is signaled by having the first range specification have an explicit start value, e.g. npt=45- or npt=45-60, in which case the server stops playout at the current playout point and then starts delivering media according to the Range header. This is equivalent to having the client first send a PAUSE and then a new PLAY request that isn't based on the pause point. In the case of continuation the first range specifier has an implicit start point and an explicit stop value (Z), e.g. npt=-60, which indicate that it MUST convert the range specifier being played prior to this PLAY request (X to Y) into (X to Z) and continue as this was the request originally played. If the current delivery point is beyond the stop point, the server SHALL immediately pause delivery. As the request has been completed successfully it shall be responded with 200 OK. A PLAY_NOTIFY with end-of-stream is also sent to indicate the actual stop point. The pause point is set to the requested stop point.

Following is an example of this behavior: The server has received requests to play ranges 10 to 15. If the new PLAY request arrives at the server 4 seconds after the previous one, it will take effect while the server still plays the first range (10-15). The server changes the current play to continue to 25 seconds, i.e. the equivalent single request would be PLAY with range: npt=10-25.


```
C->S: PLAY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 834
      Session: 12345678
      Range: npt=10-15
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 834
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: npt=10-15
      Seek-Style: Next
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
                 ssrc=0D12F123:seq=5712;rtptime=934207921,
                 url="rtsp://example.com/fizzle/videotrack"
                 ssrc=789DAF12:seq=57654;rtptime=2792482193
      Session: 12345678

C->S: PLAY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 835
      Session: 12345678
      Range: npt=-25
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 835
      Date: Thu, 23 Jan 1997 15:35:09 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: npt=14-25
      Seek-Style: Next
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
                 ssrc=0D12F123:seq=5712;rtptime=934239921,
                 url="rtsp://example.com/fizzle/videotrack"
                 ssrc=789DAF12:seq=57654;rtptime=2792842193
```

A common use of a PLAY request while in Play state is changing the scale of the media, i.e., entering or leaving from fast forward or fast rewind. The client can issue an updating PLAY request that is either a continuation or a complete replacement, as discussed above this section. We give an example of a client that is requesting a fast forward (scale=2) without giving a stop point and then change from fast forward to regular playout (scale = 1).

```
C->S: PLAY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 2034
      Session: 12345678
      Range: npt=now-
      Scale: 2.0
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 2034
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: npt=2:17:21.394-
      Seek-Style: Next
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
                ssrc=0D12F123:seq=5712;rtptime=934207921,
                url="rtsp://example.com/fizzle/videotrack"
                ssrc=789DAF12:seq=57654;rtptime=2792482193
```

[playing in fast forward and now returning to scale = 1]

```
C->S: PLAY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 2035
      Session: 12345678
      Range: npt=now-
      Scale: 1.0
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 2035
      Date: Thu, 23 Jan 1997 15:35:09 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: npt=2:17:27.144-
      Seek-Style: Next
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
                ssrc=0D12F123:seq=5712;rtptime=934239921,
                url="rtsp://example.com/fizzle/videotrack"
                ssrc=789DAF12:seq=57654;rtptime=2792842193
```

13.4.4. Playing On-Demand Media

On-demand media is indicated by the content of the Media-Properties header in the SETUP response by (see also Section 16.28):

- o Random-Access property is set to Random Access;

- o Content Modifications set to Immutable;
- o Retention set to Unlimited or Time-Limited.

Playing on-demand media follows the general usage as described in Section 13.4.1.

13.4.5. Playing Dynamic On-Demand Media

Dynamic on-demand media is indicated by the content of the Media-Properties header in the SETUP response by (see also Section 16.28):

- o RandomAccess set to Random-Access;
- o Content Modifications set to Dynamic;
- o Retention set to Unlimited or Time-Limited.

Playing on-demand media follows the general usage as described in Section 13.4.1 as long as the media has not been changed.

There are two ways for the client to be informed about changes of media resources in Play state. The client will receive a PLAY_NOTIFY request with Notify-Reason header set to media-properties-update (see Section 13.5.2). The client can use the value of the Media-Range to decide further actions, if the Media-Range header is present in the PLAY_NOTIFY request. The second way is that the client issues a GET_PARAMETER request without a body but including a Media-Range header. The 200 OK response MUST include the current Media-Range header (see Section 16.29).

13.4.6. Playing Live Media

Live media is indicated by the content of the Media-Properties header in the SETUP response by (see also Section 16.28):

- o Random-Access set to No-Seeking;
- o Content Modifications set to Time-Progressing;
- o Retention with Time-Duration set to 0.0.

For live media, the SETUP response 200 OK MUST include the Media-Range header (see Section 16.29).

A client MAY send PLAY requests without the Range header. If the request includes the Range header it MUST use a symbolic value representing "now". For NPT that range specification is "npt=now-".

The server MUST include the Range header in the response and it MUST indicate an explicit time value and not a symbolic value. In other words, "npt=now-" is not valid to be used in the response. Instead the time since session start is recommended expressed as an open interval, e.g. "npt=96.23-". An absolute time value (clock) for the corresponding time MAY be given, i.e. "clock=20030213T143205Z-". The UTC clock format can only be used if client has shown support for it using the Accept-Ranges header.

13.4.7. Playing Live with Recording

Certain media servers may offer recording services of live sessions to their clients. This recording would normally be from the beginning of the media session. Clients can randomly access the media between now and the beginning of the media session. This live media with recording is indicated by the content of the Media-Properties header in the SETUP response by (see also Section 16.28):

- o Random-Access set to Random-Access;
- o Content Modifications set to Time-Progressing;
- o Retention set to Time-limited or Unlimited

The SETUP response 200 OK MUST include the Media-Range header (see Section 16.29) for this type of media. For live media with recording, the Range header indicates the current delivery point in the media and the Media-Range header indicates the currently available media window around the current time. This window can cover recorded content in the past (seen from current time in the media) or recorded content in the future (seen from current time in the media). The server adjusts the delivery point to the requested border of the window, if the client requests a delivery point that is located outside the recording windows, e.g., if requested to far in the past, the server selects the oldest range in the recording. The considerations in Section 13.5.3 apply, if a client requests delivery with Scale (Section 16.44) values other than 1.0 (Normal playback rate) while delivering live media with recording.

13.4.8. Playing Live with Time-Shift

Certain media servers may offer time-shift services to their clients. This time shift records a fixed interval in the past, i.e., a sliding window recording mechanism, but not past this interval. Clients can randomly access the media between now and the interval. This live media with recording is indicated by the content of the Media-Properties header in the SETUP response by (see also Section 16.28):

- o Random-Access set to Random-Access;
- o Content Modifications set to Time-Progressing;
- o Retention set to Time-Duration and a value indicating the recording interval (>0).

The SETUP response 200 OK MUST include the Media-Range header (see Section 16.29) for this type of media. For live media with recording the Range header indicates the current time in the media and the Media Range indicates a window around the current time. This window can cover recorded content in the past (seen from current time in the media) or recorded content in the future (seen from current time in the media). The server adjusts the play point to the requested border of the window, if the client requests a play point that is located outside the recording windows, e.g., if requested too far in the past, the server selects the oldest range in the recording. The considerations in Section 13.5.3 apply, if a client requests delivery using a Scale (Section 16.44) value other than 1.0 (Normal playback rate) while delivering live media with time-shift.

13.5. PLAY_NOTIFY

The PLAY_NOTIFY method is issued by a server to inform a client about an asynchronous event for a session in Play state. The Session header MUST be presented in a PLAY_NOTIFY request and indicates the scope of the request. Sending of PLAY_NOTIFY requests requires a persistent connection between server and client, otherwise there is no way for the server to send this request method to the client.

PLAY_NOTIFY requests have an end-to-end (i.e. server to client) scope, as they carry the Session header, and apply only to the given session. The client SHOULD immediately return a response to the server.

PLAY_NOTIFY requests MAY be used with a message body, depending on the value of the Notify-Reason header. It is described in the particular section for each Notify-Reason if a message body is used. However, currently there is no Notify-Reason that allows using a message body. In this case, there is a need to obey some limitations when adding new Notify-Reasons that intend to use a message body: the server can send any type of message body, but it is not ensured that the client can understand the received message body. This is related to DESCRIBE (see Section 13.2), but in this particular case the client can state its acceptable message bodies by using the Accept header. In the case of PLAY_NOTIFY, the server does not know which message bodies are understood by the client.

The Notify-Reason header (see Section 16.31) specifies the reason why the server sends the PLAY_NOTIFY request. This is extensible and new reasons MAY be added in the future (see Section 22.8). In case the client does not understand the reason for the notification it MUST respond with an 465 (Notification Reason Unknown) (Section 15.4.30) error code. Servers can send PLAY_NOTIFY with these types:

- o end-of-stream (see Section 13.5.1);
- o media-properties-update (see Section 13.5.2);
- o scale-change (see Section 13.5.3).

13.5.1. End-of-Stream

A PLAY_NOTIFY request with Notify-Reason header set to end-of-stream indicates the completion or near completion of the PLAY request and the ending delivery of the media stream(s). The request MUST NOT be issued unless the server is in the Play state. The end of the media stream delivery notification may be used to indicate either a successful completion of the PLAY request currently being served, or to indicate some error resulting in failure to complete the request. The Request-Status header (Section 16.40) MUST be included to indicate which request the notification is for and its completion status. The message response status codes (Section 8.1.1) are used to indicate how the PLAY request concluded. The sender of a PLAY_NOTIFY can issue an updated PLAY_NOTIFY, in the case of a PLAY_NOTIFY sent with wrong information. For instance, a PLAY_NOTIFY was issued before reaching the end-of-stream, but some error occurred resulting in that the previously sent PLAY_NOTIFY contained a wrong time when the stream will end. In this case a new PLAY_NOTIFY MUST be sent including the correct status for the completion and all additional information.

PLAY_NOTIFY requests with Notify-Reason header set to end-of-stream MUST include a Range header and the Scale header if the scale value is not 1. The Range header indicates the point in the stream or streams where delivery is ending with the timescale that was used by the server in the PLAY response for the request being fulfilled. The server MUST NOT use the "now" constant in the Range header; it MUST use the actual numeric end position in the proper timescale. When end-of-stream notifications are issued prior to having sent the last media packets, this is evident as the end time in the Range header is beyond the current time in the media being received by the client, e.g., npt=-15, if npt is currently at 14.2 seconds. The Scale header is to be included so that it is evident if the media time scale is moving backwards and/or have a non-default pace. The end-of-stream notification does not prevent the client from sending a new PLAY

request.

If RTP is used as media transport, a RTP-Info header MUST be included, and the RTP-Info header MUST indicate the last sequence number in the seq parameter.

A PLAY_NOTIFY request with Notify-Reason header set to end-of-stream MUST NOT carry a message body.

This example request notifies the client about a future end-of-stream event:

```
S->C: PLAY_NOTIFY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 854
      Notify-Reason: end-of-stream
      Request-Status: cseq=853 status=200 reason="OK"
      Range: npt=-145
      RTP-Info:url="rtsp://example.com/audio"
               ssrc=0D12F123:seq=14783;rtptime=2345962545
      Session: uZ3ci0K+Ld-M
      Date: Mon, 08 Mar 2010 13:37:16 GMT
```

```
C->S: RTSP/2.0 200 OK
      CSeq: 854
      User-Agent: PhonyClient/1.2
      Session: uZ3ci0K+Ld-M
```

13.5.2. Media-Properties-Update

A PLAY_NOTIFY request with Notify-Reason header set to media-properties-update indicates an update of the media properties for the given session (see Section 16.28) and/or the available media range that can be played as indicated by Media-Range (Section 16.29). PLAY_NOTIFY requests with Notify-Reason header set to media-properties-update MUST include a Media-Properties and Date header and SHOULD include a Media-Range header.

This notification MUST be sent for media that are time-progressing every time an event happens that changes the basis for making estimates on how the media range progress. In addition it is RECOMMENDED that the server sends these notifications every 5 minutes for time-progressing content to ensure the long-term stability of the client estimation and allowing for clock skew detection by the client. Requests for the just mentioned reasons MUST include Media-Range header to provide current Media duration and the Range header to indicate the current playing point and any remaining parts of the requested range.

The recommendation for sending updates every 5 minutes is due to any clock skew issues. In 5 minutes the clock skew should not become too significant as this is not used for media playback and synchronization, only for determining which content is available to the user.

A PLAY_NOTIFY request with Notify-Reason header set to media-properties-update MUST NOT carry a message body.

```
S->C: PLAY_NOTIFY rtsp://example.com/fizzle/foo RTSP/2.0
      Date: Tue, 14 Apr 2008 15:48:06 GMT
      CSeq: 854
      Notify-Reason: media-properties-update
      Session: uZ3ci0K+Ld-M
      Media-Properties: Time-Progressing,
                       Time-Limited=20080415T153919.36Z, Random-Access=5.0
      Media-Range: npt=0-1:37:21.394
      Range: npt=1:15:49.873-

C->S: RTSP/2.0 200 OK
      CSeq: 854
      User-Agent: PhonyClient/1.2
      Session: uZ3ci0K+Ld-M
```

13.5.3. Scale-Change

The server may be forced to change the rate, when a client request delivery using a Scale (Section 16.44) value other than 1.0 (normal playback rate). For time progressing media with some retention, i.e. the server stores already sent content, a client requesting to play with Scale values larger than 1 may catch up with the front end of the media. The server will then be unable to continue to provide content at Scale larger than 1 as content is only made available by the server at Scale=1. Another case is when Scale < 1 and the media retention is time-duration limited. In this case the delivery point can reach the oldest media unit available, and further playback at this scale becomes impossible as there will be no media available. To avoid having the client lose any media, the scale will need to be adjusted to the same rate at which the media is removed from the storage buffer, commonly Scale = 1.0.

Another case is when the content itself consists of spliced pieces or is dynamically updated. In these cases the server may be required to change from one supported scale value (different than Scale=1.0) to another. In this case the server will pick the closest value and inform the client of what it has picked. In these cases the media properties will also be sent updating the supported Scale values. This enables a client to adjust the used Scale value.

To minimize impact on playback in any of the above cases the server MUST modify the playback properties and set Scale to a supportable value and continue delivery of the media. When doing this modification it MUST send a PLAY_NOTIFY message with the Notify-Reason header set to "scale-change". The request MUST contain a Range header with the media time where the change took effect, a Scale header with the new value in use, Session header with the ID for the session it applies to and a Date header with the server wallclock time of the change. For time progressing content also the Media-Range and the Media-Properties at this point in time MUST be included. The Media-Properties header MUST be included if the scale change was due to the content changing what scale values that is supported.

For media streams being delivered using RTP also a RTP-Info header MUST be included. It MUST contain the rtptime parameter with a value corresponding to the point of change in that media and optionally also the sequence number.

A PLAY_NOTIFY request with Notify-Reason header set to "Scale-Change" MUST NOT carry a message body.

```
S->C: PLAY_NOTIFY rtsp://example.com/fizzle/foo RTSP/2.0
      Date: Tue, 14 Apr 2008 15:48:06 GMT
      CSeq: 854
      Notify-Reason: scale-change
      Session: uZ3ci0K+Ld-M
      Media-Properties: Time-Progressing,
                      Time-Limited=20080415T153919.36Z, Random-Access=5.0
      Media-Range: npt=0-1:37:21.394
      Range: npt=1:37:21.394-
      Scale: 1
      RTP-Info: url="rtsp://example.com/fizzle/foo/audio"
                ssrc=0D12F123:rtptime=2345962545

C->S: RTSP/2.0 200 OK
      CSeq: 854
      User-Agent: PhonyClient/1.2
      Session: uZ3ci0K+Ld-M
```

13.6. PAUSE

The PAUSE request causes the stream delivery to immediately be interrupted (halted). A PAUSE request MUST be done either with the aggregated control URI for aggregated sessions, resulting in all media being halted, or the media URI for non-aggregated sessions. Any attempt to do muting of a single media with a PAUSE request in an aggregated session MUST be responded to with error 460 (Only

Aggregate Operation Allowed). After resuming playback, synchronization of the tracks MUST be maintained. Any server resources are kept, though servers MAY close the session and free resources after being paused for the duration specified with the timeout parameter of the Session header in the SETUP message.

Example:

```
C->S: PAUSE rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 834
      Session: 12345678
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 834
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Range: npt=45.76-75.00
```

The PAUSE request causes stream delivery to be interrupted immediately on receipt of the message and the pause point is set to the current point in the presentation. That pause point in the media stream needs to be maintained. A subsequent PLAY request without Range header resume from the pause point and plays until media end.

The pause point after any PAUSE request MUST be returned to the client by adding a Range header with what remains unplayed of the PLAY request's range. For media with random access properties, if one desires to resume playing a ranged request, one simply includes the Range header from the PAUSE response and includes the Seek-Style header with the Next policy in the PLAY request. For media that is time-progressing and has retention duration=0 the follow-up PLAY request to start media delivery again, will need to use "npt=now-" and not the answer given in the response to PAUSE.

```
C->S: PLAY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 834
      Session: 12345678
      Range: npt=10-30
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 834
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Server: PhonyServer/1.0
      Range: npt=10-30
      Seek-Style: First-Prior
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
                ssrc=0D12F123:seq=5712;rtptime=934207921,
                url="rtsp://example.com/fizzle/videotrack"
                ssrc=4FAD8726:seq=57654;rtptime=2792482193
      Session: 12345678
```

After 11 seconds, i.e. at 21 seconds into the presentation:

```
C->S: PAUSE rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 835
      Session: 12345678
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 835
      Date: 23 Jan 1997 15:35:17 GMT
      Server: PhonyServer/1.0
      Range: npt=21-30
      Session: 12345678
```

If a client issues a PAUSE request and the server acknowledges and enters the Ready state, the proper server response, if the player issues another PAUSE, is still 200 OK. The 200 OK response **MUST** include the Range header with the current pause point. See examples below:

```
C->S: PAUSE rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 834
      Session: 12345678
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 834
      Session: 12345678
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Range: npt=45.76-98.36
```

```
C->S: PAUSE rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 835
      Session: 12345678
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 835
      Session: 12345678
      Date: 23 Jan 1997 15:35:07 GMT
      Range: npt=45.76-98.36
```

13.7. TEARDOWN

13.7.1. Client to Server

The TEARDOWN client to server request stops the stream delivery for the given URI, freeing the resources associated with it. A TEARDOWN request MAY be performed on either an aggregated or a media control URI. However, some restrictions apply depending on the current state. The TEARDOWN request MUST contain a Session header indicating what session the request applies to.

A TEARDOWN using the aggregated control URI or the media URI in a session under non-aggregated control (single media session) MAY be done in any state (Ready and Play). A successful request MUST result in that media delivery being immediately halted and the session state being destroyed. This MUST be indicated through the lack of a Session header in the response.

A TEARDOWN using a media URI in an aggregated session MAY only be done in Ready state. Such a request only removes the indicated media stream and associated resources from the session. This may result in that a session returns to non-aggregated control, due to that it only contains a single media after the requests completion. A session that will exist after the processing of the TEARDOWN request MUST in the response to that TEARDOWN request contain a Session header. Thus the presence of the Session header indicates to the receiver of the

response if the session is still existing or has been removed.

Example:

```
C->S: TEARDOWN rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 892
      Session: 12345678
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 892
      Server: PhonyServer/1.0
```

13.7.2. Server to Client

The server can send TEARDOWN requests in the server to client direction to indicate that the server has been forced to terminate the ongoing session. This may happen for several reasons, such as server maintenance without available backup, or that the session has been inactive for extended periods of time. The reason is provided in the Terminate-Reason header (Section 16.50).

When a RTSP client has maintained a RTSP session that otherwise is inactive for an extended period of time the server may reclaim the resources. That is done by issuing a TEARDOWN request with the Terminate-Reason set to "Session-Timeout". This MAY be done when the client has been inactive in the RTSP session for more than one Session Timeout period (Section 16.47). However, the server is RECOMMENDED to not perform this operation until an extended period of inactivity has passed. The time period is considered extended when it is 10 times the Session Timeout period. Consideration of the application of the server and its content should be performed when configuring what is considered as extended period of time.

In case the server needs to stop providing service to the established sessions and there is no server to point at in a REDIRECT request, then TEARDOWN SHALL be used to terminate the session. This method can also be used when non-recoverable internal errors have happened and the server has no other option then to terminate the sessions.

The TEARDOWN request MUST be done only on the session aggregate control URI (i.e., it is not allowed to terminate individual media streams, if it is a session aggregate) and MUST include the following headers; Session and Terminate-Reason headers. The request only applies to the session identified in the Session header. The server may include a message to the client's user with the "user-msg" parameter.

The TEARDOWN request may alternatively be done on the wild card URI * and without any session header. The scope of such a request is limited to the next-hop (i.e. the RTSP agent in direct communication with the server) and applies, as well, to the control connection between the next-hop RTSP agent and the server. This request indicates that all sessions and pending requests being managed via the control connection are terminated. Any intervening proxies SHOULD do all of the following in the order listed:

1. respond to the TEARDOWN request
2. disconnect the control channel from the requesting server
3. pass the TEARDOWN request to each applicable client (typically those clients with an active session or an unanswered request)

Note: The proxy is responsible for accepting TEARDOWN responses from its clients; these responses MUST NOT be passed on to either the original server or the target server in the redirect.

13.8. GET_PARAMETER

The GET_PARAMETER request retrieves the value of any specified parameter or parameters for a presentation or stream specified in the URI. If the Session header is present in a request, the value of a parameter MUST be retrieved in the specified session context. There are two ways of specifying the parameters to be retrieved. The first is by including headers which have been defined such that you can use them for this purpose. Headers for this purpose should allow empty, or stripped value parts to avoid having to specify bogus data when indicating the desire to retrieve a value. The successful completion of the request should also be evident from any filled out values in the response. The Media-Range header (Section 16.29) is one such header. The other way is to specify a message body that lists the parameter(s) that are desired to be retrieved. The Content-Type header (Section 16.18) is used to specify which format the message body has.

The headers that MAY be used for retrieving their current value using GET_PARAMETER are:

- o Accept-Ranges
- o Media-Range
- o Media-Properties

- o Range
- o RTP-Info

The method MAY also be used without a message body or any header that request parameters for keep-alive purpose. The keep-alive timer has been updated for any request that is successful, i.e., a 200 OK response is received. Any non-required header present in such a request may or may not have been processed. Normally the presence of filled out values in the header will be indication that the header has been processed. However, for cases when this is difficult to determine, it is recommended to use a feature-tag and the Require header. Due to this reason it is usually easier if any parameters to be retrieved are sent in the body, rather than using any header.

Parameters specified within the body of the message must all be understood by the request receiving agent. If one or more parameters are not understood a 451 (Parameter Not Understood) MUST be sent including a body listing these parameters that weren't understood. If all parameters are understood their values are filled in and returned in the response message body.

Example:

```
S->C: GET_PARAMETER rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 431
      User-Agent: PhonyClient/1.2
      Session: 12345678
      Content-Length: 26
      Content-Type: text/parameters

      packets_received
      jitter

C->S: RTSP/2.0 200 OK
      CSeq: 431
      Session: 12345678
      Server: PhonyServer/1.1
      Date: Mon, 08 Mar 2010 13:43:23 GMT
      Content-Length: 38
      Content-Type: text/parameters

      packets_received: 10
      jitter: 0.3838
```

13.9. SET_PARAMETER

This method requests to set the value of a parameter or a set of parameters for a presentation or stream specified by the URI. The method MAY also be used without a message body. It is the RECOMMENDED method to be used in a request sent for the sole purpose of updating the keep-alive timer. If this request is successful, i.e. a 200 OK response is received, then the keep-alive timer has been updated. Any non-required header present in such a request may or may not have been processed. To allow a client to determine if any such header has been processed, it is necessary to use a feature tag and the Require header. Due to this reason it is RECOMMENDED that any parameters are sent in the body, rather than using any header.

A request is RECOMMENDED to only contain a single parameter to allow the client to determine why a particular request failed. If the request contains several parameters, the server MUST only act on the request if all of the parameters can be set successfully. A server MUST allow a parameter to be set repeatedly to the same value, but it MAY disallow changing parameter values. If the receiver of the request does not understand or cannot locate a parameter, error 451 (Parameter Not Understood) MUST be used. In the case a parameter is not allowed to change, the error code is 458 (Parameter Is Read-Only). The response body MUST contain only the parameters that have errors. Otherwise no body MUST be returned.

Note: transport parameters for the media stream MUST only be set with the SETUP command.

Restricting setting transport parameters to SETUP is for the benefit of firewalls.

The parameters are split in a fine-grained fashion so that there can be more meaningful error indications. However, it may make sense to allow the setting of several parameters if an atomic setting is desirable. Imagine device control where the client does not want the camera to pan unless it can also tilt to the right angle at the same time.

Example:


```
C->S: SET_PARAMETER rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 421
      User-Agent: PhonyClient/1.2
      Session: iixT43KLc
      Date: Mon, 08 Mar 2010 14:45:04 GMT
      Content-length: 20
      Content-type: text/parameters
```

```
      barparam: barstuff
```

```
S->C: RTSP/2.0 451 Parameter Not Understood
      CSeq: 421
      Session: iixT43KLc
      Server: PhonyServer/1.0
      Date: Mon, 08 Mar 2010 14:44:56 GMT
      Content-length: 20
      Content-type: text/parameters
```

```
      barparam: barstuff
```

13.10. REDIRECT

The REDIRECT method is issued by a server to inform a client that the service provided will be terminated and where a corresponding service can be provided instead. This may happen for different reasons. One is that the server is being administered such that it must stop providing service. Thus the client is required to connect to another server location to access the resource indicated by the Request-URI.

The REDIRECT request SHALL contain a Terminate-Reason header (Section 16.50) to inform the client of the reason for the request. Additional parameters related to the reason may also be included. The intention here is to allow a server administrator to do a controlled shutdown of the RTSP server. That requires sufficient time to inform all entities having associated state with the server and for them to perform a controlled migration from this server to a fall back server.

A REDIRECT request with a Session header has end-to-end (i.e. server to client) scope and applies only to the given session. Any intervening proxies SHOULD NOT disconnect the control channel while there are other remaining end-to-end sessions. The REQUIRED Location header MUST contain a complete absolute URI pointing to the resource to which the client SHOULD reconnect. Specifically, the Location MUST NOT contain just the host and port. A client may receive a REDIRECT request with a Session header, if and only if, an end-to-end session has been established.

A client may receive a REDIRECT request without a Session header at any time when it has communication or a connection established with a server. The scope of such a request is limited to the next-hop (i.e. the RTSP agent in direct communication with the server) and applies to all sessions controlled, as well as the control connection between the next-hop RTSP agent and the server. A REDIRECT request without a Session header indicates that all sessions and pending requests being managed via the control connection MUST be redirected. The REQUIRED Location header, if included in such a request, SHOULD contain an absolute URI with only the host address and the OPTIONAL port number of the server to which the RTSP agent SHOULD reconnect. Any intervening proxies SHOULD do all of the following in the order listed:

1. respond to the REDIRECT request
2. disconnect the control channel from the requesting server
3. connect to the server at the given host address
4. pass the REDIRECT request to each applicable client (typically those clients with an active session or an unanswered request)

Note: The proxy is responsible for accepting REDIRECT responses from its clients; these responses MUST NOT be passed on to either the original server or the redirected server.

When the server lacks any alternative server and needs to terminate a session or all sessions the TEARDOWN request SHALL be used instead.

When no Terminate-Reason "time" parameter are included in a REDIRECT request, the client SHALL perform the redirection immediately and return a response to the server. The server shall consider the session as terminated and can free any associated state after it receives the successful (2xx) response. The server MAY close the signaling connection upon receiving the response and the client SHOULD close the signaling connection after sending the 2xx response. The exception to this is when the client has several sessions on the server being managed by the given signaling connection. In this case, the client SHOULD close the connection when it has received and responded to REDIRECT requests for all the sessions managed by the signaling connection.

The Terminate-Reason header "time" parameter MAY be used to indicate the wallclock time by when the redirection MUST have taken place. To allow a client to determine that redirect time without being time synchronized with the server, the server MUST include a Date header in the request. The client should have before the redirection time-

line terminated the session and closed the control connection. The server MAY simply cease to provide service when the deadline time has been reached, or it may issue TEARDOWN requests to the remaining sessions.

If the REDIRECT request times out following the rules in Section 10.4 the server MAY terminate the session or transport connection that would be redirected by the request. This is a safeguard against misbehaving clients that refuse to respond to a REDIRECT request. That should not provide any benefit.

After a REDIRECT request has been processed, a client that wants to continue to send or receive media for the resource identified by the Request-URI will have to establish a new session with the designated host. If the URI given in the Location header is a valid resource URI, a client SHOULD issue a DESCRIBE request for the URI.

Note: The media resource indicated by the Location header can be identical, slightly different or totally different. This is the reason why a new DESCRIBE request SHOULD be issued.

If the Location header contains only a host address, the client MAY assume that the media on the new server is identical to the media on the old server, i.e. all media configuration information from the old session is still valid except for the host address. However, the usage of conditional SETUP using MTag identifiers are RECOMMENDED to verify the assumption.

This example request redirects traffic for this session to the new server at the given absolute time:

```
S->C: REDIRECT rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 732
      Location: rtsp://s2.example.com:8001
      Terminate-Reason: Server-Admin ;time=19960213T143205Z
      Session: uZ3ci0K+Ld-M
      Date: Thu, 13 Feb 1996 14:30:43 GMT

C->S: RTSP/2.0 200 OK
      CSeq: 732
      User-Agent: PhonyClient/1.2
      Session: uZ3ci0K+Ld-M
```

14. Embedded (Interleaved) Binary Data

In order to fulfill certain requirements on the network side, e.g. in conjunction with network address translators that block RTP traffic over UDP, it may be necessary to interleave RTSP messages and media stream data. This interleaving should generally be avoided unless necessary since it complicates client and server operation and imposes additional overhead. Also, head of line blocking may cause problems. Interleaved binary data **SHOULD** only be used if RTSP is carried over TCP. Interleaved data is not allowed inside RTSP messages.

Stream data such as RTP packets is encapsulated by an ASCII dollar sign (36 decimal), followed by a one-byte channel identifier, followed by the length of the encapsulated binary data as a binary, two-byte integer in network byte order. The stream data follows immediately afterwards, without a CRLF, but including the upper-layer protocol headers. Each \$ block **MUST** contain exactly one upper-layer protocol data unit, e.g., one RTP packet.

```

      0               1               2               3
      0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
| "$" = 36   | Channel ID   | Length in bytes   |
+-----+-----+-----+-----+-----+-----+-----+-----+
: Length number of bytes of binary data                                     :
+-----+-----+-----+-----+-----+-----+-----+-----+

```

The channel identifier is defined in the Transport header with the interleaved parameter (Section 16.52).

When the transport choice is RTP, RTCP messages are also interleaved by the server over the TCP connection. The usage of RTCP messages is indicated by including an interval containing a second channel in the interleaved parameter of the Transport header, see Section 16.52. If RTCP is used, packets **MUST** be sent on the first available channel higher than the RTP channel. The channels are bi-directional, using the same ChannelID in both directions, and therefore RTCP traffic are sent on the second channel in both directions.

RTCP is sometimes needed for synchronization when two or more streams are interleaved in such a fashion. Also, this provides a convenient way to tunnel RTP/RTCP packets through the TCP control connection when required by the network configuration and transfer them onto UDP when possible.

```
C->S: SETUP rtsp://example.com/bar.file RTSP/2.0
      CSeq: 2
      Transport: RTP/AVP/TCP;unicast;interleaved=0-1
      Accept-Ranges: NPT, SMPTE, UTC
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 2
      Date: Thu, 05 Jun 1997 18:57:18 GMT
      Transport: RTP/AVP/TCP;unicast;interleaved=5-6
      Session: 12345678
      Accept-Ranges: NPT
      Media-Properties: Random-Access=0.2, Immutable, Unlimited

C->S: PLAY rtsp://example.com/bar.file RTSP/2.0
      CSeq: 3
      Session: 12345678
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 3
      Session: 12345678
      Date: Thu, 05 Jun 1997 18:57:19 GMT
      RTP-Info: url="rtsp://example.com/bar.file"
                ssrc=0D12F123;seq=232433;rtptime=972948234
      Range: npt=0-56.8
      Seek-Style: RAP

S->C: $005{2 byte length}{"length" bytes data, w/RTP header}
S->C: $005{2 byte length}{"length" bytes data, w/RTP header}
S->C: $006{2 byte length}{"length" bytes RTCP packet}
```

15. Status Code Definitions

Where applicable, HTTP status [H10] codes are reused. Status codes that have the same meaning are not repeated here. See Table 4 in Section 8.1 for a listing of which status codes may be returned by which requests. All error messages, 4xx and 5xx MAY return a body containing further information about the error.

15.1. Success 1xx

15.1.1. 100 Continue

The client SHOULD continue with its request. This interim response is used to inform the client that the initial part of the request has been received and has not yet been rejected by the server. The client SHOULD continue by sending the remainder of the request or, if the request has already been completed, ignore this response. The server MUST send a final response after the request has been completed.

15.2. Success 2xx

This class of status code indicates that the client's request was successfully received, understood, and accepted.

15.2.1. 200 OK

The request has succeeded. The information returned with the response is dependent on the method used in the request.

15.3. Redirection 3xx

The notation "3rr" indicates response codes from 300 to 399 inclusive which are meant for redirection. The response code 304 is excluded from this set, as it is not used for redirection.

Within RTSP, redirection may be used for load balancing or redirecting stream requests to a server topologically closer to the client. Mechanisms to determine topological proximity are beyond the scope of this specification.

A 3rr code MAY be used to respond to any request. It is RECOMMENDED that they are used if necessary before a session is established, i.e., in response to DESCRIBE or SETUP. However, in cases where a server is not able to send a REDIRECT request to the client, the server MAY need to resort to using 3rr responses to inform a client with an established session about the need for redirecting the session. If a 3rr response is received for a request in relation to

an established session, the client SHOULD send a TEARDOWN request for the session, and MAY reestablish the session using the resource indicated by the Location.

If the Location header is used in a response it MUST contain an absolute URI pointing out the media resource the client is redirected to, the URI MUST NOT only contain the host name.

15.3.1. 301 Moved Permanently

The requested resource is moved permanently and resides now at the URI given by the location header. The user client SHOULD redirect automatically to the given URI. This response MUST NOT contain a message-body. The Location header MUST be included in the response.

15.3.2. 302 Found

The requested resource resides temporarily at the URI given by the Location header. The Location header MUST be included in the response. This response is intended to be used for many types of temporary redirects; e.g., load balancing. It is RECOMMENDED that the server set the reason phrase to something more meaningful than "Found" in these cases. The user client SHOULD redirect automatically to the given URI. This response MUST NOT contain a message-body.

This example shows a client being redirected to a different server:

```
C->S: SETUP rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 2
      Transport: RTP/AVP/TCP;unicast;interleaved=0-1
      Accept-Ranges: NPT, SMPTE, UTC
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 302 Try Other Server
      CSeq: 2
      Location: rtsp://s2.example.com:8001/fizzle/foo
```

15.3.3. 303 See Other

This status code MUST NOT be used in RTSP 2.0. However, it was allowed to use in RTSP 1.0 (RFC 2326).

15.3.4. 304 Not Modified

If the client has performed a conditional DESCRIBE or SETUP (see Section 16.24) and the requested resource has not been modified, the server SHOULD send a 304 response. This response MUST NOT contain a

message-body.

The response MUST include the following header fields:

- o Date
- o MTag and/or Content-Location, if the header(s) would have been sent in a 200 response to the same request.
- o Expires, Cache-Control, and/or Vary, if the field-value might differ from that sent in any previous response for the same variant.

This response is independent for the DESCRIBE and SETUP requests. That is, a 304 response to DESCRIBE does NOT imply that the resource content is unchanged (only the session description) and a 304 response to SETUP does NOT imply that the resource description is unchanged. The MTag and If-Match headers may be used to link the DESCRIBE and SETUP in this manner.

15.3.5. 305 Use Proxy

The requested resource MUST be accessed through the proxy given by the Location field. The Location field gives the URI of the proxy. The recipient is expected to repeat this single request via the proxy. 305 responses MUST only be generated by origin servers.

15.4. Client Error 4xx

15.4.1. 400 Bad Request

The request could not be understood by the server due to malformed syntax. The client SHOULD NOT repeat the request without modifications. If the request does not have a CSeq header, the server MUST NOT include a CSeq in the response.

15.4.2. 401 Unauthorized

The request requires user authentication. The response MUST include a WWW-Authenticate header (Section 16.57) field containing a challenge applicable to the requested resource. The client MAY repeat the request with a suitable Authorization header field. If the request already included Authorization credentials, then the 401 response indicates that authorization has been refused for those credentials. If the 401 response contains the same challenge as the prior response, and the user agent has already attempted authentication at least once, then the user SHOULD be presented the message body that was given in the response, since that message body

might include relevant diagnostic information. HTTP access authentication is explained in [RFC2617].

15.4.3. 402 Payment Required

This code is reserved for future use.

15.4.4. 403 Forbidden

The server understood the request, but is refusing to fulfill it. Authorization will not help and the request SHOULD NOT be repeated. If the server wishes to make public why the request has not been fulfilled, it SHOULD describe the reason for the refusal in the message body. If the server does not wish to make this information available to the client, the status code 404 (Not Found) can be used instead.

15.4.5. 404 Not Found

The server has not found anything matching the Request-URI. No indication is given of whether the condition is temporary or permanent. The 410 (Gone) status code SHOULD be used if the server knows, through some internally configurable mechanism, that an old resource is permanently unavailable and has no forwarding address. This status code is commonly used when the server does not wish to reveal exactly why the request has been refused, or when no other response is applicable.

15.4.6. 405 Method Not Allowed

The method specified in the request is not allowed for the resource identified by the Request-URI. The response MUST include an Allow header containing a list of valid methods for the requested resource. This status code is also to be used if a request attempts to use a method not indicated during SETUP.

15.4.7. 406 Not Acceptable

The resource identified by the request is only capable of generating response message bodies which have content characteristics not acceptable according to the Accept headers sent in the request.

The response SHOULD include a message body containing a list of available message body characteristics and location(s) from which the user or user agent can choose the one most appropriate. The message body format is specified by the media type given in the Content-Type header field. Depending upon the format and the capabilities of the user agent, selection of the most appropriate choice MAY be performed

automatically. However, this specification does not define any standard for such automatic selection.

If the response could be unacceptable, a user agent SHOULD temporarily stop receipt of more data and query the user for a decision on further actions.

15.4.8. 407 Proxy Authentication Required

This code is similar to 401 (Unauthorized) (Section 15.4.2), but indicates that the client must first authenticate itself with the proxy. The proxy MUST return a Proxy-Authenticate header field (Section 16.33) containing a challenge applicable to the proxy for the requested resource.

15.4.9. 408 Request Timeout

The client did not produce a request within the time that the server was prepared to wait. The client MAY repeat the request without modifications at any later time.

15.4.10. 410 Gone

The requested resource is no longer available at the server and the forwarding address is not known. This condition is expected to be considered permanent. If the server does not know, or has no facility to determine, whether or not the condition is permanent, the status code 404 (Not Found) SHOULD be used instead. This response is cacheable unless indicated otherwise.

The 410 response is primarily intended to assist the task of repository maintenance by notifying the recipient that the resource is intentionally unavailable and that the server owners desire that remote links to that resource be removed. Such an event is common for limited-time, promotional services and for resources belonging to individuals no longer working at the server's site. It is not necessary to mark all permanently unavailable resources as "gone" or to keep the mark for any length of time -- that is left to the discretion of the owner of the server.

15.4.11. 411 Length Required

The server refuses to accept the request without a defined Content-Length. The client MAY repeat the request if it adds a valid Content-Length header field containing the length of the message-body in the request message.

15.4.12. 412 Precondition Failed

The precondition given in one or more of the request-header fields evaluated to false when it was tested on the server. This response code allows the client to place preconditions on the current resource meta information (header field data) and thus prevent the requested method from being applied to a resource other than the one intended.

15.4.13. 413 Request Message Body Too Large

The server is refusing to process a request because the request message body is larger than the server is willing or able to process. The server MAY close the connection to prevent the client from continuing the request.

If the condition is temporary, the server SHOULD include a Retry-After header field to indicate that it is temporary and after what time the client MAY try again.

15.4.14. 414 Request-URI Too Long

The server is refusing to service the request because the Request-URI is longer than the server is willing to interpret. This rare condition is only likely to occur when a client has used a request with long query information, when the client has descended into a URI "black hole" of redirection (e.g., a redirected URI prefix that points to a suffix of itself), or when the server is under attack by a client attempting to exploit security holes present in some servers using fixed-length buffers for reading or manipulating the Request-URI.

15.4.15. 415 Unsupported Media Type

The server is refusing to service the request because the message body of the request is in a format not supported by the requested resource for the requested method.

15.4.16. 451 Parameter Not Understood

The recipient of the request does not support one or more parameters contained in the request. When returning this error message the sender SHOULD return a message body containing the offending parameter(s).

15.4.17. 452 reserved

This error code was removed from RFC 2326 [RFC2326] as it is obsolete. This error code MUST NOT be used anymore.

15.4.18. 453 Not Enough Bandwidth

The request was refused because there was insufficient bandwidth. This may, for example, be the result of a resource reservation failure.

15.4.19. 454 Session Not Found

The RTSP session identifier in the Session header is missing, invalid, or has timed out.

15.4.20. 455 Method Not Valid in This State

The client or server cannot process this request in its current state. The response **MUST** contain an Allow header to make error recovery possible.

15.4.21. 456 Header Field Not Valid for Resource

The server could not act on a required request header. For example, if PLAY contains the Range header field but the stream does not allow seeking. This error message may also be used for specifying when the time format in Range is impossible for the resource. In that case the Accept-Ranges header **MUST** be returned to inform the client of which format(s) that are allowed.

15.4.22. 457 Invalid Range

The Range value given is out of bounds, e.g., beyond the end of the presentation.

15.4.23. 458 Parameter Is Read-Only

The parameter to be set by SET_PARAMETER can be read but not modified. When returning this error message the sender **SHOULD** return a message body containing the offending parameter(s).

15.4.24. 459 Aggregate Operation Not Allowed

The requested method may not be applied on the URI in question since it is an aggregate (presentation) URI. The method may be applied on a media URI.

15.4.25. 460 Only Aggregate Operation Allowed

The requested method may not be applied on the URI in question since it is not an aggregate control (presentation) URI. The method may be applied on the aggregate control URI.

15.4.26. 461 Unsupported Transport

The Transport field did not contain a supported transport specification.

15.4.27. 462 Destination Unreachable

The data transmission channel could not be established because the client address could not be reached. This error will most likely be the result of a client attempt to place an invalid dest_addr parameter in the Transport field.

15.4.28. 463 Destination Prohibited

The data transmission channel was not established because the server prohibited access to the client address. This error is most likely the result of a client attempt to redirect media traffic to another destination with a dest_addr parameter in the Transport header.

15.4.29. 464 Data Transport Not Ready Yet

The data transmission channel to the media destination is not yet ready for carrying data. However, the responding agent still expects that the data transmission channel will be established at some point in time. Note, however, that this may result in a permanent failure like 462 "Destination Unreachable".

An example when this error may occur is in the case a client sends a PLAY request to a server prior to ensuring that the TCP connections negotiated for carrying media data was successfully established (In violation of this specification). The server would use this error code to indicate that the requested action could not be performed due to the failure of completing the connection establishment.

15.4.30. 465 Notification Reason Unknown

This indicates that the client has received a PLAY_NOTIFY (Section 13.5) with a Notify-Reason header (Section 16.31) unknown to the client.

15.4.31. 466 Key Management Error

This indicates that there has been an error in a Key Management function used in conjunction with a request. For example usage of MIKEY according to Appendix C.1.4.1 may result in this error.

15.4.32. 470 Connection Authorization Required

The secured connection attempt needs user or client authorization before proceeding. The next hops certificate is included in this response in the Accept-Credentials header.

15.4.33. 471 Connection Credentials not accepted

When performing a secure connection over multiple connections, an intermediary has refused to connect to the next hop and carry out the request due to unacceptable credentials for the used policy.

15.4.34. 472 Failure to establish secure connection

A proxy fails to establish a secure connection to the next hop RTSP agent. This is primarily caused by a fatal failure at the TLS handshake, for example due to server not accepting any cipher suites.

15.5. Server Error 5xx

Response status codes beginning with the digit "5" indicate cases in which the server is aware that it has erred or is incapable of performing the request. The server SHOULD include a message body containing an explanation of the error situation, and whether it is a temporary or permanent condition. User agents SHOULD display any included message body to the user. These response codes are applicable to any request method.

15.5.1. 500 Internal Server Error

The server encountered an unexpected condition which prevented it from fulfilling the request.

15.5.2. 501 Not Implemented

The server does not support the functionality required to fulfill the request. This is the appropriate response when the server does not recognize the request method and is not capable of supporting it for any resource.

15.5.3. 502 Bad Gateway

The server, while acting as a gateway or proxy, received an invalid response from the upstream server it accessed in attempting to fulfill the request.

15.5.4. 503 Service Unavailable

The server is currently unable to handle the request due to a temporary overloading or maintenance of the server. The implication is that this is a temporary condition which will be alleviated after some delay. If known, the length of the delay MAY be indicated in a Retry-After header. If no Retry-After is given, the client SHOULD handle the response as it would for a 500 response. The client MUST honor the length, if given in the Retry-After header.

Note: The existence of the 503 status code does not imply that a server must use it when becoming overloaded. Some servers may wish to simply refuse the connection.

15.5.5. 504 Gateway Timeout

The server, while acting as a proxy, did not receive a timely response from the upstream server specified by the URI or some other auxiliary server (e.g., DNS) it needed to access in attempting to complete the request.

15.5.6. 505 RTSP Version Not Supported

The server does not support, or refuses to support, the RTSP protocol version that was used in the request message. The server is indicating that it is unable or unwilling to complete the request using the same major version as the client other than with this error message. The response SHOULD contain a message body describing why that version is not supported and what other protocols are supported by that server.

15.5.7. 551 Option not supported

A feature-tag given in the Require or the Proxy-Require fields was not supported. The Unsupported header MUST be returned stating the feature for which there is no support.

16. Header Field Definitions

method	direction	object	acronym	Body
DESCRIBE	C -> S	P,S	DES	r
GET_PARAMETER	C -> S, S -> C	P,S	GPR	R,r
OPTIONS	C -> S, S -> C	P,S	OPT	
PAUSE	C -> S	P,S	PSE	
PLAY	C -> S	P,S	PLY	
PLAY_NOTIFY	S -> C	P,S	PNY	R
REDIRECT	S -> C	P,S	RDR	
SETUP	C -> S	S	STP	
SET_PARAMETER	C -> S, S -> C	P,S	SPR	R,r
TEARDOWN	C -> S	P,S	TRD	
	S -> C	P	TRD	

Table 8: Overview of RTSP methods, their direction, and what objects (P: presentation, S: stream) they operate on. Body notes if a method is allowed to carry body and in which direction, R = Request, r=response. Note: It is allowed for all error messages 4xx and 5xx to have a body

The general syntax for header fields is covered in Section 5.2. This section lists the full set of header fields along with notes on meaning, and usage. The syntax definition for header fields are present in Section 20.2.3. Throughout this section, we use [HX.Y] to informational refer to Section X.Y of the current HTTP/1.1 specification RFC 2616 [RFC2616]. Examples of each header field are given.

Information about header fields in relation to methods and proxy processing is summarized in Table 9, Table 10, Table 11, and Table 12.

The "where" column describes the request and response types in which the header field can be used. Values in this column are:

R: header field may only appear in requests;

r: header field may only appear in responses;

2xx, 4xx, etc.: A numerical value or range indicates response codes with which the header field can be used;

c: header field is copied from the request to the response.

An empty entry in the "where" column indicates that the header field may be present in both requests and responses.

The "proxy" column describes the operations a proxy may perform on a header field. An empty proxy column indicates that the proxy MUST NOT do any changes to that header, all allowed operations are explicitly stated:

a: A proxy can add or concatenate the header field if not present.

m: A proxy can modify an existing header field value.

d: A proxy can delete a header field value.

r: A proxy needs to be able to read the header field, and thus this header field cannot be encrypted.

The rest of the columns relate to the presence of a header field in a method. The method names when abbreviated, are according to Table 8:

c: Conditional; requirements on the header field depend on the context of the message.

m: The header field is mandatory.

m*: The header field SHOULD be sent, but clients/servers need to be prepared to receive messages without that header field.

o: The header field is optional.

*: The header field MUST be present if the message body is not empty. See Section 16.16, Section 16.18 and Section 5.3 for details.

-: The header field is not applicable.

"Optional" means that a Client/Server MAY include the header field in a request or response. The Client/Server behavior when receiving such headers varies, for some it may ignore the header field, in

other cases it is a request to process the header. This is regulated by the method and header descriptions. Example of headers that require processing are the Require and Proxy-Require header fields discussed in Section 16.41 and Section 16.35. A "mandatory" header field MUST be present in a request, and MUST be understood by the Client/Server receiving the request. A mandatory response header field MUST be present in the response, and the header field MUST be understood by the Client/Server processing the response. "Not applicable" means that the header field MUST NOT be present in a request. If one is placed in a request by mistake, it MUST be ignored by the Client/Server receiving the request. Similarly, a header field labeled "not applicable" for a response means that the Client/Server MUST NOT place the header field in the response, and the Client/Server MUST ignore the header field in the response.

An RTSP agent MUST ignore extension headers that are not understood.

The From and Location header fields contain an URI. If the URI contains a comma, or semicolon, the URI MUST be enclosed in double quotes ("). Any URI parameters are contained within these quotes. If the URI is not enclosed in double quote, any semicolon-delimited parameters are header-parameters, not URI parameters.

Header	Where	Pro xy	DE S	OPT	STP	PLY	PSE	TRD
Accept	R		o	-	-	-	-	-
Accept-Credentials	R	rm	o	o	o	o	o	o
Accept-Encoding	R	r	o	-	-	-	-	-
Accept-Language	R	r	o	-	-	-	-	-
Accept-Ranges	R	r	-	-	m	-	-	-
Accept-Ranges	r	r	-	-	m	-	-	-
Accept-Ranges	456	r	-	-	-	m	-	-
Allow	r	am	c	c	c	-	-	-
Allow	405	am	m	m	m	m	m	m
Authorization	R		o	o	o	o	o	o

Bandwidth	R		o	o	o	o	-	-
Blocksize	R		o	-	o	o	-	-
Cache-Control		r	o	-	o	-	-	-
Connection		ad	o	o	o	o	o	o
Connection-Credentials	470,407	ar	o	o	o	o	o	o
Content-Base	r		o	-	-	-	-	-
Content-Base	4xx,5xx		o	o	o	o	o	o
Content-Encoding	R	r	-	-	-	-	-	-
Content-Encoding	r	r	o	-	-	-	-	-
Content-Encoding	4xx,5xx	r	o	o	o	o	o	o
Content-Language	R	r	-	-	-	-	-	-
Content-Language	r	r	o	-	-	-	-	-
Content-Language	4xx,5xx	r	o	o	o	o	o	o
Content-Length	r	r	*	-	-	-	-	-
Content-Length	4xx,5xx	r	*	*	*	*	*	*
Content-Location	r	r	o	-	-	-	-	-
Content-Location	4xx,5xx	r	o	o	o	o	o	o
Content-Type	r	r	*	-	-	-	-	-
Content-Type	4xx,5xx	ar	*	*	*	*	*	*
CSeq	Rc	rm	m	m	m	m	m	m

Date		am	o/ *	o/*	o/*	o/*	o/*	o/*
Expires	r	r	o	-	-	-	-	-
From	R	r	o	o	o	o	o	o
If-Match	R	r	-	-	o	-	-	-
If-Modified-Since	R	r	o	-	o	-	-	-
If-None-Match	R	r	o	-	o	-	-	-
Last-Modified	r	r	o	-	o	-	-	-
Location	3rr		o	o	o	o	o	o

Table 9: Overview of RTSP header fields (A-L) related to methods
DESCRIBE, OPTIONS, SETUP, PLAY, PAUSE, and TEARDOWN.

Header	Where	Proxy	DES	OPT	STP	PLY	PSE	TRD
Media-Properties			-	-	m	m	m	-
Media-Range			-	-	m	m	m	-
MTag	r	r	o	-	o	-	-	-
Pipelined-Requests		amdr	-	o	o	o	o	o
Proxy-Authenticate	407	amr	m	m	m	m	m	m
Proxy-Authorization	R	rd	o	o	o	o	o	o
Proxy-Require	R	ar	o	o	o	o	o	o
Proxy-Require	r	r	c	c	c	c	c	c
Proxy-Supported	R	amr	c	c	c	c	c	c

Proxy-Supported	r		c	c	c	c	c	c
Public	r	amr	-	m	-	-	-	-
Public	501	amr	m	m	m	m	m	m
Range	R		-	-	-	o	-	-
Range	r		-	-	c	m	m	-
Referrer	R		o	o	o	o	o	o
Request-Status	R		-	-	-	-	-	-
Require	R		o	o	o	o	o	o
Retry-After	3rr,503		o	o	o	o	o	-
Retry-After	413		o	-	-	-	-	-
RTP-Info	r		-	-	c	c	-	-
Scale	R	r	-	-	-	o	-	-
Scale	r	amr	-	-	-	c	-	-
Seek-Style	R		-	-	-	o	-	-
Seek-Style	r		-	-	-	m	-	-
Server	R	r	-	o	-	-	-	o
Server	r	r	o	o	o	o	o	o
Session	R	r	-	o	o	m	m	m
Session	r	r	-	c	m	m	m	o
Speed	R	admr	-	-	-	o	-	-
Speed	r	admr	-	-	-	c	-	-
Supported	R	amr	o	o	o	o	o	o
Supported	r	amr	c	c	c	c	c	c

Terminate-Reason	R	r	-	-	-	-	-	-
Timestamp	R	admr	o	o	o	o	o	o
Timestamp	c	admr	m	m	m	m	m	m
Transport		mr	-	-	m	-	-	-
Unsupported	r		c	c	c	c	c	c
User-Agent	R		m*	m*	m*	m*	m*	m*
Vary	r		c	c	c	c	c	c
Via	R	amr	o	o	o	o	o	o
Via	c	dr	m	m	m	m	m	m
WWW-Authenticate	401		m	m	m	m	m	m

Table 10: Overview of RTSP header fields (M-W) related to methods DESCRIBE, OPTIONS, SETUP, PLAY, PAUSE, and TEARDOWN.

Header	Where	Proxy	GPR	SPR	RDR	PNY
Accept	R	arm	o	o	-	-
Accept-Credentials	R	rm	o	o	o	-
Accept-Encoding	TBD	TBD	TBD	TBD	TBD	TBD
Accept-Language	TBD	TBD	TBD	TBD	TBD	TBD
Accept-Ranges		rm	o	-	-	-
Allow	405	amr	m	m	m	-
Authorization	R		o	o	o	-
Bandwidth	R		-	o	-	-
Blocksize	R		-	o	-	-
Cache-Control		r	o	o	-	-

Connection			o	o	o	o
Connection-Credentials	470,407	ar	o	o	o	-
Content-Base	R		o	o	-	-
Content-Base	r		o	o	-	-
Content-Base	4xx,5xx		o	o	o	o
Content-Encoding	R	r	o	o	-	-
Content-Encoding	r	r	o	o	-	-
Content-Encoding	4xx,5xx	r	o	o	o	o
Content-Language	R	r	o	o	-	-
Content-Language	r	r	o	o	-	-
Content-Language	4xx,5xx	r	o	o	o	o
Content-Length	R	r	*	*	-	-
Content-Length	r	r	*	*	-	-
Content-Length	4xx,5xx	r	*	*	*	*
Content-Location	R		o	o	-	-
Content-Location	r		o	o	-	-
Content-Location	4xx,5xx		o	o	o	o
Content-Type	R		*	*	-	-
Content-Type	r		*	*	-	-
Content-Type	4xx,5xx		*	*	*	*
CSeq	R,c	mr	m	m	m	m
Date	R	a	o	o	m	o
Date	r	am	o	o	o	o
Expires	TBD	TBD	TBD	TBD	TBD	TBD

From	R	r	o	o	o	-
If-Match	TBD	TBD	TBD	TBD	TBD	TBD
If-Modified-Since	R	am	o	-	-	-
If-None-Match	R	am	o	-	-	-
Last-Modified	R	r	-	-	-	-
Last-Modified	r	r	o	-	-	-
Location	3rr		o	o	o	-
Location	R		-	-	m	-
Media-Properties	R	amr	o	-	-	c
Media-Properties	r	mr	c	-	-	-
Media-Range	R		o	-	-	c
Media-Range	r		c	-	-	-
MTag	TBD	TBD	TBD	TBD	TBD	TBD
Notify-Reason	R		-	-	-	m
Pipelined-Requests	R	amdr	o	o	-	-
Proxy-Authenticate	407	amr	m	m	m	-
Proxy-Authorization	R	rd	o	o	o	-
Proxy-Require	R	ar	o	o	o	-
Proxy-Require	r	r	c	c	c	-
Proxy-Supported	R	amr	c	c	c	-
Proxy-Supported	r		c	c	c	-
Public	501	admr	m	m	m	-

Table 11: Overview of RTSP header fields (A-P) related to methods GET_PARAMETER, SET_PARAMETER, REDIRECT, and PLAY_NOTIFY.

Header	Where	Proxy	GPR	SPR	RDR	PNY
Range	R		o	-	o	m
Referrer	R		o	o	o	-
Request-Status	R		-	-	-	c
Require	R	r	o	o	o	-
Retry-After	3rr,503		o	o	-	-
Retry-After	413		o	o	-	-
RTP-Info	R	r	o	-	-	C
RTP-Info	r	r	c	-	-	-
Scale			-	-	-	c
Seek-Style			-	-	-	-
Server	R	r	o	o	o	o
Server	r	r	o	o	-	-
Session	R	r	o	o	o	m
Session	r	r	c	c	o	m
Speed			-	-	-	-
Supported	R	adrm	o	o	o	-
Supported	r	adrm	c	c	c	-
Terminate-Reason	R	r	-	-	m	-
Timestamp	R	adrm	o	o	o	-
Timestamp	c	adrm	m	m	m	-
Transport	TBD	TBD	TBD	TBD	TBD	TBD
Unsupported	r	arm	c	c	c	-
User-Agent	R	r	m*	m*	-	-

User-Agent	r	r	m*	m*	m*	m*
Vary	r		c	c	-	-
Via	R	amr	o	o	o	-
Via	c	dr	m	m	m	-
WWW-Authenticate	401		m	m	m	-

Table 12: Overview of RTSP header fields (R-W) related to methods GET_PARAMETER, SET_PARAMETER, REDIRECT, and PLAY_NOTIFY.

16.1. Accept

The Accept request-header field can be used to specify certain presentation description and parameter media types [RFC4288] which are acceptable for the response to DESCRIBE and GET_PARAMETER requests.

See Section 20.2.3 for the syntax.

Example of use:

Accept: application/example ;q=1.0, application/sdp

16.2. Accept-Credentials

The Accept-Credentials header is a request header used to indicate to any trusted intermediary how to handle further secured connections to proxies or servers. See Section 19 for the usage of this header. It MUST NOT be included in server to client requests.

In a request the header MUST contain the method (User, Proxy, or Any) for approving credentials selected by the requester. The method MUST NOT be changed by any proxy, unless it is "Proxy" when a proxy MAY change it to "user" to take the role of user approving each further hop. If the method is "User" the header contains zero or more of credentials that the client accepts. The header may contain zero credentials in the first RTSP request to a RTSP server when using the "User" method. This is as the client has not yet received any credentials to accept. Each credential MUST consist of one URI identifying the proxy or server, the hash algorithm identifier, and the hash over that agent's DER encoded certificate [RFC5280] in Base64 [RFC4648]. All RTSP clients and proxies MUST implement the SHA-256[FIPS-pub-180-2] algorithm for computation of the hash of the DER encoded certificate. The SHA-256 algorithm is identified by the token "sha-256".

The intention with allowing for other hash algorithms is to enable the future retirement of algorithms that are not implemented somewhere else than here. Thus the definition of future algorithms for this purpose is intended to be extremely limited. A feature tag can be used to ensure that support for the replacement algorithm exist.

Example:

Accept-Credentials:User

"rtsp://proxy2.example.com/";sha-256;exaIl9VMbQMOFGClx5rXnPJKVNI=,

"rtsp://server.example.com/";sha-256;lurbjj5khhBONhIuOXtt4bBRH1M=

16.3. Accept-Encoding

The Accept-Encoding request-header field is similar to Accept, but restricts the content-codings (see Section 16.14), i.e. transformation codings of the message body, such as gzip compression, that are acceptable in the response.

A server tests whether a content-coding is acceptable, according to an Accept-Encoding field, using these rules:

1. If the content-coding is one of the content-codings listed in the Accept-Encoding field, then it is acceptable, unless it is accompanied by a qvalue of 0. (As defined in [H3.9], a qvalue of 0 means "not acceptable.")
2. The special "*" symbol in an Accept-Encoding field matches any available content-coding not explicitly listed in the header field.
3. If multiple content-codings are acceptable, then the acceptable content-coding with the highest non-zero qvalue is preferred.
4. The "identity" content-coding is always acceptable, i.e. no transformation at all, unless specifically refused because the Accept-Encoding field includes "identity;q=0", or because the field includes "*;q=0" and does not explicitly include the "identity" content-coding. If the Accept-Encoding field-value is empty, then only the "identity" encoding is acceptable.

If an Accept-Encoding field is present in a request, and if the server cannot send a response which is acceptable according to the Accept-Encoding header, then the server SHOULD send an error response with the 406 (Not Acceptable) status code.

If no Accept-Encoding field is present in a request, the server MAY assume that the client will accept any content coding. In this case,

if "identity" is one of the available content-codings, then the server SHOULD use the "identity" content-coding, unless it has additional information that a different content-coding is meaningful to the client.

16.4. Accept-Language

The Accept-Language request-header field is similar to Accept, but restricts the set of natural languages that are preferred as a response to the request. Note that the language specified applies to the presentation description and any reason phrases, but not the media content.

A language tag identifies a natural language spoken, written, or otherwise conveyed by human beings for communication of information to other human beings. Computer languages are explicitly excluded. The syntax and registry of RTSP 2.0 language tags is the same as that defined by [RFC5646].

Each language-range MAY be given an associated quality value which represents an estimate of the user's preference for the languages specified by that range. The quality value defaults to "q=1". For example:

Accept-Language: da, en-gb;q=0.8, en;q=0.7

would mean: "I prefer Danish, but will accept British English and other types of English." A language-range matches a language-tag if it exactly equals the full tag, or if it exactly equals a prefix of the tag, i.e., the primary-tag in the ABNF, such that the character following primary-tag is "-". The special range "*", if present in the Accept-Language field, matches every tag not matched by any other range present in the Accept-Language field.

Note: This use of a prefix matching rule does not imply that language tags are assigned to languages in such a way that it is always true that if a user understands a language with a certain tag, then this user will also understand all languages with tags for which this tag is a prefix. The prefix rule simply allows the use of prefix tags if this is the case.

In the process of selecting a language, each language-tag is assigned a qualification factor, i.e., if a language being supported by the client is actually supported by the server and what "preference" level the language achieves. The quality value (q-value) of the longest language-range in the field that matches the language-tag is assigned as the qualification factor for a particular language-tag. If no language-range in the field matches the tag, the language

qualification factor assigned is 0. If no Accept-Language header is present in the request, the server SHOULD assume that all languages are equally acceptable. If an Accept-Language header is present, then all languages which are assigned a qualification factor greater than 0 are acceptable.

16.5. Accept-Ranges

The Accept-Ranges general-header field allows indication of the format supported in the Range header. The client MUST include the header in SETUP requests to indicate which formats it support to receive in PLAY and PAUSE responses, and REDIRECT requests. The server MUST include the header in SETUP and 456 error responses to indicate the formats supported for the resource indicated by the request URI. The header MAY be included in GET_PARAMETER request and response pairs. The GET_PARAMETER request MUST contain a Session header to identify the session context the request is related to. The requester and responder will indicate their capabilities regarding Range formats respectively.

Accept-Ranges: NPT, SMPTE

The syntax is defined in Section 20.2.3.

16.6. Allow

The Allow message-header field lists the methods supported by the resource identified by the Request-URI. The purpose of this field is to strictly inform the recipient of valid methods associated with the resource. An Allow header field MUST be present in a 405 (Method Not Allowed) response. The Allow header MUST also be present in all OPTIONS responses where the content of the header will not include exactly the same methods as listed in the Public header.

The Allow MUST also be included in SETUP and DESCRIBE responses, if the methods allowed for the resource is different than the complete set of methods defined in this memo.

Example of use:

Allow: SETUP, PLAY, SET_PARAMETER, DESCRIBE

16.7. Authorization

An RTSP client that wishes to authenticate itself with a server using authentication mechanism from HTTP [RFC2617] , usually, but not necessarily, after receiving a 401 response, does so by including an Authorization request-header field with the request. The Authorization field value consists of credentials containing the

authentication information of the user agent for the realm of the resource being requested.

If a request is authenticated and a realm specified, the same credentials SHOULD be valid for all other requests within this realm (assuming that the authentication scheme itself does not require otherwise, such as credentials that vary according to a challenge value or using synchronized clocks).

When a shared cache (see Section 18) receives a request containing an Authorization field, it MUST NOT return the corresponding response as a reply to any other request, unless one of the following specific exceptions holds:

1. If the response includes the "max-age" cache-control directive, the cache MAY use that response in replying to a subsequent request. But (if the specified maximum age has passed) a proxy cache MUST first revalidate it with the origin server, using the request-headers from the new request to allow the origin server to authenticate the new request. (This is the defined behavior for max-age.) If the response includes "max-age=0", the proxy MUST always revalidate it before re-using it.
2. If the response includes the "must-revalidate" cache-control directive, the cache MAY use that response in replying to a subsequent request. But if the response is stale, all caches MUST first revalidate it with the origin server, using the request-headers from the new request to allow the origin server to authenticate the new request.
3. If the response includes the "public" cache-control directive, it MAY be returned in reply to any subsequent request.

16.8. Bandwidth

The Bandwidth request-header field describes the estimated bandwidth available to the client, expressed as a positive integer and measured in kilobits per second. The bandwidth available to the client may change during an RTSP session, e.g., due to mobility, congestion, etc.

Clients may not be able to accurately determine the available bandwidth, for example due to that first hop is not a bottleneck. For example most local area networks (LAN) will not be a bottleneck if the server is not in the same LAN. Thus link speeds of WLAN or Ethernet networks are normally not a basis for estimating the available bandwidth. Cellular devices or other devices directly connected to a modem or connection enabling device may more

accurately estimate the bottleneck bandwidth and what is reasonable share of it for RTSP controlled media. The client will also need to take into account other traffic sharing the bottleneck. For example by only assigning a certain fraction to RTSP and its media streams. It is RECOMMENDED that only clients that has accurate and explicit information about bandwidth bottlenecks uses this header.

This header is not a substitute for proper congestion control. Only a method providing an initial estimate and coarsely determine if the selected content can be delivered at all.

Example:

Bandwidth: 62360

16.9. Blocksize

The Blocksize request-header field is sent from the client to the media server asking the server for a particular media packet size. This packet size does not include lower-layer headers such as IP, UDP, or RTP. The server is free to use a blocksize which is lower than the one requested. The server MAY truncate this packet size to the closest multiple of the minimum, media-specific block size, or override it with the media-specific size if necessary. The block size MUST be a positive decimal number, measured in octets. The server only returns an error (4xx) if the value is syntactically invalid.

16.10. Cache-Control

The Cache-Control general-header field is used to specify directives that MUST be obeyed by all caching mechanisms along the request/response chain.

Cache directives MUST be passed through by a proxy or gateway application, regardless of their significance to that application, since the directives may be applicable to all recipients along the request/response chain. It is not possible to specify a cache-directive for a specific cache.

Cache-Control should only be specified in a DESCRIBE, GET_PARAMETER, SET_PARAMETER and SETUP request and its response. Note: Cache-Control does not govern only the caching of responses as for HTTP, instead it also applies to the media stream identified by the SETUP request. The RTSP requests are generally not cacheable, for further information see Section 18. Below is the description of the cache directives that can be included in the Cache-Control header.

no-cache: Indicates that the media stream **MUST NOT** be cached anywhere. This allows an origin server to prevent caching even by caches that have been configured to return stale responses to client requests. Note, there is no security function enforcing that the content can't be cached.

public: Indicates that the media stream is cacheable by any cache.

private: Indicates that the media stream is intended for a single user and **MUST NOT** be cached by a shared cache. A private (non-shared) cache may cache the media streams.

no-transform: An intermediate cache (proxy) may find it useful to convert the media type of a certain stream. A proxy might, for example, convert between video formats to save cache space or to reduce the amount of traffic on a slow link. Serious operational problems may occur, however, when these transformations have been applied to streams intended for certain kinds of applications. For example, applications for medical imaging, scientific data analysis and those using end-to-end authentication all depend on receiving a stream that is bit-for-bit identical to the original media stream. Therefore, if a response includes the no-transform directive, an intermediate cache or proxy **MUST NOT** change the encoding of the stream. Unlike HTTP, RTSP does not provide for partial transformation at this point, e.g., allowing translation into a different language.

only-if-cached: In some cases, such as times of extremely poor network connectivity, a client may want a cache to return only those media streams that it currently has stored, and not to receive these from the origin server. To do this, the client may include the only-if-cached directive in a request. If it receives this directive, a cache **SHOULD** either respond using a cached media stream that is consistent with the other constraints of the request, or respond with a 504 (Gateway Timeout) status. However, if a group of caches is being operated as a unified system with good internal connectivity, such a request **MAY** be forwarded within that group of caches.

max-stale: Indicates that the client is willing to accept a media stream that has exceeded its expiration time. If max-stale is assigned a value, then the client is willing to accept a response that has exceeded its expiration time by no more than the specified number of seconds. If no value is assigned to max-stale, then the client is willing to accept a stale response of any age.

min-fresh: Indicates that the client is willing to accept a media stream whose freshness lifetime is no less than its current age plus the specified time in seconds. That is, the client wants a response that will still be fresh for at least the specified number of seconds.

must-revalidate: When the must-revalidate directive is present in a SETUP response received by a cache, that cache **MUST NOT** use the entry after it becomes stale to respond to a subsequent request without first revalidating it with the origin server. That is, the cache is required to do an end-to-end revalidation every time, if, based solely on the origin server's Expires, the cached response is stale.

proxy-revalidate: The proxy-revalidate directive has the same meaning as the must-revalidate directive, except that it does not apply to non-shared user agent caches. It can be used on a response to an authenticated request to permit the user's cache to store and later return the response without needing to revalidate it (since it has already been authenticated once by that user), while still requiring proxies that service many users to revalidate each time (in order to make sure that each user has been authenticated). Note that such authenticated responses also need the public cache control directive in order to allow them to be cached at all.

max-age: When an intermediate cache is forced, by means of a max-age=0 directive, to revalidate its own cache entry, and the client has supplied its own validator in the request, the supplied validator might differ from the validator currently stored with the cache entry. In this case, the cache **MAY** use either validator in making its own request without affecting semantic transparency.

However, the choice of validator might affect performance. The best approach is for the intermediate cache to use its own validator when making its request. If the server replies with 304 (Not Modified), then the cache can return its now validated copy to the client with a 200 (OK) response. If the server replies with a new message body and cache validator, however, the intermediate cache can compare the returned validator with the one provided in the client's request, using the strong comparison function. If the client's validator is equal to the origin server's, then the intermediate cache simply returns 304 (Not Modified). Otherwise, it returns the new message body with a 200 (OK) response.

16.11. Connection

The Connection general-header field allows the sender to specify options that are desired for that particular connection and MUST NOT be communicated by proxies over further connections.

RTSP 2.0 proxies MUST parse the Connection header field before a message is forwarded and, for each connection-token in this field, remove any header field(s) from the message with the same name as the connection-token. Connection options are signaled by the presence of a connection-token in the Connection header field, not by any corresponding additional header field(s), since the additional header field may not be sent if there are no parameters associated with that connection option.

Message headers listed in the Connection header MUST NOT include end-to-end headers, such as Cache-Control.

RTSP 2.0 defines the "close" connection option for the sender to signal that the connection will be closed after completion of the response. For example, Connection: close in either the request or the response header fields indicates that the connection SHOULD NOT be considered 'persistent' (Section 10.2) after the current request/response is complete.

The use of the connection option "close" in RTSP messages SHOULD be limited to error messages when the server is unable to recover and therefore see it necessary to close the connection. The reason is that the client has the choice of continuing using a connection indefinitely, as long as it sends valid messages.

16.12. Connection-Credentials

The Connection-Credentials response header is used to carry the chain of credentials of any next hop that need to be approved by the requester. It MUST only be used in server to client responses.

The Connection-Credentials header in an RTSP response MUST, if included, contain the credential information (in form of a list of certificates providing the chain of certification) of the next hop that an intermediary needs to securely connect to. The header MUST include the URI of the next hop (proxy or server) and a base64 [RFC4648] encoded binary structure containing a sequence of DER encoded X.509v3 certificates[RFC5280] .

The binary structure starts with the number of certificates (NR_CERTS) included as a 16 bit unsigned integer. This is followed by NR_CERTS number of 16 bit unsigned integers providing the size in

octets of each DER encoded certificate. This is followed by NR_CERTS number of DER encoded X.509v3 certificates in a sequence (chain). The proxy or server's certificate must come first in the structure. Each following certificate must directly certify the one preceding it. Because certificate validation requires that root keys be distributed independently, the self-signed certificate which specifies the root certificate authority may optionally be omitted from the chain, under the assumption that the remote end must already possess it in order to validate it in any case.

Example:

Connection-Credentials:"rtsp://proxy2.example.com/";MIIDNTCC...

Where MIIDNTCC... is a BASE64 encoding of the following structure:

0																1																2																3															
0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1	2	3	4	5	6	7	8	9	0	1																																
Number of certificates																Size of certificate #1																																															
Size of certificate #2																Size of certificate #3																																															
: DER Encoding of Certificate #1																																:																															
: DER Encoding of Certificate #2																																:																															
: DER Encoding of Certificate #3																																:																															

16.13. Content-Base

The Content-Base message-header field may be used to specify the base URI for resolving relative URIs within the message body.

Content-Base: rtsp://media.example.com/movie/twister/

If no Content-Base field is present, the base URI of an message body is defined either by its Content-Location (if that Content-Location URI is an absolute URI) or the URI used to initiate the request, in that order of precedence. Note, however, that the base URI of the contents within the message-body may be redefined within that message-body.

16.14. Content-Encoding

The Content-Encoding header field is used as a modifier to the media-type. When present, its value indicates what additional content

codings have been applied to the message body, and thus what decoding mechanisms must be applied in order to obtain the media-type referenced by the Content-Type header field. Content-Encoding is primarily used to allow a document to be compressed without losing the identity of its underlying media type.

The content-coding is a characteristic of the message body identified by the Request-URI. Typically, the message body is stored with this encoding and is only decoded before rendering or analogous usage. However, a non-transparent proxy MAY modify the content-coding if the new coding is known to be acceptable to the recipient, unless the "no-transform" cache-control directive is present in the message.

If the content-coding of a message body is not "identity", then the response MUST include a Content-Encoding Message-body header that lists the non-identity content-coding(s) used.

If the content-coding of a message body in a request message is not acceptable to the origin server, the server SHOULD respond with a status code of 415 (Unsupported Media Type).

If multiple encodings have been applied to a message body, the content codings MUST be listed in the order in which they were applied, first to last from left to right. Additional information about the encoding parameters MAY be provided by other header fields not defined by this specification.

16.15. Content-Language

The Content-Language header field describes the natural language(s) of the intended audience for the enclosed message body. Note that this might not be equivalent to all the languages used within the message body.

Language tags are mentioned in Section 16.4. The primary purpose of Content-Language is to allow a user to identify and differentiate entities according to the user's own preferred language. Thus, if the body content is intended only for a Danish-literate audience, the appropriate field is

Content-Language: da

If no Content-Language is specified, the default is that the content is intended for all language audiences. This might mean that the sender does not consider it to be specific to any natural language, or that the sender does not know for which language it is intended.

Multiple languages MAY be listed for content that is intended for

multiple audiences. For example, a rendition of the "Treaty of Waitangi," presented simultaneously in the original Maori and English versions, would call for

Content-Language: mi, en

However, just because multiple languages are present within a message body does not mean that it is intended for multiple linguistic audiences. An example would be a beginner's language primer, such as "A First Lesson in Latin," which is clearly intended to be used by an English-literate audience. In this case, the Content-Language would properly only include "en".

Content-Language MAY be applied to any media type -- it is not limited to textual documents.

16.16. Content-Length

The Content-Length general-header field contains the length of the message body of the RTSP message (i.e. after the double CRLF following the last header). Unlike HTTP, it MUST be included in all messages that carry a message body beyond the header portion of the RTSP message. If it is missing, a default value of zero is assumed. Any Content-Length greater than or equal to zero is a valid value.

16.17. Content-Location

The Content-Location header field MAY be used to supply the resource location for the message body enclosed in the message when that body is accessible from a location separate from the requested resource's URI. A server SHOULD provide a Content-Location for the variant corresponding to the response message body; especially in the case where a resource has multiple variants associated with it, and those entities actually have separate locations by which they might be individually accessed, the server SHOULD provide a Content-Location for the particular variant which is returned.

The Content-Location value is not a replacement for the original requested URI; it is only a statement of the location of the resource corresponding to this particular variant at the time of the request. Future requests MAY specify the Content-Location URI as the request URI if the desire is to identify the source of that particular variant. This is useful if the RTSP agent desires to verify if the resource variant is current through a conditional request.

A cache cannot assume that a message body with a Content-Location different from the URI used to retrieve it can be used to respond to later requests on that Content-Location URI. However, the Content-

Location can be used to differentiate between multiple variants retrieved from a single requested resource.

If the Content-Location is a relative URI, the relative URI is interpreted relative to the Request-URI.

Note, that Content-Location can be used in some cases to derive the base-URI for relative URI present in session description formats. This needs to be taken into account when Content-Location is used. The easiest way to avoid needing to consider that issue is to include the Content-Base whenever the Content-Location is included.

Note also, when using Media Tags in conjunction with Content-Location it is important that the different versions have different MTags, even if provided under different Content-Location URIs. This as they have still been provided under the same request URI.

Note also, as in most cases the URI used in the DESCRIBE and the SETUP requests are different, the URI provided in a DESCRIBE Content-Location response can't directly be used in a SETUP request. Instead the extra step of resolving URIs combined with the media descriptions indication, like with SDP's a=control attribute.

16.18. Content-Type

The Content-Type header indicates the media type of the message body sent to the recipient. Note that the content types suitable for RTSP are likely to be restricted in practice to presentation descriptions and parameter-value types.

16.19. CSeq

The CSeq general-header field specifies the sequence number for an RTSP request-response pair. This field **MUST** be present in all requests and responses. For every RTSP request containing the given sequence number, the corresponding response will have the same number. Any retransmitted request **MUST** contain the same sequence number as the original (i.e., the sequence number is not incremented for retransmissions of the same request). For each new RTSP request the CSeq value **MUST** be incremented by one. The initial sequence number **MAY** be any number, however, it is **RECOMMENDED** to start at 0. Each sequence number series is unique between each requester and responder, i.e., the client has one series for its request to a server and the server has another when sending request to the client. Each requester and responder is identified with its socket address (IP address and port number).

Proxies that aggregate several sessions on the same transport will

have to ensure that the requests sent towards a particular server have a joint sequence number space, i.e., they will regularly need to renumber the CSeq header field in requests (from proxy to server) and responses (from server to proxy) to fulfill the rules for the header. The proxy **MUST** increase the CSeq by one for each request it transmits, without regard of different sessions.

Example:
CSeq: 239

16.20. Date

The Date header field represents the date and time at which the message was originated. The inclusion of the Date header in RTSP message follows these rules:

- o An RTSP message, sent either by the client or the server, containing a body **MUST** include a Date header, if the sending host has a clock;
- o Clients and servers are **RECOMMENDED** to include a Date header in all other RTSP messages, if the sending host has a clock;
- o If the server does not have a clock that can provide a reasonable approximation of the current time, its responses **MUST NOT** include a Date header field. In this case, this rule **MUST** be followed: Some origin server implementations might not have a clock available. An origin server without a clock **MUST NOT** assign Expires or Last-Modified values to a response, unless these values were associated with the resource by a system or user with a reliable clock. It **MAY** assign an Expires value that is known, at or before server configuration time, to be in the past (this allows "pre-expiration" of responses without storing separate Expires values for each resource).

A received message that does not have a Date header field **MUST** be assigned one by the recipient if the message will be cached by that recipient. An RTSP implementation without a clock **MUST NOT** cache responses without revalidating them on every use. An RTSP cache, especially a shared cache, **SHOULD** use a mechanism, such as NTP, to synchronize its clock with a reliable external standard.

The RTSP-date sent in a Date header **SHOULD NOT** represent a date and time subsequent to the generation of the message. It **SHOULD** represent the best available approximation of the date and time of message generation, unless the implementation has no means of generating a reasonably accurate date and time. In theory, the date ought to represent the moment just before the message body is

generated. In practice, the date can be generated at any time during the message origination without affecting its semantic value.

16.21. Expires

The Expires message-header field gives a date and time after which the description or media-stream should be considered stale. The interpretation depends on the method:

DESCRIBE response: The Expires header indicates a date and time after which the presentation description (body) SHOULD be considered stale.

SETUP response: The Expires header indicate a date and time after which the media stream SHOULD be considered stale.

A stale cache entry may not normally be returned by a cache (either a proxy cache or an user agent cache) unless it is first validated with the origin server (or with an intermediate cache that has a fresh copy of the message body). See Section 18 for further discussion of the expiration model.

The presence of an Expires field does not imply that the original resource will change or cease to exist at, before, or after that time.

The format is an absolute date and time as defined by RTSP-date. An example of its use is

Expires: Thu, 01 Dec 1994 16:00:00 GMT

RTSP/2.0 clients and caches MUST treat other invalid date formats, especially including the value "0", as having occurred in the past (i.e., already expired).

To mark a response as "already expired," an origin server should use an Expires date that is equal to the Date header value. To mark a response as "never expires," an origin server SHOULD use an Expires date approximately one year from the time the response is sent. RTSP/2.0 servers SHOULD NOT send Expires dates more than one year in the future.

16.22. From

The From request-header field, if given, SHOULD contain an Internet e-mail address for the human user who controls the requesting user agent. The address SHOULD be machine-usable, as defined by "mailbox" in [RFC1123].

This header field MAY be used for logging purposes and as a means for identifying the source of invalid or unwanted requests. It SHOULD NOT be used as an insecure form of access protection. The interpretation of this field is that the request is being performed on behalf of the person given, who accepts responsibility for the method performed. In particular, robot agents SHOULD include this header so that the person responsible for running the robot can be contacted if problems occur on the receiving end.

The Internet e-mail address in this field MAY be separate from the Internet host which issued the request. For example, when a request is passed through a proxy the original issuer's address SHOULD be used.

The client SHOULD NOT send the From header field without the user's approval, as it might conflict with the user's privacy interests or their site's security policy. It is strongly recommended that the user be able to disable, enable, and modify the value of this field at any time prior to a request.

16.23. If-Match

The If-Match request-header field is especially useful for ensuring the integrity of the presentation description, independent of how the presentation description was received. The presentation description can be fetched via means external to RTSP (such as HTTP) or via the DESCRIBE message. In the case of retrieving the presentation description via RTSP, the server implementation is guaranteeing the integrity of the description between the time of the DESCRIBE message and the SETUP message. By including the MTag given in or with the session description in an If-Match header part of the SETUP request, the client ensures that resources set up are matching the description. A SETUP request with the If-Match header for which the MTag validation check fails, MUST response using 412 (Precondition Failed).

This validation check is also very useful if a session has been redirected from one server to another.

16.24. If-Modified-Since

The If-Modified-Since request-header field is used with the DESCRIBE and SETUP methods to make them conditional. If the requested variant has not been modified since the time specified in this field, a description will not be returned from the server (DESCRIBE) or a stream will not be set up (SETUP). Instead, a 304 (Not Modified) response MUST be returned without any message-body.

An example of the field is:

 If-Modified-Since: Sat, 29 Oct 1994 19:43:31 GMT

16.25. If-None-Match

This request header can be used with one or several message body tags to make DESCRIBE requests conditional. A client that has one or more message bodies previously obtained from the resource, can verify that none of those entities is current by including a list of their associated message body tags in the If-None-Match header field. The purpose of this feature is to allow efficient updates of cached information with a minimum amount of transaction overhead. As a special case, the value "*" matches any current entity of the resource.

If any of the message body tags match the message body tag of the message body that would have been returned in the response to a similar DESCRIBE request (without the If-None-Match header) on that resource, or if "*" is given and any current entity exists for that resource, then the server **MUST NOT** perform the requested method, unless required to do so because the resource's modification date fails to match that supplied in an If-Modified-Since header field in the request. Instead, if the request method was DESCRIBE, the server **SHOULD** respond with a 304 (Not Modified) response, including the cache-related header fields (particularly MTag) of one of the message bodies that matched. For all other request methods, the server **MUST** respond with a status of 412 (Precondition Failed).

See Section 18.1.3 for rules on how to determine if two message body tags match.

If none of the message body tags match, then the server **MAY** perform the requested method as if the If-None-Match header field did not exist, but **MUST** also ignore any If-Modified-Since header field(s) in the request. That is, if no message body tags match, then the server **MUST NOT** return a 304 (Not Modified) response.

If the request would, without the If-None-Match header field, result in anything other than a 2xx or 304 status, then the If-None-Match header **MUST** be ignored. (See Section 18.1.4 for a discussion of server behavior when both If-Modified-Since and If-None-Match appear in the same request.)

The result of a request having both an If-None-Match header field and an If-Match header field is unspecified and **MUST** be considered an illegal request.

16.26. Last-Modified

The Last-Modified message-header field indicates the date and time at which the origin server believes the presentation description or media stream was last modified. For the method DESCRIBE, the header field indicates the last modification date and time of the description, for SETUP that of the media stream.

An origin server **MUST NOT** send a Last-Modified date which is later than the server's time of message origination. In such cases, where the resource's last modification would indicate some time in the future, the server **MUST** replace that date with the message origination date.

An origin server **SHOULD** obtain the Last-Modified value of the message body as close as possible to the time that it generates the Date value of its response. This allows a recipient to make an accurate assessment of the message body's modification time, especially if the message body changes near the time that the response is generated.

RTSP servers **SHOULD** send Last-Modified whenever feasible.

16.27. Location

The Location response-header field is used to redirect the recipient to a location other than the Request-URI for completion of the request or identification of a new resource. For 3xx responses, the location **SHOULD** indicate the server's preferred URI for automatic redirection to the resource. The field value consists of a single absolute URI.

Note: The Content-Location header field (Section 16.17) differs from Location in that the Content-Location identifies the original location of the message body enclosed in the request. It is therefore possible for a response to contain header fields for both Location and Content-Location. Also, see Section 18.2 for cache requirements of some methods.

16.28. Media-Properties

This general header is used in SETUP response or PLAY_NOTIFY requests to indicate the media's properties that currently are applicable to the RTSP session. PLAY_NOTIFY **MAY** be used to modify these properties at any point. However, the client **SHOULD** have received the update prior to any action related to the new media properties take effect. For aggregated sessions, the Media-Properties header will be returned in each SETUP response. The header received in the latest response is the one that applies on the whole session from this point until

any future update. The header MAY be included without value in GET_PARAMETER requests to the server with a Session header included to query the current Media-Properties for the session. The responder MUST include the current session's media properties.

The media properties expressed by this header is the one applicable to all media in the RTSP session. For aggregated sessions, the header expressed the combined media-properties. As a result, aggregation of media MAY result in a change of the media properties, and thus the content of the Media-Properties header contained in subsequent SETUP responses.

The header contains a list of property values that are applicable to the currently setup media or aggregate of media as indicated by the RTSP URI in the request. No ordering is enforced within the header. Property values should be grouped into a single group that handles a particular orthogonal property. Values or groups that express multiple properties SHOULD NOT be used. The list of properties that can be expressed MAY be extended at any time. Unknown property values MUST be ignored.

This specification defines the following 4 groups and their property values:

Random Access:

Random-Access: Indicates that random access is possible. May optionally include a floating point value in seconds indicating the longest duration between any two random access points in the media.

Beginning-Only: Seeking is limited to the beginning only.

No-Seeking: No seeking is possible.

Content Modifications:

Immutable: The content will not be changed during the life-time of the RTSP session.

Dynamic: The content may be changed based on external methods or triggers

Time-Progressing The media accessible progresses as wallclock time progresses.

Retention:

Unlimited: Content will be retained for the duration of the lifetime of the RTSP session.

Time-Limited: Content will be retained at least until the specified wallclock time. The time must be provided in the absolute time format specified in Section 4.6.

Time-Duration Each individual media unit is retained for at least the specified time duration. This definition allows for retaining data with a time based sliding window. The time duration is expressed as floating point number in seconds. 0.0 is a valid value as this indicates that no data is retained in a time-progressing session.

Supported Scale:

Scales: A quoted comma separated list of one or more decimal values or ranges of scale values supported by the content in arbitrary order. A range has a start and stop value separated by a colon. A range indicates that the content supports fine grained selection of scale values. Fine grained allows for steps at least as small as one tenth of a scale value. Negative values are supported. The value 0 has no meaning and MUST NOT be used.

Examples of this header for on-demand content and a live stream without recording are:

On-demand:

Media-Properties: Random-Access=2.5s, Unlimited, Immutable,
Scales="-20, -10, -4, 0.5:1.5, 4, 8, 10, 15, 20"

Live stream without recording/timeshifting:

Media-Properties: No-Seeking, Time-Progressing, Time-Duration=0.0

16.29. Media-Range

The Media-Range general header is used to give the range of the media at the time of sending the RTSP message. This header MUST be included in SETUP response, and PLAY and PAUSE response for media that are Time-Progressing, and PLAY and PAUSE response after any change for media that are Dynamic, and in PLAY_NOTIFY request that are sent due to Media-Property-Update. Media-Range header without any range specifications MAY be included in GET_PARAMETER requests to the server to request the current range. The server MUST in this case include the current range at the time of sending the response.

The header **MUST** include range specifications for all time formats supported for the media, as indicated in Accept-Ranges header (Section 16.5) when setting up the media. The server **MAY** include more than one range specification of any given time format to indicate media that has non-continuous range.

For media that has the Time-Progressing property, the Media-Range values will only be valid for the particular point in time when it was issued. As wallclock progresses so will also the media range. However, it shall be assumed that media time progresses in direct relationship to wallclock time (with the exception of clock skew) so that a reasonably accurate estimation of the media range can be calculated.

16.30. MTag

The MTag response header **MAY** be included in DESCRIBE, GET_PARAMETER or SETUP responses. The message body tags (Section 4.8) returned in a DESCRIBE response, and the one in SETUP refers to the presentation, i.e. both the returned session description and the media stream. This allows for verification that one has the right session description to a media resource at the time of the SETUP request. However, it has the disadvantage that a change in any of the parts results in invalidation of all the parts.

If the MTag is provided both inside the message body, e.g. within the "a=mtag" attribute in SDP, and in the response message, then both tags **MUST** be identical. It is **RECOMMENDED** that the MTag is primarily given in the RTSP response message, to ensure that caches can use the MTag without requiring content inspection. However, for session descriptions that are distributed outside of RTSP, for example using HTTP, etc. it will be necessary to include the message body tag in the session description as specified in Appendix D.1.9.

SETUP and DESCRIBE requests can be made conditional upon the MTag using the headers If-Match (Section 16.23) and If-None-Match (Section 16.25).

16.31. Notify-Reason

The Notify Reason header is solely used in the PLAY_NOTIFY method. It indicates the reason why the server has sent the asynchronous PLAY_NOTIFY request (see Section 13.5).

16.32. Pipelined-Requests

The Pipelined-Requests general header is used to indicate that a request is to be executed in the context created by a previous

request(s). The primary usage of this header is to allow pipelining of SETUP requests so that any additional SETUP request after the first one does not need to wait for the session ID to be sent back to the requesting agent. The header contains a unique identifier that is scoped by the persistent connection used to send the requests.

Upon receiving a request with the Pipelined-Requests the responding agent MUST look up if there exists a binding between this Pipelined-Requests identifier for the current persistent connection and an RTSP session ID. If that exists then the received request is processed the same way as if it contained the Session header with the found session ID. If there does not exist a mapping and no Session header is included in the request, the responding agent MUST create a binding upon the successful completion of a session creating request, i.e. SETUP. A binding MUST NOT be created, if the request failed to create an RTSP session. In case the request contains both a Session header and the Pipelined-Requests header the Pipelined-Requests MUST be ignored.

Note: Based on the above definition at least the first request containing a new unique Pipelined-Requests will be required to be a SETUP request (unless the protocol is extended with new methods of creating a session). After that first one, additional SETUP requests or request of any type using the RTSP session context may include the Pipelined-Requests header.

When responding to any request that contained the Pipelined-Requests header the server MUST also include the Session header when a binding to a session context exist. An RTSP agent that knows the session ID SHOULD NOT use the Pipelined-Requests header in any request and only use the Session header. This as the Session identifier is persistent across transport contexts, like TCP connections, which the Pipelined-Requests identifier is not.

The RTSP agent sending the request with a Pipelined-Requests header has the responsibility for using a unique and previously unused identifier within the transport context. Currently only a TCP connection is defined as such transport context. A server MUST delete the Pipelined-Requests identifier and its binding to a session upon the termination of that session. Despite the previous mandate, RTSP agents are RECOMMENDED to not reuse identifiers to allow for better error handling and logging.

RTSP Proxies may need to translate Pipelined-Requests identifier values from incoming request to outgoing to allow for aggregation of requests onto a persistent connection.

16.33. Proxy-Authenticate

The Proxy-Authenticate response-header field MUST be included as part of a 407 (Proxy Authentication Required) response. The field value consists of a challenge that indicates the authentication scheme and parameters applicable to the proxy for this Request-URI.

The HTTP access authentication process is described in [RFC2617]. Unlike WWW-Authenticate, the Proxy-Authenticate header field applies only to the current connection and SHOULD NOT be passed on to downstream agents. However, an intermediate proxy might need to obtain its own credentials by requesting them from the downstream agent, which in some circumstances will appear as if the proxy is forwarding the Proxy-Authenticate header field.

16.34. Proxy-Authorization

The Proxy-Authorization request-header field allows the client to identify itself (or its user) to a proxy which requires authentication. The Proxy-Authorization field value consists of credentials containing the authentication information of the user agent for the proxy and/or realm of the resource being requested.

The HTTP access authentication process is described in [RFC2617]. Unlike Authorization, the Proxy-Authorization header field applies only to the next outbound proxy that demanded authentication using the Proxy-Authenticate field. When multiple proxies are used in a chain, the Proxy-Authorization header field is consumed by the first outbound proxy that was expecting to receive credentials. A proxy MAY relay the credentials from the client request to the next proxy if that is the mechanism by which the proxies cooperatively authenticate a given request.

16.35. Proxy-Require

The Proxy-Require request-header field is used to indicate proxy-sensitive features that MUST be supported by the proxy. Any Proxy-Require header features that are not supported by the proxy MUST be negatively acknowledged by the proxy to the client using the Unsupported header. The proxy MUST use the 551 (Option Not Supported) status code in the response. Any feature-tag included in the Proxy-Require does not apply to the end-point (server or client). To ensure that a feature is supported by both proxies and servers the tag needs to be included in also a Require header.

See Section 16.41 for more details on the mechanics of this message and a usage example. See discussion in the proxies section (Section 17.1) about when to consider that a feature requires proxy

support.

Example of use:

Proxy-Require: play.basic

16.36. Proxy-Supported

The Proxy-Supported header field enumerates all the extensions supported by the proxy using feature-tags. The header carries the intersection of extensions supported by the forwarding proxies. The Proxy-Supported header MAY be included in any request by a proxy. It MUST be added by any proxy if the Supported header is present in a request. When present in a request, the receiver MUST in the response copy the received Proxy-Supported header.

The Proxy-Supported header field contains a list of feature-tags applicable to proxies, as described in Section 4.7. The list is the intersection of all feature-tags understood by the proxies. To achieve an intersection, the proxy adding the Proxy-Supported header includes all proxy feature-tags it understands. Any proxy receiving a request with the header, MUST check the list and removes any feature-tag(s) it does not support. A Proxy-Supported header present in the response MUST NOT be touched by the proxies.

Example:

```
C->P1: OPTIONS rtsp://example.com/ RTSP/2.0
      Supported: foo, bar, blech
      User-Agent: PhonyClient/1.2
```

```
P1->P2: OPTIONS rtsp://example.com/ RTSP/2.0
      Supported: foo, bar, blech
      Proxy-Supported: proxy-foo, proxy-bar, proxy-blech
      Via: 2.0 pro.example.com
```

```
P2->S:  OPTIONS rtsp://example.com/ RTSP/2.0
      Supported: foo, bar, blech
      Proxy-Supported: proxy-foo, proxy-blech
      Via: 2.0 pro.example.com, 2.0 prox2.example.com
```

```
S->C:  RTSP/2.0 200 OK
      Supported: foo, bar, baz
      Proxy-Supported: proxy-foo, proxy-blech
      Public: OPTIONS, SETUP, PLAY, PAUSE, TEARDOWN
      Via: 2.0 pro.example.com, 2.0 prox2.example.com
```

16.37. Public

The Public response header field lists the set of methods supported by the response sender. This header applies to the general capabilities of the sender and its only purpose is to indicate the sender's capabilities to the recipient. The methods listed may or may not be applicable to the Request-URI; the Allow header field (Section 16.6) MAY be used to indicate methods allowed for a particular URI.

Example of use:

Public: OPTIONS, SETUP, PLAY, PAUSE, TEARDOWN

In the event that there are proxies between the sender and the recipient of a response, each intervening proxy MUST modify the Public header field to remove any methods that are not supported via that proxy. The resulting Public header field will contain an intersection of the sender's methods and the methods allowed through by the intervening proxies.

In general, proxies should allow all methods to transparently pass through from the sending RTSP agent to the receiving RTSP agent, but there may be cases where this is not desirable for a given proxy. Modification of the Public response header field by the intervening proxies ensures that the request sender gets an accurate response indicating the methods that can be used on the target agent via the proxy chain.

16.38. Range

The Range header specifies a time range in PLAY (Section 13.4), PAUSE (Section 13.6), SETUP (Section 13.3), REDIRECT (Section 13.10), and PLAY_NOTIFY (Section 13.5) requests and responses. It MAY be included in GET_PARAMETER requests from the client to the server with only a Range format and no value to request the current media position, whether the session is in Play or Ready state in the included format. The server SHALL, if supporting the range format, respond with the current playing point or pause point as the start of the range. If an explicit stop point was used in the previous PLAY request, then that value shall be included as stop point. Note that if the server is currently under any type of media playback manipulation affecting the interpretation of Range, like Scale, that is also required to be included in any GET_PARAMETER response to provide complete information.

The range can be specified in a number of units. This specification defines smpte (Section 4.4), npt (Section 4.5), and clock (Section 4.6) range units. While byte ranges [H14.35.1] and other

extended units MAY be used, their behavior is unspecified since they are not normally meaningful in RTSP. Servers supporting the Range header MUST understand the NPT range format and SHOULD understand the SMPTE range format. If the Range header is sent in a time format that is not understood, the recipient SHOULD return 456 (Header Field Not Valid for Resource) and include an Accept-Ranges header indicating the supported time formats for the given resource.

Example:

Range: clock=19960213T143205Z-

The Range header contains a range of one single range format. A range is a half-open interval with a start and an end point, including the start point, but excluding the end point. A range may either be fully specified with explicit values for start point and end point, or have either start or end point be implicit. An implicit start point indicates the session's pause point, and if no pause point is set the start of the content. An implicit end point indicates the end of the content. The usage of both implicit start and end point is not allowed in the same range header, however, the exclusion of the range header has that meaning, i.e. from pause point (or start) until end of content.

Regarding the half-open intervals; a range of A-B starts exactly at time A, but ends just before B. Only the start time of a media unit such as a video or audio frame is relevant. For example, assume that video frames are generated every 40 ms. A range of 10.0-10.1 would include a video frame starting at 10.0 or later time and would include a video frame starting at 10.08, even though it lasted beyond the interval. A range of 10.0-10.08, on the other hand, would exclude the frame at 10.08.

Please note the difference between NPT time scales' "now" and an implicit start value. Implicit value reference the current pause-point. While "now" is the currently ongoing time. In a time-progressing session with recording (retention for some or full time) the pause point may be 2 min into the session while now could be 1 hour into the session.

By default, range intervals increase, where the second point is larger than the first point.

Example:

Range: npt=10-15

However, range intervals can also decrease if the Scale header (see Section 16.44) indicates a negative scale value. For example, this would be the case when a playback in reverse is desired.

Example:

Scale: -1
Range: npt=15-10

Decreasing ranges are still half open intervals as described above. Thus, for range A-B, A is closed and B is open. In the above example, 15 is closed and 10 is open. An exception to this rule is the case when B=0 in a decreasing range. In this case, the range is closed on both ends, as otherwise there would be no way to reach 0 on a reverse playback for formats that have such a notion, like NPT and SMPTE.

Example:

Scale: -1
Range: npt=15-0

In this range both 15 and 0 are closed.

A decreasing range interval without a corresponding negative Scale header is not valid.

16.39. Referrer

The Referrer request-header field allows the client to specify, for the server's benefit, the address (URI) of the resource from which the Request-URI was obtained. The URI refers to that of the presentation description, typically retrieved via HTTP. The Referrer request-header allows a server to generate lists of back-links to resources for interest, logging, optimized caching, etc. It also allows obsolete or mistyped links to be traced for maintenance. The Referrer field MUST NOT be sent if the Request-URI was obtained from a source that does not have its own URI, such as input from the user keyboard.

If the field value is a relative URI, it SHOULD be interpreted relative to the Request-URI. The URI MUST NOT include a fragment.

Because the source of a link might be private information or might reveal an otherwise private information source, it is strongly recommended that the user be able to select whether or not the Referrer field is sent. For example, a streaming client could have a toggle switch for openly/anonymously, which would respectively enable/disable the sending of Referrer and From information.

Clients SHOULD NOT include a Referrer header field in a (non-secure) RTSP request if the referring page was transferred with a secure protocol.

16.40. Request-Status

This request header is used to indicate the end result for requests that takes time to complete, such a PLAY (Section 13.4). It is sent in PLAY_NOTIFY (Section 13.5) with the end-of-stream reason to report how the PLAY request concluded, either in success or in failure. The header carries a reference to the request it reports on using the CSeq number for the session indicated by the Session header in the request. It provides both a numerical status code (according to Section 8.1.1) and a human readable reason phrase.

Example:

Request-Status: cseq=63 status=500 reason="Media data unavailable"

16.41. Require

The Require request-header field is used by clients or servers to ensure that the other end-point supports features that are required in respect to this request. It can also be used to query if the other end-point supports certain features, however, the use of the Supported (Section 16.49) is much more effective in this purpose. The server MUST respond to this header by using the Unsupported header to negatively acknowledge those feature-tags which are NOT supported. The response MUST use the error code 551 (Option Not Supported). This header does not apply to proxies, for the same functionality in respect to proxies see Proxy-Require header (Section 16.35) with the exception of media modifying proxies. Media modifying proxies, due to their nature of handling media in a way that is very similar to a server, do need to understand also the server features to correctly serve the client.

This is to make sure that the client-server interaction will proceed without delay when all features are understood by both sides, and only slow down if features are not understood (as in the example below). For a well-matched client-server pair, the interaction proceeds quickly, saving a round-trip often required by negotiation mechanisms. In addition, it also removes state ambiguity when the client requires features that the server does not understand.

Example (Not complete):

```
C->S:  SETUP rtsp://server.com/foo/bar/baz.rm RTSP/2.0
        CSeq: 302
        Require: funky-feature
        Funky-Parameter: funkystuff

S->C:  RTSP/2.0 551 Option not supported
        CSeq: 302
        Unsupported: funky-feature
```

In this example, "funky-feature" is the feature-tag which indicates to the client that the fictional Funky-Parameter field is required. The relationship between "funky-feature" and Funky-Parameter is not communicated via the RTSP exchange, since that relationship is an immutable property of "funky-feature" and thus should not be transmitted with every exchange.

Proxies and other intermediary devices **MUST** ignore this header. If a particular extension requires that intermediate devices support it, the extension should be tagged in the Proxy-Require field instead (see Section 16.35). See discussion in the proxies section (Section 17.1) about when to consider that a feature requires proxy support.

16.42. Retry-After

The Retry-After response-header field can be used with a 503 (Service Unavailable) response to indicate how long the service is expected to be unavailable to the requesting client. This field **MAY** also be used with any 3xx (Redirection) response to indicate the minimum time the user-agent is asked to wait before issuing the redirected request. The value of this field can be either an RTSP-date or an integer number of seconds (in decimal) after the time of the response.

```
Example:
Retry-After: Fri, 31 Dec 1999 23:59:59 GMT
Retry-After: 120
```

In the latter example, the delay is 2 minutes.

16.43. RTP-Info

The RTP-Info general header field is used to set RTP-specific parameters in the PLAY and GET_PARAMETER responses or a PLAY_NOTIFY and GET_PARAMETER requests. For streams using RTP as transport protocol the RTP-Info header **SHOULD** be part of a 200 response to PLAY.

The exclusion of the RTP-Info in a PLAY response for RTP transported media will result in that a client needs to synchronize the media streams using RTCP. This may have negative impact as the RTCP can be lost, and does not need to be particularly timely in its arrival. Also functionality as informing the client from which packet a seek has occurred is affected.

The RTP-Info MAY be included in SETUP responses to provide synchronization information when changing transport parameters, see Section 13.3. The RTP-Info header and the Range header MAY be included in a GET_PARAMETER request from client to server without any values to request the current playback point and corresponding RTP synchronization information. When the RTP-Info header is included in a Request also the Range header MUST be included (Note, Range header only MAY be used). The server response SHALL include both the Range header and the RTP-Info header. If the session is in Play state, then the value of the Range header SHALL be filled in with the current playback point and with the corresponding RTP-Info values. If the server is another state, no values are included in the RTP-Info header. The header is included in PLAY_NOTIFY requests with the Notify-Reason of end-of-stream to provide RTP information about the end of the stream.

The header can carry the following parameters:

- url: Indicates the stream URI for which the following RTP parameters correspond, this URI MUST be the same as used in the SETUP request for this media stream. Any relative URI MUST use the Request-URI as base URI. This parameter MUST be present.
- ssrc: The Synchronization source (SSRC) that the RTP timestamp and sequence number provided applies to. This parameter MUST be present.
- seq: Indicates the sequence number of the first packet of the stream that is direct result of the request. This allows clients to gracefully deal with packets when seeking. The client uses this value to differentiate packets that originated before the seek from packets that originated after the seek. Note that a client may not receive the packet with the expressed sequence number, and instead packets with a higher sequence number, due to packet loss or reordering. This parameter is RECOMMENDED to be present.

rtptime: MUST indicate the RTP timestamp value corresponding to the start time value in the Range response header, or if not explicitly given the implied start point. The client uses this value to calculate the mapping of RTP time to NPT or other media timescale. This parameter SHOULD be present to ensure inter-media synchronization is achieved. There exists no requirement that any received RTP packet will have the same RTP timestamp value as the one in the parameter used to establish synchronization.

A mapping from RTP timestamps to NTP timestamps (wallclock) is available via RTCP. However, this information is not sufficient to generate a mapping from RTP timestamps to media clock time (NPT, etc.). Furthermore, in order to ensure that this information is available at the necessary time (immediately at startup or after a seek), and that it is delivered reliably, this mapping is placed in the RTSP control channel.

In order to compensate for drift for long, uninterrupted presentations, RTSP clients should additionally map NPT to NTP, using initial RTCP sender reports to do the mapping, and later reports to check drift against the mapping.

Example:

```
Range:npt=3.25-15
RTP-Info:url="rtsp://example.com/foo/audio" ssrc=0A13C760:seq=45102;
          rtptime=12345678,url="rtsp://example.com/foo/video"
          ssrc=9A9DE123:seq=30211;rtptime=29567112
```

Lets assume that Audio uses a 16kHz RTP timestamp clock and Video a 90kHz RTP timestamp clock. Then the media synchronization is depicted in the following way.

NPT	3.0---	3.1---	3.2-X-	3.3---	3.4---	3.5---	3.6
Audio			PA A				
Video			V		PV		

X: NPT time value = 3.25, from Range header.
A: RTP timestamp value for Audio from RTP-Info header (12345678).
V: RTP timestamp value for Video from RTP-Info header (29567112).
PA: RTP audio packet carrying an RTP timestamp of 12344878. Which corresponds to NPT = $(12344878 - A) / 16000 + 3.25 = 3.2$
PV: RTP video packet carrying an RTP timestamp of 29573412. Which corresponds to NPT = $(29573412 - V) / 90000 + 3.25 = 3.32$

16.44. Scale

A scale value of 1 indicates normal play at the normal forward viewing rate. If not 1, the value corresponds to the rate with respect to normal viewing rate. For example, a ratio of 2 indicates twice the normal viewing rate ("fast forward") and a ratio of 0.5 indicates half the normal viewing rate. In other words, a ratio of 2 has content time increase at twice the playback time. For every second of elapsed (wallclock) time, 2 seconds of content time will be delivered. A negative value indicates reverse direction. For certain media transports this may require certain considerations to work consistent, see Appendix C.1 for description on how RTP handles this.

The transmitted data rate SHOULD NOT be changed by selection of a different scale value. The resulting bit-rate should be reasonably close to the nominal bit-rate of the content for Scale = 1. The server has to actively manipulate the data when needed to meet the bitrate constraints. Implementation of scale changes depends on the server and media type. For video, a server may, for example, deliver only key frames or selected frames. For audio, it may time-scale the audio while preserving pitch or, less desirably, deliver fragments of audio, or completely mute the audio.

The server and content may restrict the range of scale values that it supports. The supported values are indicated by the Media-Properties header (Section 16.28). The client SHOULD only indicate values indicated to be supported. However, as the values may change as the content progresses a requested value may no longer be valid when the request arrives. Thus, a non-supported value in a request does not generate an error, only forces the server to choose the closest value. The response MUST always contain the actual scale value chosen by the server.

If the server does not implement the possibility to scale, it will not return a Scale header. A server supporting Scale operations for PLAY MUST indicate this with the use of the "play.scale" feature-tag.

When indicating a negative scale for a reverse playback, the Range header MUST indicate a decreasing range as described in Section 16.38.

Example of playing in reverse at 3.5 times normal rate:

Scale: -3.5
Range: npt=15-10

16.45. Seek-Style

When a client sends a PLAY request with a Range header to perform a random access to the media, the client does not know if the server will pick the first media samples or the first random access point prior to the request range. Depending on use case, the client may have a strong preference. To express this preference and provide the client with information on how the server actually acted on that preference the Seek-Style header is defined.

Seek-Style is a general header that MAY be included in any PLAY request to indicate the client's preference for any media stream that has random access properties. The server MUST always include the header in any PLAY response for media with random access properties to indicate what policy was applied. A server that receives an unknown Seek-Style policy MUST ignore it and select the server default policy. A client receiving an unknown policy MUST ignore it and use the Range header and any media synchronization information as basis to determine what the server did.

This specification defines the following seek policies that may be requested (see also Section 4.9.1):

RAP: Random Access Point (RAP) is the behavior of requesting the server to locate the closest previous random access point that exists in the media aggregate and deliver from that. By requesting a RAP, media quality will be the best possible as all media will be delivered from a point where full media state can be established in the media decoder.

CoRAP: Conditional Random Access Point (CoRAP) is a variant of the above RAP behavior. This policy is primarily intended for cases where there is larger distance between the random access points in the media. CoRAP is conditioned on that there is a Random Access Point closer to the requested start point than to the current pause point. This policy assumes that the media state existing prior to the pause is usable if delivery is continued. If the client or server knows that this is not the fact the RAP policy should be used. In other words: in most cases when the client requests a start point prior to the current pause point, a valid decoding dependency chain from the media delivered prior to the pause and to the requested media unit will not exist. If the server searched to a random access point the server MUST return the CoRAP policy in the Seek-Style header and adjust the Range header to reflect the position of the picked RAP. In case the random access point is further away and the server selects to continue from the current pause point it MUST include the "Next" policy in the Seek-Style header and adjust the Range header start

point to the current pause point.

First-Prior: The first-prior policy will start delivery with the media unit that has a playout time first prior to the requested time. For discrete media that would only include media units that would still be rendered at the request time. For continuous media that is media that will be rendered during the requested start time of the range.

Next: The next media units after the provided start time of the range. For continuous framed media that would mean the first next frame after the provided time. For discrete media the first unit that is to be rendered after the provided time. The main usage for this case is when the client knows it has all media up to a certain point and would like to continue delivery so that a complete non-interrupted media playback can be achieved. Example of such scenarios include switching from a broadcast/multicast delivery to a unicast based delivery. This policy **MUST** only be used on the client's explicit request.

Please note that these expressed preferences exist for optimizing the startup time or the media quality. The "Next" policy breaks the normal definition of the Range header to enable a client to request media with minimal overlap, although some may still occur for aggregated sessions. RAP and First-Prior both fulfill the requirement of providing media from the requested range and forward. However, unless RAP is used, the media quality for many media codecs using predictive methods can be severely degraded unless additional data is available as, for example, already buffered, or through other side channels.

16.46. Server

The Server response-header field contains information about the software used by the origin server to handle the request. The field can contain multiple product tokens and comments identifying the server and any significant subproducts. The product tokens are listed in order of their significance for identifying the application.

Example:

Server: PhonyServer/1.0

If the response is being forwarded through a proxy, the proxy application **MUST NOT** modify the Server response-header. Instead, it **SHOULD** include a Via field (Section 16.56). If the response is generated by the proxy, the proxy application **MUST** return the Server response-header as previously returned by the server.

16.47. Session

The Session request-header and response-header field identifies an RTSP session. An RTSP session is created by the server as a result of a successful SETUP request and in the response the session identifier is given to the client. The RTSP session exists until destroyed by a TEARDOWN, REDIRECT or timed out by the server.

The session identifier is chosen by the server (see Section 4.3) and MUST be returned in the SETUP response. Once a client receives a session identifier, it MUST be included in any request related to that session. This means that the Session header MUST be included in a request, using the following methods: PLAY, PAUSE, and TEARDOWN, and MAY be included in SETUP, OPTIONS, SET_PARAMETER, GET_PARAMETER, and REDIRECT, and MUST NOT be included in DESCRIBE. The Session header MUST NOT be included in the following methods, if these requests are pipelined and if the session identifier is not yet known: PLAY, PAUSE, TEARDOWN, SETUP, OPTIONS SET_PARAMETER, and GET_PARAMETER.

In an RTSP response the session header MUST be included in methods, SETUP, PLAY, and PAUSE, and MAY be included in methods, TEARDOWN, and REDIRECT, and if included in the request of the following methods it MUST also be included in the response, OPTIONS, GET_PARAMETER, and SET_PARAMETER, and MUST NOT be included in DESCRIBE responses.

Note that a session identifier identifies an RTSP session across transport sessions or connections. RTSP requests for a given session can use different URIs (Presentation and media URIs). Note, that there are restrictions depending on the session which URIs that are acceptable for a given method. However, multiple "user" sessions for the same URI from the same client will require use of different session identifiers.

The session identifier is needed to distinguish several delivery requests for the same URI coming from the same client.

The response 454 (Session Not Found) MUST be returned if the session identifier is invalid.

The header MAY include the session timeout period. If not explicitly provided this value is set to 60 seconds. As this affects how often session keep-alives are needed values smaller than 30 seconds are not recommended. However, larger than default values can be useful in applications of RTSP that have inactive but established sessions for longer time periods.

60 seconds was chosen as session timeout value due to: Resulting in not too frequent keep-alive messages and having low sensitivity to variations in request response timing. If one reduces the timeout value to below 30 seconds the corresponding request response timeout becomes a significant part of the session timeout. 60 seconds also allows for reasonably rapid recovery of committed server resources in case of client failure.

16.48. Speed

The Speed request-header field requests the server to deliver specific amounts of nominal media time per unit of delivery time, contingent on the server's ability and desire to serve the media stream at the given speed. The client requests the delivery speed to be within a given range with a lower and upper bound. The server SHALL deliver at the highest possible speed within the range, but not faster than the upper-bound, for which the underlying network path can support the resulting transport data rates. As long as any speed value within the given range can be provided the server SHALL NOT modify the media quality. Only if the server is unable to deliver media at the speed value provided by the lower bound shall it reduce the media quality.

Implementation of the Speed functionality by the server is OPTIONAL. The server can indicate its support through a feature-tag, play.speed. The lack of a Speed header in the response is an indication of lack of support of this functionality.

The speed parameter values are expressed as a positive decimal value, e.g., a value of 2.0 indicates that data is to be delivered twice as fast as normal. A speed value of zero is invalid. The range is specified in the form "lower bound - upper bound". The lower bound value may be smaller or equal to the upper bound. All speeds may not be possible to support. Therefore the server MAY modify the requested values to the closest supported. The actual supported speed MUST be included in the response. Note, however, that the use cases may vary and that Speed value ranges such as 0.7 - 0.8, 0.3-2.0, 1.0-2.5, 2.5-2.5 all have their usage.

Example:

Speed: 1.0-2.5

Use of this header changes the bandwidth used for data delivery. It is meant for use in specific circumstances where delivery of the presentation at a higher or lower rate is desired. The main use cases are buffer operations or local scale operations. Implementors should keep in mind that bandwidth for the session may be negotiated

beforehand (by means other than RTSP), and therefore re-negotiation may be necessary. To perform Speed operations the server needs to ensure that the network path can support the resulting bit-rate. Thus the media transport needs to support feedback so that the server can react and adapt to the available bitrate.

16.49. Supported

The Supported header enumerates all the extensions supported by the client or server using feature tags. The header carries the extensions supported by the message sending client or server. The Supported header MAY be included in any request. When present in a request, the receiver MUST respond with its corresponding Supported header. Note that the Supported header is also included in 4xx and 5xx responses.

The Supported header contains a list of feature-tags, described in Section 4.7, that are understood by the client or server.

Example:

```
C->S: OPTIONS rtsp://example.com/ RTSP/2.0
      Supported: foo, bar, blech
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      Supported: bar, blech, baz
```

16.50. Terminate-Reason

The Terminate-Reason request header allows the server when sending a REDIRECT or TEARDOWN request to provide a reason for the session termination and any additional information. This specification identifies three reasons for Redirections and may be extended in the future:

Server-Admin: The server needs to be shutdown for some administrative reason.

Session-Timeout: A client's session is kept alive for extended periods of time and the server has determined that it needs to reclaim the resources associated with this session.

Internal-Error An internal error that is impossible to recover from has occurred forcing the server to terminate the session.

The Server may provide additional parameters containing information around the redirect. This specification defines the following ones.

time: Provides a wallclock time when the server will stop provide any service.

user-msg: An UTF-8 text string with a message from the server to the user. This message SHOULD be displayed to the user.

16.51. Timestamp

The Timestamp general-header describes when the agent sent the request. The value of the timestamp is of significance only to the agent and may use any timescale. The responding agent MUST echo the exact same value and MAY, if it has accurate information about this, add a floating point number indicating the number of seconds that has elapsed since it has received the request. The timestamp can be used by the agent to compute the round-trip time to the responding agent so that it can adjust the timeout value for retransmissions when running over an unreliable protocol. It also resolves retransmission ambiguities for unreliable transport of RTSP.

Note that the present specification provides only for reliable transport of RTSP messages. The Timestamp general-header is specified in case the protocol is extended in the future to use unreliable transport.

16.52. Transport

The Transport request and response header indicates which transport protocol is to be used and configures its parameters such as destination address, compression, multicast time-to-live and destination port for a single stream. It sets those values not already determined by a presentation description.

A Transport request header MAY contain a list of transport options acceptable to the client, in the form of multiple transport specification entries. Transport specifications are comma separated, listed in decreasing order of preference. Parameters may be added to each transport specification, separated by a semicolon. The server MUST return a Transport response-header in the response to indicate the values actually chosen if any. If the transport specification is not supported, no transport header is returned and the request MUST be responded using the status code 461 (Unsupported Transport) (Section 15.4.26). In case more than one transport specification was present in the request, the server MUST return the single (transport-spec) which was actually chosen, if any. The number of transport-spec entries is expected to be limited as the client will get guidance on what configurations that are possible from the presentation description.

The Transport header MAY also be used in subsequent SETUP requests to change transport parameters. A server MAY refuse to change parameters of an existing stream.

A transport specification may only contain one of any given parameter within it. Parameters MAY be given in any order. Additionally, it may only contain either of the unicast or the multicast transport type parameter. All parameters need to be understood in a transport specification, if not, the transport specification MUST be ignored. RTSP proxies of any type that uses or modifies the transport specification, e.g. access proxy or security proxy, MUST remove specifications with unknown parameters before forwarding the RTSP message. If that result in no remaining transport specification the proxy SHALL send a 461 (Unsupported Transport) (Section 15.4.26) response without any Transport header.

The Transport header is restricted to describing a single media stream. (RTSP can also control multiple streams as a single entity.) Making it part of RTSP rather than relying on a multitude of session description formats greatly simplifies designs of firewalls.

The general syntax for the transport specifier is a list of slash separated tokens:

Value1/Value2/Value3...

Which for RTP transports take the form:

RTP/profile/lower-transport.

The default value for the "lower-transport" parameters is specific to the profile. For RTP/AVP, the default is UDP.

There are two different methods for how to specify where the media should be delivered for unicast transport:

dest_addr: The presence of this parameter and its values indicates the destination address or addresses (host address and port pairs for IP flows) necessary for the media transport.

No dest_addr: The lack of the dest_addr parameter indicates that the server MUST send media to same address for which the RTSP messages originates.

The choice of method for indicating where the media is to be delivered depends on the use case. In some cases the only allowed method will be to use no explicit address indication and have the server deliver media to the source of the RTSP messages.

For Multicast there is several methods for specifying addresses but

they are different in how they work compared with unicast:

dest_addr with client picked address: The address and relevant parameters like TTL (scope) for the actual multicast group to deliver the media to. There are security implications (Section 21) with this method that needs to be addressed if using this method because a RTSP server can be used as a DoS attacker on an existing multicast group.

dest_addr using Session Description Information: The information included in the transport header can all be coming from the session description, e.g. the SDP c= and m= line. This mitigates some of the security issues of the previous methods as it is the session provider that picks the multicast group and scope. The client **MUST** include the information if it is available in the session description.

No dest_addr: The behavior when no explicit multicast group is present in a request is not defined.

An RTSP proxy will need to take care. If the media is not desired to be routed through the proxy, the proxy will need to introduce the destination indication.

Below are the configuration parameters associated with transport:

General parameters:

unicast / multicast: This parameter is a mutually exclusive indication of whether unicast or multicast delivery will be attempted. One of the two values **MUST** be specified. Clients that are capable of handling both unicast and multicast transmission needs to indicate such capability by including two full transport-specs with separate parameters for each.

layers: The number of multicast layers to be used for this media stream. The layers are sent to consecutive addresses starting at the dest_addr address. If the parameter is not included, it defaults to a single layer.

dest_addr: A general destination address parameter that can contain one or more address specifications. Each combination of protocol/profile/lower transport needs to have the format and interpretation of its address specification defined. For RTP/AVP/UDP and RTP/AVP/TCP, the address specification is a tuple containing a host address and port. Note, only a single destination parameter per transport spec is intended. The usage of multiple destinations to distribute a single media to

multiple entities is unspecified.

The client originating the RTSP request MAY specify the destination address of the stream recipient with the host address part of the tuple. When the destination address is specified, the recipient may be a different party than the originator of the request. To avoid becoming the unwitting perpetrator of a remote-controlled denial-of-service attack, a server MUST perform security checks (see Section 21.1) and SHOULD log such attempts before allowing the client to direct a media stream to a recipient address not chosen by the server. Implementations cannot rely on TCP as reliable means of client identification. If the server does not allow the host address part of the tuple to be set, it MUST return 463 (Destination Prohibited).

The host address part of the tuple MAY be empty, for example ":58044", in cases when only destination port is desired to be specified. Responses to requests including the Transport header with a dest_addr parameter SHOULD include the full destination address that is actually used by the server. The server MUST NOT remove address information present already in the request when responding unless the protocol requires it.

src_addr: A general source address parameter that can contain one or more address specifications. Each combination of protocol/profile/lower transport needs to have the format and interpretation of its address specification defined. For RTP/AVP/UDP and RTP/AVP/TCP, the address specification is a tuple containing a host address and port.

This parameter MUST be specified by the server if it transmits media packets from another address than the one RTSP messages are sent to. This will allow the client to verify source address and give it a destination address for its RTCP feedback packets, if RTP is used. The address or addresses indicated in the src_addr parameter SHOULD be used both for sending and receiving of the media streams data packets. The main reasons are threefold: First, indicating the port and source address(s) lets the receiver know where from the packets is expected to originate. Secondly, traversal of NATs is greatly simplified when traffic is flowing symmetrically over an NAT binding. Thirdly, certain NAT traversal mechanisms, needs to know to which address and port to send so called "binding packets" from the receiver to the sender, thus creating an address binding in the NAT that the sender to receiver packet flow can use.

This information may also be available through SDP. However, since this is more a feature of transport than media initialization, the authoritative source for this information should be in the SETUP response.

mode: The mode parameter indicates the methods to be supported for this session. Currently defined valid values are "PLAY". If not provided, the default is "PLAY". The "RECORD" value was defined in RFC 2326 and is in this specification unspecified but reserved. RECORD and other values may be specified in the future.

interleaved: The interleaved parameter implies mixing the media stream with the control stream in whatever protocol is being used by the control stream, using the mechanism defined in Section 14. The argument provides the channel number to be used in the \$ block Section 14 and MUST be present. This parameter MAY be specified as a interval, e.g., interleaved=4-5 in cases where the transport choice for the media stream requires it, e.g., for RTP with RTCP. The channel number given in the request is only a guidance from the client to the server on what channel number(s) to use. The server MAY set any valid channel number in the response. The declared channel(s) are bi-directional, so both end-parties MAY send data on the given channel. One example of such usage is the second channel used for RTCP, where both server and client send RTCP packets on the same channel.

This allows RTP/RTCP to be handled similarly to the way that it is done with UDP, i.e., one channel for RTP and the other for RTCP.

MIKEY: This parameter is used in conjunction with transport specifications that can utilize MIKEY for security context establishment. So far only the SRTP based RTP profiles SAVP and SAVPF can utilize MIKEY and this is defined in Appendix C.1.4.1. This parameter can be included both in request and response messages. The binary MIKEY message SHALL be BASE64 [RFC4648] encoded before being included in the value part of the parameter.

Multicast-specific:

ttl: multicast time-to-live for IPv4. When included in requests the value indicate the TTL value that the client request the server to use. In a response, the value actually being used by the server is returned. A server will need to consider what values

that are reasonable and also the authority of the user to set this value. Corresponding functions are not needed for IPv6 as the scoping is part of the address.

RTP-specific:

These parameters MAY only be used if the media transport protocol is RTP.

ssrc: The ssrc parameter, if included in a SETUP response, indicates the RTP SSRC [RFC3550] value(s) that will be used by the media server for RTP packets within the stream. It is expressed as an eight digit hexadecimal value.

The ssrc parameter MUST NOT be specified in requests. The functionality of specifying the ssrc parameter in a SETUP request is deprecated as it is incompatible with the specification of RTP in RFC 3550[RFC3550]. If the parameter is included in the Transport header of a SETUP request, the server SHOULD ignore it, and choose appropriate SSRCs for the stream. The server SHOULD set the ssrc parameter in the Transport header of the response.

RTCP-mux: Use to negotiate the usage of RTP and RTCP multiplexing [RFC5761] on a single underlying transport stream / flow. The presence of this parameter in a SETUP request indicates the clients support and requires the server to use RTP and RTCP multiplexing. The client SHALL only include one transport stream in the Transport header specification. To provide the server with a choice between using RTP/RTCP multiplexing or not, two different transport header specifications must be included.

The parameters setup and connection defined below MAY only be used if the media transport protocol of the lower-level transport is connection-oriented (such as TCP). However, these parameters MUST NOT be used when interleaving data over the RTSP control connection.

setup: Clients use the setup parameter on the Transport line in a SETUP request, to indicate the roles it wishes to play in a TCP connection. This parameter is adapted from [RFC4145]. We discuss the use of this parameter in RTP/AVP/TCP non-interleaved transport in Appendix C.2.2; the discussion below is limited to syntactic issues. Clients may specify the following values for the setup parameter: ["active":] The client will initiate an outgoing connection. ["passive":] The client will accept an incoming connection. ["actpass":] The client is willing to accept an incoming connection or to

initiate an outgoing connection.

If a client does not specify a setup value, the "active" value is assumed.

In response to a client SETUP request where the setup parameter is set to "active", a server's 2xx reply MUST assign the setup parameter to "passive" on the Transport header line.

In response to a client SETUP request where the setup parameter is set to "passive", a server's 2xx reply MUST assign the setup parameter to "active" on the Transport header line.

In response to a client SETUP request where the setup parameter is set to "actpass", a server's 2xx reply MUST assign the setup parameter to "active" or "passive" on the Transport header line.

Note that the "holdconn" value for setup is not defined for RTSP use, and MUST NOT appear on a Transport line.

connection: Clients use the setup parameter on the Transport line in a SETUP request, to indicate the SETUP request prefers the reuse of an existing connection between client and server (in which case the client sets the "connection" parameter to "existing"), or that the client requires the creation of a new connection between client and server (in which case the client sets the "connection" parameter to "new"). Typically, clients use the "new" value for the first SETUP request for a URL, and "existing" for subsequent SETUP requests for a URL.

If a client SETUP request assigns the "new" value to "connection", the server response MUST also assign the "new" value to "connection" on the Transport line.

If a client SETUP request assigns the "existing" value to "connection", the server response MUST assign a value of "existing" or "new" to "connection" on the Transport line, at its discretion.

The default value of "connection" is "existing", for all SETUP requests (initial and subsequent).

The combination of transport protocol, profile and lower transport needs to be defined. A number of combinations are defined in the Appendix C.

Below is a usage example, showing a client advertising the capability

to handle multicast or unicast, preferring multicast. Since this is a unicast-only stream, the server responds with the proper transport parameters for unicast.

```
C->S: SETUP rtsp://example.com/foo/bar/baz.rm RTSP/2.0
      CSeq: 302
      Transport: RTP/AVP;multicast;mode="PLAY",
                RTP/AVP;unicast;dest_addr="192.0.2.5:3456"/
                "192.0.2.5:3457";mode="PLAY"
      Accept-Ranges: NPT, SMPTE, UTC
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 302
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Session: 47112344
      Transport: RTP/AVP;unicast;dest_addr="192.0.2.5:3456"/
                "192.0.2.5:3457";src_addr="192.0.2.224:6256"/
                "192.0.2.224:6257";mode="PLAY"
      Accept-Ranges: NPT
      Media-Properties: Random-Access=0.6, Dynamic,
                      Time-Limited=20081128T165900
```

16.53. Unsupported

The Unsupported response-header lists the features not supported by the responding RTSP agent. In the case where the feature was specified via the Proxy-Require field (Section 16.35), if there is a proxy on the path between the client and the server, the proxy **MUST** send a response message with a status code of 551 (Option Not Supported). The request **MUST NOT** be forwarded.

See Section 16.41 for a usage example.

16.54. User-Agent

The User-Agent general-header field contains information about the user agent originating the request. This is for statistical purposes, the tracing of protocol violations, and automated recognition of user agents for the sake of tailoring responses to avoid particular user agent limitations. User agents **SHOULD** include this field with requests. The field can contain multiple product tokens and comments identifying the agent and any subproducts which form a significant part of the user agent. By convention, the product tokens are listed in order of their significance for identifying the application.

Example:

User-Agent: PhonyClient/1.2

16.55. Vary

The Vary field value indicates the set of request-header fields that fully determines, while the response is fresh, whether a cache is permitted to use the response to reply to a subsequent request without revalidation. For uncacheable or stale responses, the Vary field value advises the user agent about the criteria that were used to select the representation. A Vary field value of "*" implies that a cache cannot determine from the request headers of a subsequent request whether this response is the appropriate representation.

An RTSP server SHOULD include a Vary header field with any cacheable response that is subject to server-driven negotiation. Doing so allows a cache to properly interpret future requests on that resource and informs the user agent about the presence of negotiation on that resource. A server MAY include a Vary header field with a non-cacheable response that is subject to server-driven negotiation, since this might provide the user agent with useful information about the dimensions over which the response varies at the time of the response.

A Vary field value consisting of a list of field-names signals that the representation selected for the response is based on a selection algorithm which considers ONLY the listed request-header field values in selecting the most appropriate representation. A cache MAY assume that the same selection will be made for future requests with the same values for the listed field names, for the duration of time for which the response is fresh.

The field-names given are not limited to the set of standard request-header fields defined by this specification. Field names are case-insensitive.

A Vary field value of "*" signals that unspecified parameters not limited to the request-headers (e.g., the network address of the client), play a role in the selection of the response representation. The "*" value MUST NOT be generated by a proxy server; it may only be generated by an origin server.

16.56. Via

The Via general-header field MUST be used by proxies to indicate the intermediate protocols and recipients between the user agent and the server on requests, and between the origin server and the client on responses. The field is intended to be used for tracking message forwards, avoiding request loops, and identifying the protocol

capabilities of all senders along the request/response chain.

Multiple Via field values represents each proxy that has forwarded the message. Each recipient MUST append its information such that the end result is ordered according to the sequence of forwarding applications.

Proxies (e.g., Access Proxy or Translator Proxy) SHOULD NOT, by default, forward the names and ports of hosts within the private/protected region. This information SHOULD only be propagated if explicitly enabled. If not enabled, the via-received of any host behind the firewall/NAT SHOULD be replaced by an appropriate pseudonym for that host.

For organizations that have strong privacy requirements for hiding internal structures, a proxy MAY combine an ordered subsequence of Via header field entries with identical sent-protocol values into a single such entry. Applications MUST NOT combine entries which have different received-protocol values.

16.57. WWW-Authenticate

The WWW-Authenticate response-header field MUST be included in 401 (Unauthorized) response messages. The field value consists of at least one challenge that indicates the authentication scheme(s) and parameters applicable to the Request-URI.

The HTTP access authentication process is described in [RFC2617]. User agents are advised to take special care in parsing the WWW-Authenticate field value as it might contain more than one challenge, or if more than one WWW-Authenticate header field is provided, the contents of a challenge itself can contain a comma-separated list of authentication parameters.

17. Proxies

RTSP Proxies are RTSP agents that are located in between a client and a server. A proxy can take on both the role as a client and as server depending on what it tries to accomplish. Proxies are also introduced for several different reasons and the below listed are often combined.

In general there are two categories of RTSP proxies, transparent (of which the client is not aware) and the non-transparent proxies (of which the client is aware). Transparent proxies are not visible to the client in terms of that the transport layer connection, e.g., TCP for RTSP, as there is only a single transport connection which is terminated at the RTSP client and the RTSP server. In the case of non-transparent proxies, there are two transport layer connections, one from the RTSP client to the RTSP proxy and a second from the RTSP proxy to the RTSP server.

There are these types of RTSP proxies:

Caching Proxy: This type of proxy is used to reduce the workload on servers and connections. By caching the description and media streams, i.e., the presentation, the proxy can serve a client with content, but without requesting it from the server once it has been cached and has not become stale. See the caching Section 18. This type of proxy is also expected to understand RTSP end-point functionality, i.e., functionality identified in the Require header in addition to what Proxy-Require demands.

Translator Proxy: This type of proxy is used to ensure that an RTSP client gets access to servers and content on an external network or using content encodings not supported by the client. The proxy performs the necessary translation of addresses, protocols or encodings. This type of proxy is expected to also understand RTSP end-point functionality, i.e. functionality identified in the Require header in addition to what Proxy-Require demands.

Access Proxy: This type of proxy is used to ensure that an RTSP clients get access to servers on an external network. Thus this proxy is placed on the border between two domains, e.g. a private address space and the public Internet. The proxy performs the necessary translation, usually addresses. This type of proxy is required to redirect the media to itself or a controlled gateway that performs the translation before the media can reach the client.

Security Proxy: This type of proxy is used to help facilitate security functions around RTSP. For example when having a firewalled network, the security proxy request that the necessary pinholes in the firewall are opened when a client in the protected network wants to access media streams on the external side. This proxy can also limit the clients access to certain types of content. This proxy can perform its function without redirecting the media between the server and client. However, in deployments with private address spaces this proxy is likely to be combined with the access proxy. Anyway, the functionality of this proxy is usually closely tied into understanding all aspects of the media transport.

Auditing Proxy: RTSP proxies can also provide network owners with a logging and audit point for RTSP sessions, e.g. for corporations that track their employees usage of the network. This type of proxy can perform its function without inserting itself or any other node in the media transport. This proxy type can also accept unknown methods as it doesn't interfere with the clients' requests.

All types of proxies can be used also when using secured communication with TLS as RTSP 2.0 allows the client to approve certificate chains used for connection establishment from a proxy, see Section 19.3.2. However, that trust model may not be suitable for all types of deployment. In those cases, the secured sessions do by-pass of the proxies.

Access proxies SHOULD NOT be used in equipment like NATs and firewalls that aren't expected to be regularly maintained, like home or small office equipment. In these cases it is better to use the NAT traversal procedures defined for RTSP 2.0 [I-D.ietf-mmusic-rtsp-nat]. The reason for these recommendations is that any extensions of RTSP resulting in new media transport protocols or profiles, new parameters, etc. may fail in a proxy that isn't maintained. This would impede RTSP's future development and usage.

17.1. Proxies and Protocol Extensions

The existence of proxies must always be considered when developing new RTSP extensions. Most types of proxies will need to implement any new method to operate correctly in the presence of that extension. New headers can be introduced and will not be blocked by older proxies. However, it is important to consider if this header and its function is required to be understood by the proxy or can be forwarded. If the header needs to be understood, a feature-tag representing the functionality MUST be included in the Proxy-Require

header. Below are guidelines for analysis if the header needs to be understood. The transport header and its parameters also shows that headers that are extensible and require correct interpretation in the proxy also require handling rules.

Whether a proxy needs to understand a header is not easy to determine, as they serve a broad variety of functions. When evaluating if a header needs to be understood, one can divide the functionality into three main categories:

Media modifying: The caching and translator proxies are modifying the actual media and therefore needs to understand also request directed to the server that affects how the media is rendered. Thus, this type of proxy needs to also understand the server side functionality.

Transport modifying: The access and the security proxy both need to understand how the transport is performed, either for opening pinholes or to translate the outer headers, e.g. IP and UDP.

Non-modifying: The audit proxy is special in that it does not modify the messages in other ways than to insert the Via header. That makes it possible for this type to forward RTSP messages that contain different types of unknown methods, headers or header parameters.

Based on the above classification, one should evaluate if the new functionality requires the Transport modifying type of proxies to understand it or not.

17.2. Multiplexing and Demultiplexing of Messages

RTSP proxies may have to multiplex multiple RTSP sessions from their clients towards RTSP servers. This requires that RTSP requests from multiple clients are multiplexed onto a common connection for requests outgoing to an RTSP server and on the way back the responses are demultiplexed from the server to per client responses. On the protocol level this requires that request and response messages are handled in both ways, requiring that there is a mechanism to correlate what request/response pair exchanged between proxy and server is mapped to what client (or client request).

This multiplexing of requests and demultiplexing of responses is done by using the CSeq header field (see Section 16.19). The proxy has to rewrite the CSeq in requests to the server and responses from the server and remember what CSeq is mapped to what client.

18. Caching

In HTTP, request-response pairs are cached. RTSP differs significantly in that respect. Responses are not cacheable, with the exception of the presentation description returned by DESCRIBE. (Since the responses for anything but DESCRIBE and GET_PARAMETER do not return any data, caching is not really an issue for these requests.) However, it is desirable for the continuous media data, typically delivered out-of-band with respect to RTSP, to be cached, as well as the session description.

On receiving a SETUP or PLAY request, a proxy ascertains whether it has an up-to-date copy of the continuous media content and its description. It can determine whether the copy is up-to-date by issuing a SETUP or DESCRIBE request, respectively, and comparing the Last-Modified header with that of the cached copy. If the copy is not up-to-date, it modifies the SETUP transport parameters as appropriate and forwards the request to the origin server. Subsequent control commands such as PLAY or PAUSE then pass the proxy unmodified. The proxy delivers the continuous media data to the client, while possibly making a local copy for later reuse. The exact allowed behavior of the cache is given by the cache-response directives described in Section 16.10. A cache MUST answer any DESCRIBE requests if it is currently serving the stream to the requester, as it is possible that low-level details of the stream description may have changed on the origin-server.

Note that an RTSP cache, is of the "cut-through" variety. Rather than retrieving the whole resource from the origin server, the cache simply copies the streaming data as it passes by on its way to the client. Thus, it does not introduce additional latency.

To the client, an RTSP proxy cache appears like a regular media server, to the media origin server like a client. Just as an HTTP cache has to store the content type, content language, and so on for the objects it caches, a media cache has to store the presentation description. Typically, a cache eliminates all transport-references (e.g., multicast information) from the presentation description, since these are independent of the data delivery from the cache to the client. Information on the encodings remains the same. If the cache is able to translate the cached media data, it would create a new presentation description with all the encoding possibilities it can offer.

18.1. Validation Model

When a cache has a stale entry that it would like to use as a response to a client's request, it first has to check with the origin

server (or possibly an intermediate cache with a fresh response) to see if its cached entry is still usable. We call this "validating" the cache entry. Since we do not want to have to pay the overhead of retransmitting the full response if the cached entry is good, and we do not want to pay the overhead of an extra round trip if the cached entry is invalid, the RTSP protocol supports the use of conditional methods.

The key protocol features for supporting conditional methods are those concerned with "cache validators." When an origin server generates a full response, it attaches some sort of validator to it, which is kept with the cache entry. When a client (user agent or proxy cache) makes a conditional request for a resource for which it has a cache entry, it includes the associated validator in the request.

The server then checks that validator against the current validator for the requested resource, and, if they match (see Section 18.1.3), it responds with a special status code (usually, 304 (Not Modified)) and no message body. Otherwise, it returns a full response (including message body). Thus, we avoid transmitting the full response if the validator matches, and we avoid an extra round trip if it does not match.

In RTSP, a conditional request looks exactly the same as a normal request for the same resource, except that it carries a special header (which includes the validator) that implicitly turns the method (usually DESCRIBE or SETUP) into a conditional.

The protocol includes both positive and negative senses of cache-validating conditions. That is, it is possible to request either that a method be performed if and only if a validator matches or if and only if no validators match.

Note: a response that lacks a validator may still be cached, and served from cache until it expires, unless this is explicitly prohibited by a cache-control directive (see Section 16.10). However, a cache cannot do a conditional retrieval if it does not have a validator for the resource, which means it will not be refreshable after it expires.

Media streams that are being adapted based on the transport capacity between the server and the cache makes caching more difficult. A server needs to consider how it views caching of media streams that it adapts and potentially instruct any caches to not cache such streams.

18.1.1.1. Last-Modified Dates

The Last-Modified header (Section 16.26) value is often used as a cache validator. In simple terms, a cache entry is considered to be valid if the content has not been modified since the Last-Modified value.

18.1.1.2. Message Body Tag Cache Validators

The MTag response-header field value, a message body tag, provides for an "opaque" cache validator. This might allow more reliable validation in situations where it is inconvenient to store modification dates, where the one-second resolution of RTSP-date values is not sufficient, or where the origin server wishes to avoid certain paradoxes that might arise from the use of modification dates.

Message body tags are described in Section 4.8

18.1.1.3. Weak and Strong Validators

Since both origin servers and caches will compare two validators to decide if they represent the same or different entities, one normally would expect that if the message body (i.e., the presentation description) or any associated message body headers changes in any way, then the associated validator would change as well. If this is true, then we call this validator a "strong validator." We call message body (i.e., the presentation description) or any associated message body headers an entity for a better understanding.

However, there might be cases when a server prefers to change the validator only on semantically significant changes, and not when insignificant aspects of the entity change. A validator that does not always change when the resource changes is a "weak validator."

Message body tags are normally "strong validators," but the protocol provides a mechanism to tag a message body tag as "weak." One can think of a strong validator as one that changes whenever the bits of an entity changes, while a weak value changes whenever the meaning of an entity changes. Alternatively, one can think of a strong validator as part of an identifier for a specific entity, while a weak validator is part of an identifier for a set of semantically equivalent entities.

Note: One example of a strong validator is an integer that is incremented in stable storage every time an entity is changed.

An entity's modification time, if represented with one-second resolution, could be a weak validator, since it is possible that the resource might be modified twice during a single second.

Support for weak validators is optional. However, weak validators allow for more efficient caching of equivalent objects.

A "use" of a validator is either when a client generates a request and includes the validator in a validating header field, or when a server compares two validators.

Strong validators are usable in any context. Weak validators are only usable in contexts that do not depend on exact equality of an entity. For example, either kind is usable for a conditional DESCRIBE of a full entity. However, only a strong validator is usable for a sub-range retrieval, since otherwise the client might end up with an internally inconsistent entity.

Clients MAY issue DESCRIBE requests with either weak validators or strong validators. Clients MUST NOT use weak validators in other forms of requests.

The only function that the RTSP protocol defines on validators is comparison. There are two validator comparison functions, depending on whether the comparison context allows the use of weak validators or not:

- o The strong comparison function: in order to be considered equal, both validators MUST be identical in every way, and both MUST NOT be weak.
- o The weak comparison function: in order to be considered equal, both validators MUST be identical in every way, but either or both of them MAY be tagged as "weak" without affecting the result.

A message body tag is strong unless it is explicitly tagged as weak.

A Last-Modified time, when used as a validator in a request, is implicitly weak unless it is possible to deduce that it is strong, using the following rules:

- o The validator is being compared by an origin server to the actual current validator for the entity and,

- o That origin server reliably knows that the associated entity did not change more than once during the second covered by the presented validator.

OR

- o The validator is about to be used by a client in an If-Modified-Since, because the client has a cache entry for the associated entity, and
- o That cache entry includes a Date value, which gives the time when the origin server sent the original response, and
- o The presented Last-Modified time is at least 60 seconds before the Date value.

OR

- o The validator is being compared by an intermediate cache to the validator stored in its cache entry for the entity, and
- o That cache entry includes a Date value, which gives the time when the origin server sent the original response, and
- o The presented Last-Modified time is at least 60 seconds before the Date value.

This method relies on the fact that if two different responses were sent by the origin server during the same second, but both had the same Last-Modified time, then at least one of those responses would have a Date value equal to its Last-Modified time. The arbitrary 60-second limit guards against the possibility that the Date and Last-Modified values are generated from different clocks, or at somewhat different times during the preparation of the response. An implementation MAY use a value larger than 60 seconds, if it is believed that 60 seconds is too short.

If a client wishes to perform a sub-range retrieval on a value for which it has only a Last-Modified time and no opaque validator, it MAY do this only if the Last-Modified time is strong in the sense described here.

18.1.1.4. Rules for When to Use Message Body Tags and Last-Modified Dates

We adopt a set of rules and recommendations for origin servers, clients, and caches regarding when various validator types ought to be used, and for what purposes.

RTSP origin servers:

- o SHOULD send a message body tag validator unless it is not feasible to generate one.
- o MAY send a weak message body tag instead of a strong message body tag, if performance considerations support the use of weak message body tags, or if it is unfeasible to send a strong message body tag.
- o SHOULD send a Last-Modified value if it is feasible to send one, unless the risk of a breakdown in semantic transparency that could result from using this date in an If-Modified-Since header would lead to serious problems.

In other words, the preferred behavior for an RTSP origin server is to send both a strong message body tag and a Last-Modified value.

In order to be legal, a strong message body tag MUST change whenever the associated entity value changes in any way. A weak message body tag SHOULD change whenever the associated entity changes in a semantically significant way.

Note: in order to provide semantically transparent caching, an origin server MUST avoid reusing a specific strong message body tag value for two different entities, or reusing a specific weak message body tag value for two semantically different entities. Cache entries might persist for arbitrarily long periods, regardless of expiration times, so it might be inappropriate to expect that a cache will never again attempt to validate an entry using a validator that it obtained at some point in the past.

RTSP clients:

- o If a message body tag has been provided by the origin server, MUST use that message body tag in any cache-conditional request (using If-Match or If-None-Match).
- o If only a Last-Modified value has been provided by the origin server, SHOULD use that value in non-subrange cache-conditional requests (using If-Modified-Since).
- o If both a message body tag and a Last-Modified value have been provided by the origin server, SHOULD use both validators in cache-conditional requests.

An RTSP origin server, upon receiving a conditional request that includes both a Last-Modified date (e.g., in an If-Modified-Since

header) and one or more message body tags (e.g., in an If-Match, If-None-Match, or If-Range header field) as cache validators, MUST NOT return a response status of 304 (Not Modified) unless doing so is consistent with all of the conditional header fields in the request.

Note: The general principle behind these rules is that RTSP servers and clients should transmit as much non-redundant information as is available in their responses and requests. RTSP systems receiving this information will make the most conservative assumptions about the validators they receive.

18.1.5. Non-validating Conditionals

The principle behind message body tags is that only the service author knows the semantics of a resource well enough to select an appropriate cache validation mechanism, and the specification of any validator comparison function more complex than byte-equality would open up a can of worms. Thus, comparisons of any other headers are never used for purposes of validating a cache entry.

18.2. Invalidation After Updates or Deletions

The effect of certain methods performed on a resource at the origin server might cause one or more existing cache entries to become non-transparently invalid. That is, although they might continue to be "fresh," they do not accurately reflect what the origin server would return for a new request on that resource.

There is no way for the RTSP protocol to guarantee that all such cache entries are marked invalid. For example, the request that caused the change at the origin server might not have gone through the proxy where a cache entry is stored. However, several rules help reduce the likelihood of erroneous behavior.

In this section, the phrase "invalidate an entity" means that the cache will either remove all instances of that entity from its storage, or will mark these as "invalid" and in need of a mandatory revalidation before they can be returned in response to a subsequent request.

Some RTSP methods MUST cause a cache to invalidate an entity. This is either the entity referred to by the Request-URI, or by the Location or Content-Location headers (if present). These methods are:

- o DESCRIBE

o SETUP

In order to prevent denial of service attacks, an invalidation based on the URI in a Location or Content-Location header **MUST** only be performed if the host part is the same as in the Request-URI.

A cache that passes through requests for methods it does not understand **SHOULD** invalidate any entities referred to by the Request-URI.

19. Security Framework

The RTSP security framework consists of two high level components: the pure authentication mechanisms based on HTTP authentication, and the message transport protection based on TLS, which is independent of RTSP. Because of the similarity in syntax and usage between RTSP servers and HTTP servers, the security for HTTP is re-used to a large extent.

19.1. RTSP and HTTP Authentication

RTSP and HTTP share common authentication schemes, and thus follow the same usage guidelines as specified in [RFC2617] and also in [H15]. Servers SHOULD implement both basic and digest [RFC2617] authentication. Clients MUST implement both basic and digest authentication [RFC2617] so that a server that requires the client to authenticate can trust that the capability is present.

It should be stressed that using the HTTP authentication alone does not provide full control message security. Therefore, in environments requiring tighter security for the control messages, TLS SHOULD be used, see Section 19.2.

19.2. RTSP over TLS

RTSP MUST follow the same guidelines with regards to TLS [RFC5246] usage as specified for HTTP, see [RFC2818]. RTSP over TLS is separated from unsecured RTSP both on URI level and port level. Instead of using the "rtsp" scheme identifier in the URI, the "rtsp" scheme identifier MUST be used to signal RTSP over TLS. If no port is given in a URI with the "rtsp" scheme, port 322 MUST be used for TLS over TCP/IP.

When a client tries to setup an insecure channel to the server (using the "rtsp" URI), and the policy for the resource requires a secure channel, the server MUST redirect the client to the secure service by sending a 301 redirect response code together with the correct Location URI (using the "rtsp" scheme). A user or client MAY upgrade a non secured URI to a secured by changing the scheme from "rtsp" to "rtsp". A server implementing support for "rtsp" MUST allow this.

It should be noted that TLS allows for mutual authentication (when using both server and client certificates). Still, one of the more common ways TLS is used is to only provide server side authentication (often to avoid client certificates). TLS is then used in addition to HTTP authentication, providing transport security and server authentication, while HTTP Authentication is used to authenticate the

client.

RTSP includes the possibility to keep a TCP session up between the client and server, throughout the RTSP session lifetime. It may be convenient to keep the TCP session, not only to save the extra setup time for TCP, but also the extra setup time for TLS (even if TLS uses the resume function, there will be almost two extra round trips). Still, when TLS is used, such behavior introduces extra active state in the server, not only for TCP and RTSP, but also for TLS. This may increase the vulnerability to DoS attacks.

In addition to these recommendations, Section 19.3 gives further recommendations of TLS usage with proxies.

19.3. Security and Proxies

The nature of a proxy is often to act as a "man-in-the-middle", while security is often about preventing the existence of a "man-in-the-middle". This section provides clients with the possibility to use proxies even when applying secure transports (TLS) between the RTSP agents. The TLS proxy mechanism allows for server and proxy identification using certificates. However, the client can not be identified based on certificates. The client needs to select between using the procedure specified below or using a TLS connection directly (by-passing any proxies) to the server. The choice may be dependent on policies.

There are basically two categories of proxies, the transparent proxies (of which the client is not aware) and the non-transparent proxies (of which the client is aware), see Section Section 17 for an introduction to RTSP proxies. An infrastructure based on proxies requires that the trust model is such that both client and servers can trust the proxies to handle the RTSP messages correctly. To be able to trust a proxy, the client and server also needs to be aware of the proxy. Hence, transparent proxies cannot generally be seen as trusted and will not work well with security (unless they work only at transport layer). In the rest of this section any reference to proxy will be to a non-transparent proxy, which inspects or manipulate the RTSP messages.

HTTP Authentication is built on the assumption of proxies and can provide user-proxy authentication and proxy-proxy/server authentication in addition to the client-server authentication.

When TLS is applied and a proxy is used, the client will connect to the proxy's address when connecting to any RTSP server. This implies that for TLS, the client will authenticate the proxy server and not the end server. Note that when the client checks the server

certificate in TLS, it MUST check the proxy's identity (URI or possibly other known identity) against the proxy's identity as presented in the proxy's Certificate message.

The problem is that for a proxy accepted by the client, the proxy needs to be provided information on which grounds it should accept the next-hop certificate. Both the proxy and the user may have rules for this, and the user have the possibility to select the desired behavior. To handle this case, the Accept-Credentials header (See Section 16.2) is used, where the client can force the proxy/proxies to relay back the chain of certificates used to authenticate any intermediate proxies as well as the server. Given the assumption that the proxies are viewed as trusted, it gives the user a possibility to enforce policies to each trusted proxy of whether it should accept the next agent in the chain.

A proxy MUST use TLS for the next hop if the RTSP request includes a "rtsps" URI. TLS MAY be applied on intermediate links (e.g. between client and proxy, or between proxy and proxy), even if the resource and the end server are not required to use it. The proxy MUST, when initiating the next hop TLS connection, use the incoming TLS connections cipher suite list, only modified by removing any cipher suits that the proxy does not support. In case a proxy fails to establish a TLS connection due to cipher suite mismatch between proxy and next hop proxy or server, this is indicated using error code 472 (Failure to establish secure connection).

19.3.1. Accept-Credentials

The Accept-Credentials header can be used by the client to distribute simple authorization policies to intermediate proxies. The client includes the Accept-Credentials header to dictate how the proxy treats the server/next proxy certificate. There are currently three methods defined:

Any, which means that the proxy (or proxies) MUST accept whatever certificate presented. This is of course not a recommended option to use, but may be useful in certain circumstances (such as testing).

Proxy, which means that the proxy (or proxies) MUST use its own policies to validate the certificate and decide whether to accept it or not. This is convenient in cases where the user has a strong trust relation with the proxy. Reason why a strong trust relation may exist are; personal/company proxy, proxy has a out-of-band policy configuration mechanism.

User, which means that the proxy (or proxies) MUST send credential information about the next hop to the client for authorization. The client can then decide whether the proxy should accept the certificate or not. See Section 19.3.2 for further details.

If the Accept-Credentials header is not included in the RTSP request from the client, then the "Proxy" method MUST be used as default. If another method than the "Proxy" is to be used, then the Accept-Credentials header MUST be included in all of the RTSP requests from the client. This is because it cannot be assumed that the proxy always keeps the TLS state or the users previous preference between different RTSP messages (in particular if the time interval between the messages is long).

With the "Any" and "Proxy" methods the proxy will apply the policy as defined for each method. If the policy does not accept the credentials of the next hop, the proxy MUST respond with a message using status code 471 (Connection Credentials not accepted).

An RTSP request in the direction server to client MUST NOT include the Accept-Credentials header. As for the non-secured communication, the possibility for these requests depends on the presence of a client established connection. However, if the server to client request is in relation to a session established over a TLS secured channel, it MUST be sent in a TLS secured connection. That secured connection MUST also be the one used by the last client to server request. If no such transport connection exist at the time when the server desires to send the request, the server MUST discard the message.

Further policies MAY be defined and registered, but should be done so with caution.

19.3.2. User approved TLS procedure

For the "User" method, each proxy MUST perform the following procedure for each RTSP request:

- o Setup the TLS session to the next hop if not already present (i.e. run the TLS handshake, but do not send the RTSP request).
- o Extract the peer certificate chain for the TLS session.
- o Check if a matching identity and hash of the peer certificate is present in the Accept-Credentials header. If present, send the message to the next hop, and conclude these procedures. If not, go to the next step.

- o The proxy responds to the RTSP request with a 470 or 407 response code. The 407 response code MAY be used when the proxy requires both user and connection authorization from user or client. In this message the proxy MUST include a Connection-Credentials header, see Section 16.12 with the next hop's identity and certificate.

The client MUST upon receiving a 470 or 407 response with Connection-Credentials header take the decision on whether to accept the certificate or not (if it cannot do so, the user SHOULD be consulted). If the certificate is accepted, the client has to again send the RTSP request. In that request the client has to include the Accept-Credentials header including the hash over the DER encoded certificate for all trusted proxies in the chain.

Example:

```
C->P: SETUP rtsp://test.example.org/secret/audio RTSP/2.0
      CSeq: 2
      Transport: RTP/AVP;unicast;dest_addr="192.0.2.5:4588"/
                "192.0.2.5:4589"
      Accept-Ranges: NPT, SMPTE, UTC
      Accept-Credentials: User

P->C: RTSP/2.0 470 Connection Authorization Required
      CSeq: 2
      Connection-Credentials: "rtsp://test.example.org";
      MIIDNTCCAp...

C->P: SETUP rtsp://test.example.org/secret/audio RTSP/2.0
      CSeq: 3
      Transport: RTP/AVP;unicast;dest_addr="192.0.2.5:4588"/
                "192.0.2.5:4589"
      Accept-Credentials: User "rtsp://test.example.org";sha-256;
      dPYD7txpoGTbAqZZQJ+vaeOkYH4=
      Accept-Ranges: NPT, SMPTE, UTC

P->S: SETUP rtsp://test.example.org/secret/audio RTSP/2.0
      CSeq: 3
      Transport: RTP/AVP;unicast;dest_addr="192.0.2.5:4588"/
                "192.0.2.5:4589"
      Via: RTSP/2.0 proxy.example.org
      Accept-Credentials: User "rtsp://test.example.org";sha-256;
      dPYD7txpoGTbAqZZQJ+vaeOkYH4=
      Accept-Ranges: NPT, SMPTE, UTC
```

One implication of this process is that the connection for secured RTSP messages may take significantly more round-trip times for the

first message. A complete extra message exchange between the proxy connecting to the next hop and the client results because of the process for approval for each hop. However, if each message contains the chain of proxies that the requester accepts, the remaining message exchange should not be delayed. The procedure of including the credentials in each request rather than building state in each proxy, avoids the need for revocation procedures.

20. Syntax

The RTSP syntax is described in an Augmented Backus-Naur Form (ABNF) as defined in RFC 5234 [RFC5234]. It uses the basic definitions present in RFC 5234.

Please note that ABNF strings, e.g. "Accept", are case insensitive as specified in section 2.3 of RFC 5234.

20.1. Base Syntax

RTSP header values can be folded onto multiple lines if the continuation line begins with a space or horizontal tab. All linear white space, including folding, has the same semantics as SP. A recipient MAY replace any linear white space with a single SP before interpreting the field value or forwarding the message downstream. This is intended to behave exactly as HTTP/1.1 as described in RFC 2616 [RFC2616]. The SWS construct is used when linear white space is optional, generally between tokens and separators.

To separate the header name from the rest of value, a colon is used, which, by the above rule, allows whitespace before, but no line break, and whitespace after, including a line break. The HCOLON defines this construct.

OCTET	=	%x00-FF ; any 8-bit sequence of data
CHAR	=	%x01-7F ; any US-ASCII character (octets 1 - 127)
UPALPHA	=	%x41-5A ; any US-ASCII uppercase letter "A".. "Z"
LOALPHA	=	%x61-7A ; any US-ASCII lowercase letter "a".. "z"
ALPHA	=	UPALPHA / LOALPHA
DIGIT	=	%x30-39 ; any US-ASCII digit "0".. "9"
CTL	=	%x00-1F / %x7F ; any US-ASCII control character ; (octets 0 - 31) and DEL (127)
CR	=	%x0D ; US-ASCII CR, carriage return (13)
LF	=	%x0A ; US-ASCII LF, linefeed (10)
SP	=	%x20 ; US-ASCII SP, space (32)
HT	=	%x09 ; US-ASCII HT, horizontal-tab (9)
DQ	=	%x22 ; US-ASCII double-quote mark (34)
BACKSLASH	=	%x5C ; US-ASCII backslash (92)
CRLF	=	CR LF

```

LWS           = [CRLF] 1*( SP / HT ) ; Line-breaking White Space
SWS           = [LWS] ; Separating White Space
HCOLON        = *( SP / HT ) ":" SWS
TEXT          = %x20-7E / %x80-FF ; any OCTET except CTLs
tspecials     = "(" / ")" / "<" / ">" / "@"
               / "," / ";" / ":" / BACKSLASH / DQ
               / "/" / "[" / "]" / "?" / "="
               / "{" / "}" / SP / HT
token         = 1*(%x21 / %x23-27 / %x2A-2B / %x2D-2E / %x30-39
               / %x41-5A / %x5E-7A / %x7C / %x7E)
               ; 1*<any CHAR except CTLs or tspecials>
quoted-string = ( DQ *qdtex DQ )
qdtex         = %x20-21 / %x23-7E / %x80-FF / UTF8-NONASCII
               ; any UTF-8 TEXT except <">
quoted-pair   = BACKSLASH CHAR
ctex          = %x20-27 / %x2A-7E
               / %x80-FF ; any OCTET except CTLs, "(" and ")"
generic-param = token [ EQUAL gen-value ]
gen-value     = token / host / quoted-string

safe          = "$" / "-" / "_" / "." / "+"
extra         = "!" / "*" / "'" / "(" / ")" / ","
rtsp-extra    = "!" / "*" / "'" / "(" / ")"

HEX           = DIGIT / "A" / "B" / "C" / "D" / "E" / "F"
               / "a" / "b" / "c" / "d" / "e" / "f"
LHEX          = DIGIT / "a" / "b" / "c" / "d" / "e" / "f"
               ; lowercase "a-f" Hex
reserved      = ";" / "/" / "?" / ":" / "@" / "&" / "="

unreserved    = ALPHA / DIGIT / safe / extra
rtsp-unreserved = ALPHA / DIGIT / safe / rtsp-extra

base64        = *base64-unit [base64-pad]
base64-unit   = 4base64-char
base64-pad    = (2base64-char "==") / (3base64-char "=")
base64-char   = ALPHA / DIGIT / "+" / "/"

```

SLASH = SWS "/" SWS ; slash
EQUAL = SWS "=" SWS ; equal
LPAREN = SWS "(" SWS ; left parenthesis
RPAREN = SWS ")" SWS ; right parenthesis
COMMA = SWS "," SWS ; comma
SEMI = SWS ";" SWS ; semicolon
COLON = SWS ":" SWS ; colon
MINUS = SWS "-" SWS ; minus/dash
LDQUOT = SWS DQ ; open double quotation mark
RDQUOT = DQ SWS ; close double quotation mark
RAQUOT = ">" SWS ; right angle quote
LAQUOT = SWS "<" ; left angle quote

TEXT-UTF8char = %x21-7E / UTF8-NONASCII
UTF8-NONASCII = %xC0-DF 1UTF8-CONT
 / %xE0-EF 2UTF8-CONT
 / %xF0-F7 3UTF8-CONT
 / %xF8-FB 4UTF8-CONT
 / %xFC-FD 5UTF8-CONT
UTF8-CONT = %x80-BF

POS-FLOAT = 1*12DIGIT ["." 1*9DIGIT]
FLOAT = ["-"] POS-FLOAT

20.2. RTSP Protocol Definition

20.2.1. Generic Protocol elements

```

RTSP-IRI      = schemes ":" IRI-rest
IRI-rest      = ihier-part [ "?" iquery ] [ "#" ifragment ]
ihier-part    = "//" iauthority ipath-abempty
RTSP-IRI-ref  = RTSP-IRI / irelative-ref
irelative-ref = irelative-part [ "?" iquery ] [ "#" ifragment ]
irelative-part = "//" iauthority ipath-abempty
               / ipath-absolute
               / ipath-noscheme
               / ipath-empty

iauthority    = < As defined in RFC 3987>
ipath         = ipath-abempty ; begins with "/" or is empty
               / ipath-absolute ; begins with "/" but not "//"
               / ipath-noscheme ; begins with a non-colon segment
               / ipath-rootless ; begins with a segment
               / ipath-empty ; zero characters

ipath-abempty = *( "/" isegment )
ipath-absolute = "/" [ isegment-nz *( "/" isegment ) ]
ipath-noscheme = isegment-nz-nc *( "/" isegment )
ipath-rootless = isegment-nz *( "/" isegment )
ipath-empty    = 0<ipchar>

isegment      = *ipchar [ ";" *ipchar ]
isegment-nz   = 1*ipchar [ ";" *ipchar ]
               / ";" *ipchar
isegment-nz-nc = (1*ipchar-nc [ ";" *ipchar-nc ]
               / ";" *ipchar-nc
               ; non-zero-length segment without any colon ":"

ipchar        = iunreserved / pct-encoded / sub-delims / ":" / "@"
ipchar-nc     = iunreserved / pct-encoded / sub-delims / "@"

iquery        = < As defined in RFC 3987>
ifragment     = < As defined in RFC 3987>
iunreserved   = < As defined in RFC 3987>
pct-encoded   = < As defined in RFC 3987>

```

```

RTSP-URI      = schemes ":" URI-rest
RTSP-REQ-URI  = schemes ":" URI-req-rest
RTSP-URI-Ref  = RTSP-URI / RTSP-Relative
RTSP-REQ-Ref  = RTSP-REQ-URI / RTSP-REQ-Rel
schemes       = "rtsp" / "rtsps" / scheme
scheme        = < As defined in RFC 3986>
URI-rest      = hier-part [ "?" query ] [ "#" fragment ]
URI-req-rest   = hier-part [ "?" query ]
                ; Note fragment part not allowed in requests
hier-part     = "://" authority path-abempty

RTSP-Relative = relative-part [ "?" query ] [ "#" fragment ]
RTSP-REQ-Rel  = relative-part [ "?" query ]
relative-part = "://" authority path-abempty
                / path-absolute
                / path-noscheme
                / path-empty

authority      = < As defined in RFC 3986>
query         = < As defined in RFC 3986>
fragment      = < As defined in RFC 3986>

path          = path-abempty      ; begins with "/" or is empty
                / path-absolute   ; begins with "/" but not "/"
                / path-noscheme   ; begins with a non-colon segment
                / path-rootless   ; begins with a segment
                / path-empty      ; zero characters

path-abempty  = *( "/" segment )
path-absolute = "/" [ segment-nz *( "/" segment ) ]
path-noscheme = segment-nz-nc *( "/" segment )
path-rootless = segment-nz *( "/" segment )
path-empty    = 0<pchar>

segment       = *pchar [ ";" *pchar ]
segment-nz    = ( 1*pchar [ ";" *pchar ] ) / ( ";" *pchar )
segment-nz-nc = ( 1*pchar-nc [ ";" *pchar-nc ] ) / ( ";" *pchar-nc )
                ; non-zero-length segment without any colon ":"

pchar         = unreserved / pct-encoded / sub-delims / ":" / "@"
pchar-nc      = unreserved / pct-encoded / sub-delims / "@"

sub-delims    = "!" / "$" / "&" / "'" / "(" / ")"
                / "*" / "+" / "," / "="

```

```

smpte-range      = smpte-type ["=" smpte-range-spec]
                  ; See section 3.4
smpte-range-spec = ( smpte-time "-" [ smpte-time ] )
                  / ( "-" smpte-time )
smpte-type       = "smpte" / "smpte-30-drop"
                  / "smpte-25" / smpte-type-extension
                  ; other timecodes may be added
smpte-type-extension = "smpte" token
smpte-time       = 1*2DIGIT ":" 1*2DIGIT ":" 1*2DIGIT
                  [ ":" 1*2DIGIT [ "." 1*2DIGIT ] ]


npt-range      = "npt" ["=" npt-range-spec]
npt-range-spec = ( npt-time "-" [ npt-time ] ) / ( "-" npt-time )
npt-time       = "now" / npt-sec / npt-hhmmss
npt-sec        = 1*19DIGIT [ "." 1*9DIGIT ]
npt-hhmmss     = npt-hh ":" npt-mm ":" npt-ss [ "." 1*9DIGIT ]
npt-hh         = 1*19DIGIT ; any positive number
npt-mm         = 1*2DIGIT ; 0-59
npt-ss         = 1*2DIGIT ; 0-59


utc-range      = "clock" ["=" utc-range-spec]
utc-range-spec = ( utc-time "-" [ utc-time ] ) / ( "-" utc-time )
utc-time       = utc-date "T" utc-clock "Z"
utc-date       = 8DIGIT
utc-clock      = 6DIGIT [ "." 1*9DIGIT ]


feature-tag    = token

session-id     = 1*256( ALPHA / DIGIT / safe )

extension-header = header-name HCOLON header-value
header-name     = token
header-value    = *(TEXT-UTF8char / UTF8-CONT / LWS)

```

20.2.2. Message Syntax

```
RTSP-message = Request / Response ; RTSP/2.0 messages

Request      = Request-Line
               *((general-header
                  / request-header
                  / message-header) CRLF)
               CRLF
               [ message-body-data ]

Response     = Status-Line
               *((general-header
                  / response-header
                  / message-header) CRLF)
               CRLF
               [ message-body-data ]

Request-Line = Method SP Request-URI SP RTSP-Version CRLF

Status-Line  = RTSP-Version SP Status-Code SP Reason-Phrase CRLF
Method       = "DESCRIBE"
              / "GET_PARAMETER"
              / "OPTIONS"
              / "PAUSE"
              / "PLAY"
              / "PLAY_NOTIFY"
              / "REDIRECT"
              / "SETUP"
              / "SET_PARAMETER"
              / "TEARDOWN"
              / extension-method

extension-method = token

Request-URI    = "*" / RTSP-REQ-URI
RTSP-Version   = "RTSP/" 1*DIGIT "." 1*DIGIT

message-body-data = 1*OCTET

Status-Code    = "100" ; Continue
                / "200" ; OK
                / "301" ; Moved Permanently
                / "302" ; Found
                / "303" ; See Other
                / "304" ; Not Modified
                / "305" ; Use Proxy
                / "400" ; Bad Request
                / "401" ; Unauthorized
                / "402" ; Payment Required
```



```
/ "403" ; Forbidden
/ "404" ; Not Found
/ "405" ; Method Not Allowed
/ "406" ; Not Acceptable
/ "407" ; Proxy Authentication Required
/ "408" ; Request Time-out
/ "410" ; Gone
/ "411" ; Length Required
/ "412" ; Precondition Failed
/ "413" ; Request Message Body Too Large
/ "414" ; Request-URI Too Large
/ "415" ; Unsupported Media Type
/ "451" ; Parameter Not Understood
/ "452" ; reserved
/ "453" ; Not Enough Bandwidth
/ "454" ; Session Not Found
/ "455" ; Method Not Valid in This State
/ "456" ; Header Field Not Valid for Resource
/ "457" ; Invalid Range
/ "458" ; Parameter Is Read-Only
/ "459" ; Aggregate operation not allowed
/ "460" ; Only aggregate operation allowed
/ "461" ; Unsupported Transport
/ "462" ; Destination Unreachable
/ "463" ; Destination Prohibited
/ "464" ; Data Transport Not Ready Yet
/ "465" ; Notification Reason Unknown
/ "466" ; Key Management Error
/ "470" ; Connection Authorization Required
/ "471" ; Connection Credentials not accepted
/ "472" ; Failure to establish secure connection
/ "500" ; Internal Server Error
/ "501" ; Not Implemented
/ "502" ; Bad Gateway
/ "503" ; Service Unavailable
/ "504" ; Gateway Time-out
/ "505" ; RTSP Version not supported
/ "551" ; Option not supported
/ extension-code
```

extension-code = 3DIGIT

Reason-Phrase = 1*(TEXT-UTF8char / HT / SP)

```
general-header = Cache-Control
                / Connection
                / CSeq
                / Date
                / Media-Properties
                / Media-Range
                / Pipelined-Requests
                / Proxy-Supported
                / Seek-Style
                / Server
                / Supported
                / Timestamp
                / User-Agent
                / Via
                / extension-header

request-header  = Accept
                / Accept-Credentials
                / Accept-Encoding
                / Accept-Language
                / Authorization
                / Bandwidth
                / Blocksize
                / From
                / If-Match
                / If-Modified-Since
                / If-None-Match
                / Notify-Reason
                / Proxy-Require
                / Range
                / Referrer
                / Request-Status
                / Require
                / Scale
                / Session
                / Speed
                / Supported
                / Terminate-Reason
                / Transport
                / extension-header
```

```

response-header = Accept-Credentials
                  / Accept-Ranges
                  / Connection-Credentials
                  / MTag
                  / Location
                  / Proxy-Authenticate
                  / Public
                  / Range
                  / Retry-After
                  / RTP-Info
                  / Scale
                  / Session
                  / Speed
                  / Transport
                  / Unsupported
                  / Vary
                  / WWW-Authenticate
                  / extension-header

```

```

message-header  = Allow
                  / Content-Base
                  / Content-Encoding
                  / Content-Language
                  / Content-Length
                  / Content-Location
                  / Content-Type
                  / Expires
                  / Last-Modified
                  / extension-header

```

20.2.3. Header Syntax

```

Accept          = "Accept" HCOLON
                  [ accept-range *(COMMA accept-range) ]
accept-range    = media-type-range [SEMI accept-params]
media-type-range = ( "*"/*"
                  / ( m-type SLASH "*" )
                  / ( m-type SLASH m-subtype )
                  ) *( SEMI m-parameter )
accept-params   = "q" EQUAL qvalue *(SEMI generic-param )
qvalue          = ( "0" [ "." *3DIGIT ] )
                  / ( "1" [ "." *3("0") ] )
Accept-Credentials = "Accept-Credentials" HCOLON cred-decision
cred-decision   = ("User" [LWS cred-info])
                  / "Proxy"
                  / "Any"
                  / (token [LWS 1*header-value])

```

```

                                ; For future extensions
cred-info      = cred-info-data *(COMMA cred-info-data)

cred-info-data = DQ RTSP-REQ-URI DQ SEMI hash-alg SEMI base64
hash-alg       = "sha-256" / extension-alg
extension-alg   = token
Accept-Encoding = "Accept-Encoding" HCOLON
                [ encoding *(COMMA encoding) ]
encoding       = codings [SEMI accept-params]
codings        = content-coding / "*"
content-coding = token
Accept-Language = "Accept-Language" HCOLON
                language *(COMMA language)
language       = language-range [SEMI accept-params]
language-range = language-tag / "*"
language-tag   = primary-tag *( "-" subtag )
primary-tag    = 1*8ALPHA
subtag         = 1*8ALPHA
Accept-Ranges  = "Accept-Ranges" HCOLON acceptable-ranges
acceptable-ranges = (range-unit *(COMMA range-unit))
range-unit     = "NPT" / "SMPTE" / "UTC" / extension-format
extension-format = token
Allow          = "Allow" HCOLON Method *(COMMA Method)
Authorization   = "Authorization" HCOLON credentials
credentials    = ("Digest" LWS digest-response)
                / other-response
digest-response = dig-resp *(COMMA dig-resp)
dig-resp       = username / realm / nonce / digest-uri
                / dresponse / algorithm / cnonce
                / opaque / message-qop
                / nonce-count / auth-param
username       = "username" EQUAL username-value
username-value = quoted-string
digest-uri     = "uri" EQUAL LDQUOTE digest-uri-value RDQUOTE
digest-uri-value = RTSP-REQ-URI
message-qop    = "qop" EQUAL qop-value
cnonce        = "cnonce" EQUAL cnonce-value
cnonce-value   = nonce-value
nonce-count    = "nc" EQUAL nc-value
nc-value       = 8LHEX
dresponse      = "response" EQUAL request-digest
request-digest = LDQUOTE 32LHEX RDQUOTE
auth-param     = auth-param-name EQUAL
                ( token / quoted-string )
auth-param-name = token
other-response  = auth-scheme LWS auth-param
                *(COMMA auth-param)
auth-scheme     = token

```

```

Bandwidth           = "Bandwidth" HCOLON 1*19DIGIT

Blocksize           = "Blocksize" HCOLON 1*9DIGIT

Cache-Control       = "Cache-Control" HCOLON cache-directive
                      *(COMMA cache-directive)
cache-directive     = cache-rqst-directive
                      / cache-rspns-directive

cache-rqst-directive = "no-cache"
                      / "max-stale" [EQUAL delta-seconds]
                      / "min-fresh" EQUAL delta-seconds
                      / "only-if-cached"
                      / cache-extension

cache-rspns-directive = "public"
                      / "private"
                      / "no-cache"
                      / "no-transform"
                      / "must-revalidate"
                      / "proxy-revalidate"
                      / "max-age" EQUAL delta-seconds
                      / cache-extension

cache-extension     = token [EQUAL (token / quoted-string)]
delta-seconds       = 1*19DIGIT

Connection          = "Connection" HCOLON connection-token
                      *(COMMA connection-token)
connection-token    = "close" / token

Connection-Credentials = "Connection-Credentials" HCOLON cred-chain
cred-chain          = DQ RTSP-REQ-URI DQ SEMI base64

Content-Base        = "Content-Base" HCOLON RTSP-URI
Content-Encoding    = "Content-Encoding" HCOLON
                      content-coding *(COMMA content-coding)
Content-Language    = "Content-Language" HCOLON
                      language-tag *(COMMA language-tag)
Content-Length      = "Content-Length" HCOLON 1*19DIGIT
Content-Location    = "Content-Location" HCOLON RTSP-REQ-Ref
Content-Type        = "Content-Type" HCOLON media-type
media-type          = m-type SLASH m-subtype *(SEMI m-parameter)
m-type              = discrete-type / composite-type
discrete-type       = "text" / "image" / "audio" / "video"
                      / "application" / extension-token
composite-type      = "message" / "multipart" / extension-token
extension-token     = ietf-token / x-token

```

```

ietf-token      = token
x-token         = "x-" token
m-subtype       = extension-token / iana-token
iana-token      = token
m-parameter     = m-attribute EQUAL m-value
m-attribute     = token
m-value         = token / quoted-string

CSeq            = "CSeq" HCOLON cseq-nr
cseq-nr         = 1*9DIGIT
Date            = "Date" HCOLON RTSP-date
RTSP-date       = rfc1123-date ; HTTP-date
rfc1123-date    = wkday "," SP datel SP time SP "GMT"
datel           = 2DIGIT SP month SP 4DIGIT
                 ; day month year (e.g., 02 Jun 1982)
time            = 2DIGIT ":" 2DIGIT ":" 2DIGIT
                 ; 00:00:00 - 23:59:59
wkday           = "Mon" / "Tue" / "Wed"
                 / "Thu" / "Fri" / "Sat" / "Sun"
month           = "Jan" / "Feb" / "Mar" / "Apr"
                 / "May" / "Jun" / "Jul" / "Aug"
                 / "Sep" / "Oct" / "Nov" / "Dec"

Expires         = "Expires" HCOLON RTSP-date
From            = "From" HCOLON from-spec
from-spec       = ( name-addr / addr-spec ) *( SEMI from-param )
name-addr       = [ display-name ] LAQUOT addr-spec RAQUOT
addr-spec       = RTSP-REQ-URI / absolute-URI
absolute-URI    = < As defined in RFC 3986>
display-name    = *(token LWS) / quoted-string
from-param      = tag-param / generic-param
tag-param       = "tag" EQUAL token
If-Match        = "If-Match" HCOLON ("*" / message-tag-list)
message-tag-list = message-tag *(COMMA message-tag)
message-tag     = [ weak ] opaque-tag
weak            = "W/"
opaque-tag      = quoted-string
If-Modified-Since = "If-Modified-Since" HCOLON RTSP-date
If-None-Match   = "If-None-Match" HCOLON ("*" / message-tag-list)
Last-Modified   = "Last-Modified" HCOLON RTSP-date
Location        = "Location" HCOLON RTSP-REQ-URI
Media-Properties = "Media-Properties" HCOLON [media-prop-list]
media-prop-list = media-prop-value *(COMMA media-prop-value)
media-prop-value = ("Random-Access" [EQUAL POS-FLOAT])
                 / "Begining-Only"
                 / "No-Seeking"
                 / "Immutable"
                 / "Dynamic"

```

```

/ "Time-Progressing"
/ "Unlimited"
/ ("Time-Limited" EQUAL utc-time)
/ ("Time-Duration" EQUAL POS-FLOAT)
/ ("Scales" EQUAL scale-value-list)
/ media-prop-ext
media-prop-ext = token [EQUAL (1*rtsp-unreserved / quoted-string)]
scale-value-list = DQ scale-entry *(COMMA scale-entry) DQ
scale-entry = scale-value / (scale-value COLON scale-value)
scale-value = FLOAT
Media-Range = "Media-Range" HCOLON [ranges-list]
ranges-list = ranges-spec *(COMMA ranges-spec)
MTag = "MTag" HCOLON message-tag
Notify-Reason = "Notify-Reason" HCOLON Notify-Reas-val
Notify-Reas-val = "end-of-stream"
/ "media-properties-update"
/ "scale-change"
/ Notify-Reason-extension
Notify-Reason-extension = token
Pipelined-Requests = "Pipelined-Requests" HCOLON startup-id
startup-id = 1*8DIGIT
```

```

Proxy-Authenticate = "Proxy-Authenticate" HCOLON challenge-list
challenge-list    = challenge *(COMMA challenge)
challenge         = ("Digest" LWS digest-cln *(COMMA digest-cln))
                  / other-challenge
other-challenge   = auth-scheme LWS auth-param
                  *(COMMA auth-param)
digest-cln       = realm / domain / nonce
                  / opaque / stale / algorithm
                  / qop-options / auth-param
realm            = "realm" EQUAL realm-value
realm-value      = quoted-string
domain           = "domain" EQUAL LDQUOTE RTSP-REQ-Ref
                  *(1*SP RTSP-REQ-Ref ) RDQUOTE
nonce            = "nonce" EQUAL nonce-value
nonce-value      = quoted-string
opaque           = "opaque" EQUAL quoted-string
stale            = "stale" EQUAL ( "true" / "false" )
algorithm        = "algorithm" EQUAL ("MD5" / "MD5-sess" / token)
qop-options      = "qop" EQUAL LDQUOTE qop-value
                  *("," qop-value) RDQUOTE
qop-value        = "auth" / "auth-int" / token
Proxy-Require    = "Proxy-Require" HCOLON feature-tag-list
feature-tag-list = feature-tag *(COMMA feature-tag)
Proxy-Supported  = "Proxy-Supported" HCOLON [feature-tag-list]

Public           = "Public" HCOLON Method *(COMMA Method)

Range            = "Range" HCOLON ranges-spec

ranges-spec      = npt-range / utc-range / smpte-range
                  / range-ext
range-ext        = extension-format ["=" range-value]
range-value      = 1*(rtsp-unreserved / quoted-string / ":" )

Referrer         = "Referrer" HCOLON (absolute-URI / RTSP-URI-Ref)
Request-Status   = "Request-Status" HCOLON req-status-info
req-status-info  = cseq-info LWS status-info LWS reason-info
cseq-info        = "cseq" EQUAL cseq-nr
status-info      = "status" EQUAL Status-Code
reason-info      = "reason" EQUAL DQ Reason-Phrase DQ
Require         = "Require" HCOLON feature-tag-list

```



```
RTP-Info          = "RTP-Info" HCOLON [rtsp-info-spec
                        *(COMMA rtsp-info-spec)]
rtsp-info-spec    = stream-url 1*ssrc-parameter
stream-url        = "url" EQUAL DQ RTSP-REQ-Ref DQ
ssrc-parameter    = LWS "ssrc" EQUAL ssrc HCOLON
                        ri-parameter *(SEMI ri-parameter)
ri-parameter      = ("seq" EQUAL 1*5DIGIT)
                        / ("rtptime" EQUAL 1*10DIGIT)
                        / generic-param

Retry-After        = "Retry-After" HCOLON ( RTSP-date / delta-seconds )
Scale              = "Scale" HCOLON scale-value
Seek-Style         = "Seek-Style" HCOLON Seek-S-values
Seek-S-values      = "RAP"
                        / "CoRAP"
                        / "First-Prior"
                        / "Next"
                        / Seek-S-value-ext
Seek-S-value-ext   = token

Server             = "Server" HCOLON ( product / comment )
                        *(LWS (product / comment))
product            = token [SLASH product-version]
product-version    = token
comment            = LPAREN *( ctext / quoted-pair) RPAREN

Session           = "Session" HCOLON session-id
                        [ SEMI "timeout" EQUAL delta-seconds ]

Speed              = "Speed" HCOLON lower-bound MINUS upper-bound
lower-bound        = POS-FLOAT
upper-bound        = POS-FLOAT

Supported          = "Supported" HCOLON [feature-tag-list]
```

```
Terminate-Reason      = "Terminate-Reason" HCOLON TR-Info
TR-Info               = TR-Reason *(SEMI TR-Parameter)
TR-Reason             = "Session-Timeout"
                      / "Server-Admin"
                      / "Internal-Error"
                      / token
TR-Parameter         = TR-time / TR-user-msg / generic-param
TR-time              = "time" EQUAL utc-time
TR-user-msg          = "user-msg" EQUAL quoted-string

Timestamp            = "Timestamp" HCOLON timestamp-value [LWS delay]
timestamp-value      = *19DIGIT [ "." *9DIGIT ]
delay                = *9DIGIT [ "." *9DIGIT ]

Transport            = "Transport" HCOLON transport-spec
                      *(COMMA transport-spec)
transport-spec       = transport-id *trns-parameter
transport-id         = trans-id-rtp / other-trans
trans-id-rtp         = "RTP/" profile [ "/" lower-transport ]
                      ; no LWS is allowed inside transport-id
other-trans          = token *("/" token)
```

```

profile           = "AVP" / "SAVP" / "AVPF" / "SAVPF" / token
lower-transport   = "TCP" / "UDP" / token
trns-parameter    = (SEMI ( "unicast" / "multicast" ))
                  / (SEMI "interleaved" EQUAL channel [ "-" channel ])
                  / (SEMI "ttl" EQUAL ttl)
                  / (SEMI "layers" EQUAL 1*DIGIT)
                  / (SEMI "ssrc" EQUAL ssrc *(SLASH ssrc))
                  / (SEMI "mode" EQUAL mode-spec)
                  / (SEMI "dest_addr" EQUAL addr-list)
                  / (SEMI "src_addr" EQUAL addr-list)
                  / (SEMI "setup" EQUAL contrans-setup)
                  / (SEMI "connection" EQUAL contrans-con)
                  / (SEMI "RTCP-mux")
                  / (SEMI "MIKEY" EQUAL MIKEY-Value)
                  / (SEMI trn-param-ext)
contrans-setup     = "active" / "passive" / "actpass"
contrans-con       = "new" / "existing"
trn-param-ext      = par-name [EQUAL trn-par-value]
par-name           = token
trn-par-value      = *(rtsp-unreserved / quoted-string)
ttl                = 1*3DIGIT ; 0 to 255
ssrc               = 8HEX
channel            = 1*3DIGIT ; 0 to 255
MIKEY-Value        = base64
mode-spec          = ( DQ mode *(COMMA mode) DQ )
mode               = "PLAY" / token
addr-list          = quoted-addr *(SLASH quoted-addr)
quoted-addr        = DQ (host-port / extension-addr) DQ
host-port          = ( host [":" port] )
                  / ( ":" port )
extension-addr     = 1*qdttext
host               = < As defined in RFC 3986>
port               = < As defined in RFC 3986>

```

```
Unsupported      = "Unsupported" HCOLON feature-tag-list

User-Agent       = "User-Agent" HCOLON ( product / comment )
                  *(LWS (product / comment))

Vary             = "Vary" HCOLON ( "*" / field-name-list)
field-name-list  = field-name *(COMMA field-name)
field-name       = token
Via             = "Via" HCOLON via-parm *(COMMA via-parm)
via-parm        = sent-protocol LWS sent-by *( SEMI via-params )
via-params       = via-ttl / via-maddr
                  / via-received / via-extension
via-ttl          = "ttl" EQUAL ttl
via-maddr        = "maddr" EQUAL host
via-received     = "received" EQUAL (IPv4address / IPv6address)
IPv4address      = < As defined in RFC 3986>
IPv6address      = < As defined in RFC 3986>
via-extension    = generic-param
sent-protocol    = protocol-name SLASH protocol-version
                  SLASH transport-prot
protocol-name    = "RTSP" / token
protocol-version = token
transport-prot   = "UDP" / "TCP" / "TLS" / other-transport
other-transport  = token
sent-by          = host [ COLON port ]

WWW-Authenticate = "WWW-Authenticate" HCOLON challenge-list
```

20.3. SDP extension Syntax

This section defines in ABNF the SDP extensions defined for RTSP.
See Appendix D for the definition of the extensions in text.

```
control-attribute = "a=control:" *SP RTSP-REQ-Ref CRLF

a-range-def       = "a=range:" ranges-spec CRLF

a-mtag-def        = "a=mtag:" message-tag CRLF
```

21. Security Considerations

Because of the similarity in syntax and usage between RTSP servers and HTTP servers, the security considerations outlined in [H15] apply also.

Specifically, please note the following:

Abuse of Server Log Information: RTSP and HTTP servers will presumably have similar logging mechanisms, and thus should be equally guarded in protecting the contents of those logs, thus protecting the privacy of the users of the servers. See [H15.1.1] for HTTP server recommendations regarding server logs.

Transfer of Sensitive Information: There is no reason to believe that information transferred or controlled via RTSP may be any less sensitive than that normally transmitted via HTTP. Therefore, all of the precautions regarding the protection of data privacy and user privacy apply to implementors of RTSP clients, servers, and proxies. See [H15.1.2] for further details.

Attacks Based On File and Path Names: Though RTSP URIs are opaque handles that do not necessarily have file system semantics, it is anticipated that many implementations will translate portions of the Request-URIs directly to file system calls. In such cases, file systems SHOULD follow the precautions outlined in [H15.5], such as checking for ".." in path components.

Personal Information: RTSP clients are often privy to the same information that HTTP clients are (user name, location, etc.) and thus should be equally sensitive. See [H15.1] for further recommendations.

Privacy Issues Connected to Accept Headers: Since many of the same "Accept" headers exist in RTSP as in HTTP, the same caveats outlined in [H15.1.4] with regards to their use should be followed.

DNS Spoofing: Presumably, given the longer connection times typically associated to RTSP sessions relative to HTTP sessions, RTSP client DNS optimizations should be less prevalent. Nonetheless, the recommendations provided in [H15.3] are still relevant to any implementation which attempts to rely on a DNS-to-IP mapping to hold beyond a single use of the mapping.

Location Headers and Spoofing: If a single server supports multiple organizations that do not trust each another, then it needs to check the values of Location and Content-Location header fields in responses that are generated under control of said organizations to make sure that they do not attempt to invalidate resources over which they have no authority. ([H15.4])

In addition to the recommendations in the current HTTP specification (RFC 2616 [RFC2616], as of this writing) and also of the previous RFC 2068 [RFC2068], future HTTP specifications may provide additional guidance on security issues.

The following are added considerations for RTSP implementations.

Concentrated denial-of-service attack: The protocol offers the opportunity for a remote-controlled denial-of-service attack. See Section 21.1.

Session hijacking: Since there is no or little relation between a transport layer connection and an RTSP session, it is possible for a malicious client to issue requests with random session identifiers which would affect unsuspecting clients. The server SHOULD use a large, random and non-sequential session identifier to minimize the possibility of this kind of attack. However, unless the RTSP signaling is always confidentiality protected, e.g. using TLS, an on-path attacker will be able to hijack a session. For real session security, client authentication needs to be performed.

Authentication: Servers SHOULD implement both basic and digest [RFC2617] authentication. In environments requiring tighter security for the control messages, the transport layer mechanism TLS [RFC5246] SHOULD be used.

Stream issues: RTSP only provides for stream control. Stream delivery issues are not covered in this section, nor in the rest of this draft. RTSP implementations will most likely rely on other protocols such as RTP, IP multicast, RSVP and IGMP, and should address security considerations brought up in those and other applicable specifications.

Persistently suspicious behavior: RTSP servers SHOULD return error code 403 (Forbidden) upon receiving a single instance of behavior which is deemed a security risk. RTSP servers SHOULD also be aware of attempts to probe the server for weaknesses and entry points and MAY arbitrarily disconnect and ignore further requests from clients which are deemed to be in

violation of local security policy.

Scope of Multicast: If RTSP is used to control the transmission of media onto a multicast network it is needed to consider the scope that delivery has. RTSP supports the TTL Transport header parameter to indicate this scope for IPv4. However, such scope control is risks, as it may be set too large and distribute media beyond the intended scope.

TLS through proxies: If one uses the possibility to connect TLS in multiple legs (Section 19.3) one really needs to be aware of the trust model. That procedure requires full faith and trust in all proxies that one allows to connect through. They are men in the middle and have access to all that goes on over the TLS connection. Thus it is important to consider if that trust model is acceptable in the actual application.

Resource Exhaustion: As RTSP is a stateful protocol and establish resource usage on the server there is a clear possibility to attack the server by trying to overbook these resources to perform a denial of service attack. This attack can be both against ongoing sessions and to prevent others from establishing sessions. RTSP agents will need to have mechanisms to prevent single peers from consuming extensive amounts of resources.

21.1. Remote denial of Service Attack

The attacker may initiate traffic flows to one or more IP addresses by specifying them as the destination in SETUP requests. While the attacker's IP address may be known in this case, this is not always useful in prevention of more attacks or ascertaining the attackers identity. Thus, an RTSP server MUST only allow client-specified destinations for RTSP-initiated traffic flows if the server has ensured that the specified destination address accepts receiving media through different security mechanisms. Security mechanisms that are acceptable in an increased generality are:

- o Verification of the client's identity against a database of known users using RTSP authentication mechanisms (preferably digest authentication or stronger)
- o A list of addresses that accept to be media destinations, especially considering user identity
- o Media path based verification

The server SHOULD NOT allow the destination field to be set unless a

mechanism exists in the system to authorize the request originator to direct streams to the recipient. It is preferred that this authorization be performed by the media recipient (destination) itself and the credentials passed along to the server. However, in certain cases, such as when recipient address is a multicast group, or when the recipient is unable to communicate with the server in an out-of-band manner, this may not be possible. In these cases the server may choose another method such as a server-resident authorization list to ensure that the request originator has the proper credentials to request stream delivery to the recipient.

One solution that performs the necessary verification of acceptance of media suitable for unicast based delivery is the ICE based NAT traversal method described in [I-D.ietf-mmusic-rtsp-nat]. By using random passwords and username the probability of unintended indication as a valid media destination is very low. If the server include in its STUN requests a cookie (consisting of random material) that the destination echoes back the solution is also safe against having a off-path attacker being able to spoof the STUN checks. This leaves this solution vulnerable only to on-path attackers that can see the STUN requests go to the target of attack.

For delivery to multicast addresses there is a need for another solution which is not specified in this memo.

22. IANA Considerations

This section sets up a number of registries for RTSP 2.0 that should be maintained by IANA. These registries are separate from any registries existing for RTSP 1.0. For each registry there is a description on what it is required to contain, what specification is needed when adding an entry with IANA, and finally the entries that this document needs to register. See also the Section 2.7 "Extending RTSP". There is also an IANA registration of two SDP attributes.

Registries or entries in registries which have been made for RTSP 1.0 are not moved to RTSP 2.0. The registries and entries in registries of RTSP 1.0 and RTSP 2.0 are independent. If any registry or entry in a registry is also required in RTSP 2.0, it must follow the below defined procedure to allocated the registry or entry in a registry.

The sections describing how to register an item uses some of the requirements level described in RFC 5226 [RFC5226], namely "First Come, First Served", "Expert Review", "Specification Required", and "Standards Action".

In case a registry requires a contact person, the authors are the contact person for any entries created by this document.

A registration request to IANA MUST contain the following information:

- o A name of the item to register according to the rules specified by the intended registry.
- o Indication of who has change control over the feature (for example, IETF, ISO, ITU-T, other international standardization bodies, a consortium, a particular company or group of companies, or an individual);
- o A reference to a further description, if available, for example (in decreasing order of preference) an RFC, a published standard, a published paper, a patent filing, a technical report, documented source code or a computer manual;
- o For proprietary features, contact information (postal and email address);

22.1. Feature-tags

22.1.1. Description

When a client and server try to determine what part and functionality of the RTSP specification and any future extensions that its counterpart implements there is need for a namespace. This registry contains named entries representing certain functionality.

The usage of feature-tags is explained in Section 11 and Section 13.1.

22.1.2. Registering New Feature-tags with IANA

The registering of feature-tags is done on a first come, first served basis.

The name of the feature MUST follow these rules: The name may be of any length, but SHOULD be no more than twenty characters long. The name MUST NOT contain any spaces, or control characters. The registration MUST indicate if the feature-tag applies to clients, servers, or proxies only or any combinations of these. Any proprietary feature MUST have as the first part of the name a vendor tag, which identifies the organization. The registry entries consists of the feature tag, a one paragraph description of what it represents, its applicability (server, client, proxy, any combination) and a reference to its specification where applicable.

22.1.3. Registered entries

The following feature-tags are defined in this specification and hereby registered. The change control belongs to the IETF.

play.basic: The implementation for delivery and playback operations according to the core RTSP specification, as defined in this memo. Applies for both clients, servers and proxies.

play.scale: Support of scale operations for media playback. Applies only for servers.

play.speed: Support of the speed functionality for media delivery. Applies only for servers.

setup.rtp.rtcp.mux Support of the RTP and RTCP multiplexing as discussed in Appendix C.1.6.4. Applies for both client and servers and any media caching proxy.

This should be represented by IANA as table with the feature tags, contact person and their references.

22.2. RTSP Methods

22.2.1. Description

What a method is, is described in Section Section 13. Extending the protocol with new methods allow for totally new functionality.

22.2.2. Registering New Methods with IANA

A new method MUST be registered through an IETF Standards Action. The reason is that new methods may radically change the protocol's behavior and purpose.

A specification for a new RTSP method MUST consist of the following items:

- o A method name which follows the ABNF rules for methods.
- o A clear specification what a request using the method does and what responses are expected. Which directions the method is used, C->S or S->C or both. How the use of headers, if any, modifies the behavior and effect of the method.
- o A list or table specifying which of the IANA registered headers that are allowed to be used with the method in request or/and response. The list or table SHOULD follow the format of tables in Section Section 16.
- o Describe how the method relates to network proxies.

22.2.3. Registered Entries

This specification, RFCXXXX, registers 10 methods: DESCRIBE, GET_PARAMETER, OPTIONS, PAUSE, PLAY, PLAY_NOTIFY, REDIRECT, SETUP, SET_PARAMETER, and TEARDOWN. The initial table of the registry is below provided.

Method	Directionality	Reference
-----	-----	-----
DESCRIBE	C->S	[RFCXXXX]
GET_PARAMETER	C->S, S->C	[RFCXXXX]
OPTIONS	C->S, S->C	[RFCXXXX]
PAUSE	C->S	[RFCXXXX]
PLAY	C->S	[RFCXXXX]
PLAY_NOTIFY	S->C	[RFCXXXX]
REDIRECT	S->C	[RFCXXXX]
SETUP	C->S	[RFCXXXX]
SET_PARAMETER	C->S, S->C	[RFCXXXX]
TEARDOWN	C->S, S->C	[RFCXXXX]

22.3. RTSP Status Codes

22.3.1. Description

A status code is the three digit numbers used to convey information in RTSP response messages, see Section 8. The number space is limited and care should be taken not to fill the space.

22.3.2. Registering New Status Codes with IANA

A new status code registration follows the policy of IETF Review. A specification for a new status code MUST specify the following:

- o The registered number.
- o A description what the status code means and the expected behavior of the sender and receiver of the code.

22.3.3. Registered Entries

RFCXXXX, registers the numbered status code defined in the ABNF entry "Status-Code" except "extension-code" (that defines the syntax allowed for future extensions) in Section 20.2.2.

22.4. RTSP Headers

22.4.1. Description

By specifying new headers a method(s) can be enhanced in many different ways. An unknown header will be ignored by the receiving agent. If the new header is vital for a certain functionality, a feature-tag for the functionality can be created and demanded to be used by the counter-part with the inclusion of a Require header carrying the feature-tag.

22.4.2. Registering New Headers with IANA

Registrations in the registry can be done following the Expert Review policy. A specification SHOULD be provided, preferably an IETF RFC or other Standards Developing Organization specification. The minimal information in a registration request is the header name and the contact information.

The specification SHOULD contain the following information:

- o The name of the header.
- o An ABNF specification of the header syntax.
- o A list or table specifying when the header may be used, encompassing all methods, their request or response, the direction (C->S or S->C).
- o How the header is to be handled by proxies.
- o A description of the purpose of the header.

22.4.3. Registered entries

All headers specified in Section 16 in RFCXXXX are to be registered. The Registry is to include header name, description, and reference.

Furthermore the following RTSP headers defined in other specifications are registered:

- o x-wap-profile defined in [3gpp-26234].
- o x-wap-profile-diff defined in [3gpp-26234].
- o x-wap-profile-warning defined in [3gpp-26234].
- o x-predecbufsize defined in [3gpp-26234].
- o x-initpredecbufperiod defined in [3gpp-26234].
- o x-initpostdecbufperiod defined in [3gpp-26234].
- o 3gpp-videopostdecbufsize defined in [3gpp-26234].
- o 3GPP-Link-Char defined in [3gpp-26234].
- o 3GPP-Adaptation defined in [3gpp-26234].

- o 3GPP-QoE-Metrics defined in [3gpp-26234].
- o 3GPP-QoE-Feedback defined in [3gpp-26234].

The use of "x-" is NOT RECOMMENDED but the above headers in the register list was defined prior to the clarification.

22.5. Accept-Credentials

The security framework's TLS connection mechanism has two registrable entities.

22.5.1. Accept-Credentials policies

In Section 19.3.1 three policies for how to handle certificates are specified. Further policies may be defined and MUST be registered with IANA using the following rules:

- o Registering requires an IETF Standards Action
- o A registration is required to name a contact person.
- o Name of the policy.
- o A describing text that explains how the policy works for handling the certificates.

This specification registers the following values:

Any

Proxy

User

22.5.2. Accept-Credentials hash algorithms

The Accept-Credentials header (See Section 16.2) allows for the usage of other algorithms for hashing the DER records of accepted entities. The registration of any future algorithm is expected to be extremely rare and could also cause interoperability problems. Therefore the bar for registering new algorithms is intentionally placed high.

Any registration of a new hash algorithm MUST fulfill the following requirement:

- o Follow the IETF Standards Action policy.

- o A definition of the algorithm and its identifier meeting the "token" ABNF requirement.

The registered value is:

Hash Alg. Id	Reference

sha-256	[RFCXXXX]

22.6. Cache-Control Cache Directive Extensions

There exists a number of cache directives which can be sent in the Cache-Control header. A registry for these cache directives MUST be defined with the following rules:

- o Registering requires an IETF Standards Action or IESG Approval.
- o A registration is required to contain a contact person.
- o Name of the directive and a definition of the value, if any.
- o Specification if it is a request or response directive.
- o A describing text that explains how the cache directive is used for RTSP controlled media streams.

This specification registers the following values:

no-cache:

public:

private:

no-transform:

only-if-cached:

max-stale:

min-fresh:

must-revalidate:

proxy-revalidate:

max-age:

The registry should be represented as: Name of the directive, contact person and reference.

22.7. Media Properties

22.7.1. Description

The media streams being controlled by RTSP can have many different properties. The media properties required to cover the use cases that was in mind when writing the specification are defined. However, it can be expected that further innovation will result in new use cases or media streams with properties not covered by the ones specified here. Thus new media properties can be specified. As new media properties may need a substantial amount of new definitions to correctly specify behavior for this property the bar is intended to be high.

22.7.2. Registration Rules

Registering new media property MUST fulfill the following requirements

- o Follow the Specification Required policy and get the approval of the designated Expert.
- o Have an ABNF definition of the media property value name that meets "media-prop-ext" definition
- o A Contact Person for the Registration
- o Description of all changes to the behavior of the RTSP protocol as result of these changes.

22.7.3. Registered Values

This specification registers the 9 values listed in Section 16.28. The registry should be represented as: Name of the media property, contact person and reference.

22.8. Notify-Reason header

22.8.1. Description

Notify-Reason values are used for indicating the reason the notification was sent. Each reason has its associated rules on what headers and information that may or must be included in the

notification. New notification behaviors need to be specified to enable interoperable usage, thus a specification of each new value is required.

22.8.2. Registration Rules

Registrations for new Notify-Reason value MUST fulfill the following requirements

- o Follow the Specification Required policy and get the approval of the designated Expert.
- o An ABNF definition of the Notify reason value name that meets "Notify-Reason-extension" definition
- o A Contact Person for the Registration
- o Description of which headers shall be included in the request and response, when it should be sent, and any effect it has on the server client state.

22.8.3. Registered Values

This specification registers 3 values defined in the Notify-Reas-val ABNFSection 20.2.3:

- o end-of-stream
- o media-properties-update
- o scale-change

The registry entries should be represented in the registry as: Name, short description, contact and reference.

22.9. Range header formats

22.9.1. Description

The Range header (Section 16.38) allows for different range formats. New ones may be registered, but moderation should be applied as it makes interoperability more difficult.

22.9.2. Registration Rules

A registration MUST fulfill the following requirements:

- o Follow the Specification Required policy.
- o An ABNF definition of the range format that fulfills the "range-ext" definition.
- o A Contact person for the registration.
- o Rules for how one handles the range when using a negative Scale.

22.9.3. Registered Values

The registry should be represented as: Name of the range format, contact person and reference. This specification registers the following values.

npt: Normal Play Time

clock: UTC Clock format

smpte: SMPTE Timestamps

22.10. Terminate-Reason Header

The Terminate-Reason header (Section 16.50) has two registries for extensions.

22.10.1. Redirect Reasons

Registrations are done under the policy of Expert Review. The registered value needs to follow syntax, i.e. be a token. The specification needs to provide a definition of what procedures are to be followed when a client receives this redirect reason. This specification registers two values:

- o Session-Timeout
- o Server-Admin

The registry should be represented as: Name of the Redirect Reason, contact person and reference.

22.10.2. Terminate-Reason Header Parameters

Registrations are done under the policy of Specification Required. The registrations must define a syntax for the parameter that also follows the syntax allowed by the RTSP 2.0 specification. A contact person is also required. This specification registers:

- o time
- o user-msg

The registry should be represented as: Name of the Terminate Reason, contact person and reference.

22.11. RTP-Info header parameters

22.11.1. Description

The RTP-Info header (Section 16.43) carries one or more parameter value pairs with information about a particular point in the RTP stream. RTP extensions or new usages may need new types of information. As RTP information that could be needed is likely to be generic enough and to maximize the interoperability registration requires Specification Required.

22.11.2. Registration Rules

Registrations for new RTP-Info value MUST fulfill the following requirements

- o Follow the Specification Required policy and get the approval of the designated Expert.
- o Have an ABNF definition that meets the "generic-param" definition
- o A Contact Person for the Registration

22.11.3. Registered Values

This specification registers 2 parameter value pairs:

- o url
- o ssrc
- o seq
- o rtptime

The registry should be represented as: Name of the parameter, contact person and reference.

22.12. Seek-Style Policies

22.12.1. Description

New seek policies may be registered, however, a large number of these will complicate implementation substantially. The impact of unknown policies is that the server will not honor the unknown and use the server default policy instead.

22.12.2. Registration Rules

Registrations of new Seek-Style policies MUST fulfill the following requirements

- o Follow the Specification Required policy.
- o Have an ABNF definition of the Seek-Style policy name that meets "Seek-S-value-ext" definition
- o A Contact Person for the Registration
- o Description of which headers shall be included in the request and response, when it should be sent, and any affect it has on the server client state.

22.12.3. Registered Values

This specification registers 4 values:

- o RAP
- o CoRAP
- o First-Prior
- o Next

The registry should be represented as: Name of the Seek-Style Policy, short description, contact person and reference.

22.13. Transport Header Registries

The transport header contains a number of parameters which have possibilities for future extensions. Therefore registries for these need to be defined.

22.13.1. Transport Protocol Specification

A registry for the parameter transport-protocol specification MUST be defined with the following rules:

- o Registering uses the policy of Specification Required.
- o A contact person or organization with address and email.
- o A value definition that are following the ABNF syntax definition of "transport-id" Section 20.2.3.
- o A describing text that explains how the registered value are used in RTSP.

The registry should be represented as: The protocol ID string, contact person and reference.

This specification registers the following values:

RTP/AVP: Use of the RTP [RFC3550] protocol for media transport in combination with the "RTP profile for audio and video conferences with minimal control" [RFC3551] over UDP. The usage is explained in RFC XXXX, Appendix C.1.

RTP/AVP/UDP: the same as RTP/AVP.

RTP/AVPF: Use of the RTP [RFC3550] protocol for media transport in combination with the "Extended RTP Profile for RTCP-based Feedback (RTP/AVPF)" [RFC4585] over UDP. The usage is explained in RFC XXXX, Appendix C.1.

RTP/AVPF/UDP: the same as RTP/AVPF.

RTP/SAVP: Use of the RTP [RFC3550] protocol for media transport in combination with the "The Secure Real-time Transport Protocol (SRTP)" [RFC3711] over UDP. The usage is explained in RFC XXXX, Appendix C.1.

RTP/SAVP/UDP: the same as RTP/SAVP.

RTP/SAVPF: Use of the RTP[RFC3550] protocol for media transport in combination with the Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF) [RFC5124] over UDP. The usage is explained in RFC XXXX, Appendix C.1.

RTP/SAVPF/UDP: the same as RTP/SAVPF.

RTP/AVP/TCP: Use of the RTP [RFC3550] protocol for media transport in combination with the "RTP profile for audio and video conferences with minimal control" [RFC3551] over TCP. The usage is explained in RFC XXXX, Appendix C.2.2.

RTP/AVPF/TCP: Use of the RTP [RFC3550] protocol for media transport in combination with the "Extended RTP Profile for RTCP-based Feedback (RTP/AVPF)" [RFC4585] over TCP. The usage is explained in RFC XXXX, Appendix C.2.2.

RTP/SAVP/TCP: Use of the RTP [RFC3550] protocol for media transport in combination with the "The Secure Real-time Transport Protocol (SRTP)" [RFC3711] over TCP. The usage is explained in RFC XXXX, Appendix C.2.2.

RTP/SAVPF/TCP: Use of the RTP [RFC3550] protocol for media transport in combination with the "Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)" [RFC5124] over TCP. The usage is explained in RFC XXXX, Appendix C.2.2.

22.13.2. Transport modes

A registry for the transport parameter mode MUST be defined with the following rules:

- o Registering requires an IETF Standards Action.
- o A contact person or organization with address and email.
- o A value definition that are following the ABNF "token" definition Section 20.2.3.
- o A describing text that explains how the registered value are used in RTSP.

This specification registers 1 value:

PLAY: See RFC XXXX.

22.13.3. Transport Parameters

A registry for parameters that may be included in the Transport header MUST be defined with the following rules:

- o Registering uses the Specification Required policy.
- o A value definition that are following the ABNF "token" definition Section 20.2.3.
- o A describing text that explains how the registered value are used in RTSP.

This specification registers all the transport parameters defined in Section 16.52. This is a copy of this list:

- o unicast
- o multicast
- o interleaved
- o ttl
- o layers
- o ssrc
- o mode
- o dest_addr
- o src_addr
- o setup
- o connection
- o RTCP-mux
- o MIKEY

22.14. URI Schemes

This specification defines two URI schemes ("rtsp" and "rtsps") and reserves a third one ("rtspu"). These URI schemes are defined in existing registries which are not created by RTSP. Registrations are following RFC 4395[RFC4395].

22.14.1. The rtsp URI Scheme

URI scheme name: rtsp

Status: Permanent

URI scheme syntax: See Section 20.2.1 of RFC XXXX.

URI scheme semantics: The rtsp scheme is used to indicate resources accessible through the usage of the Real-time Streaming Protocol (RTSP). RTSP allows different operations on the resource identified by the URI, but the primary purpose is the streaming delivery of the resource to a client. However, the operations that are currently defined are: DESCRIBE, GET_PARAMETER, OPTIONS, PLAY, PLAY_NOTIFY, PAUSE, SETUP, SET_PARAMETER, and TEARDOWN.

Encoding considerations: IRIs in this scheme are defined and needs to be encoded as RTSP URIs when used within the RTSP protocol. That encoding is done according to RFC 3987.

Applications/protocols that use this URI scheme name: RTSP 1.0 (RFC 2326), RTSP 2.0 (RFC XXXX)

Interoperability considerations: The change in URI syntax performed between RTSP 1.0 and 2.0 can create interoperability issues.

Security considerations: All the security threats identified in Section 7 of RFC 3986 applies also to this scheme. They need to be reviewed and considered in any implementation utilizing this scheme.

Contact: Magnus Westerlund, magnus.westerlund@ericsson.com

Author/Change controller: IETF

References: RFC 2326, RFC 3986, RFC 3987, RFC XXXX

22.14.2. The rtsp URI Scheme

URI scheme name: rtsp

Status: Permanent

URI scheme syntax: See Section 20.2.1 of RFC XXXX.

URI scheme semantics: The rtsp scheme is used to indicate resources accessible through the usage of the Real-time Streaming Protocol (RTSP) over TLS. RTSP allows different operations on the resource identified by the URI, but the primary purpose is

the streaming delivery of the resource to a client. However, the operations that are currently defined are: DESCRIBE, GET_PARAMETER, OPTIONS, PLAY, PLAY_NOTIFY, PAUSE, SETUP, SET_PARAMETER, and TEARDOWN.

Encoding considerations: IRIs in this scheme are defined and needs to be encoded as RTSP URIs when used within the RTSP protocol. That encoding is done according to RFC 3987.

Applications/protocols that use this URI scheme name: RTSP 1.0 (RFC 2326), RTSP 2.0 (RFC XXXX)

Interoperability considerations: The change in URI syntax performed between RTSP 1.0 and 2.0 can create interoperability issues.

Security considerations: All the security threats identified in Section 7 of RFC 3986 applies also to this scheme. They need to be reviewed and considered in any implementation utilizing this scheme.

Contact: Magnus Westerlund, magnus.westerlund@ericsson.com

Author/Change controller: IETF

References: RFC 2326, RFC 3986, RFC 3987, RFC XXXX

22.14.3. The rtspu URI Scheme

URI scheme name: rtspu

Status: Permanent

URI scheme syntax: See Section 3.2 of RFC 2326.

URI scheme semantics: The rtspu scheme is used to indicate resources accessible through the usage of the Real-time Streaming Protocol (RTSP) over unreliable datagram transport. RTSP allows different operations on the resource identified by the URI, but the primary purpose is the streaming delivery of the resource to a client. However, the operations that are currently defined are: DESCRIBE, GET_PARAMETER, OPTIONS, PLAY, PLAY_NOTIFY, PAUSE, SETUP, SET_PARAMETER, and TEARDOWN.

Encoding considerations: IRIs in this scheme are not defined.

Applications/protocols that use this URI scheme name: RTSP 1.0 (RFC 2326)

Interoperability considerations: The definition of the transport mechanism of RTSP over UDP has interoperability issues. That makes the usage of this scheme problematic.

Security considerations: All the security threats identified in Section 7 of RFC 3986 applies also to this scheme. They needs to be reviewed and considered in any implementation utilizing this scheme.

Contact: Magnus Westerlund, magnus.westerlund@ericsson.com

Author/Change controller: IETF

References: RFC 2326

22.15. SDP attributes

This specification defines three SDP [RFC4566] attributes that it is requested that IANA register.

SDP Attribute ("att-field"):

Attribute name:	range
Long form:	Media Range Attribute
Type of name:	att-field
Type of attribute:	Media and session level
Subject to charset:	No
Purpose:	RFC XXXX
Reference:	RFC XXXX, RFC 2326
Values:	See ABNF definition.
Attribute name:	control
Long form:	RTSP control URI
Type of name:	att-field
Type of attribute:	Media and session level
Subject to charset:	No
Purpose:	RFC XXXX
Reference:	RFC XXXX, RFC 2326
Values:	Absolute or Relative URIs.
Attribute name:	mtag
Long form:	Message Tag
Type of name:	att-field
Type of attribute:	Media and session level
Subject to charset:	No
Purpose:	RFC XXXX
Reference:	RFC XXXX
Values:	See ABNF definition

22.16. Media Type Registration for text/parameters

Type name: text

Subtype name: parameters

Required parameters:

Optional parameters:

Encoding considerations:

Security considerations: This format may carry any type of parameters. Some can have security requirements, like privacy, confidentiality or integrity requirements. The format has no built in security protection. For the usage it was defined the transport can be protected between server and client using TLS. However, care must be take to consider if also the proxies are

trusted with the parameters in case hop-by-hop security is used. If stored as file in file system the necessary precautions needs to be taken in relation to the parameters requirements including object security such as S/MIME [RFC5751].

Interoperability considerations: This media type was mentioned as a fictional example in RFC 2326 but was not formally specified. This has resulted in usage of this media type which may not match its formal definition.

Published specification: RFC XXXX, Appendix F.

Applications that use this media type: Applications that use RTSP and have additional parameters they like to read and set using the RTSP GET_PARAMETER and SET_PARAMETER methods.

Additional information:

Magic number(s):

File extension(s):

Macintosh file type code(s):

Person & email address to contact for further information: Magnus Westerlund (magnus.westerlund@ericsson.com)

Intended usage: Common

Restrictions on usage: None

Author: Magnus Westerlund (magnus.westerlund@ericsson.com)

Change controller: IETF

Addition Notes:

23. References

23.1. Normative References

- [3gpp-26234] Third Generation Partnership Project (3GPP), "Transparent end-to-end Packet-switched Streaming Service (PSS); Protocols and codecs; Technical Specification 26.234", December 2002.
- [FIPS-pub-180-2] National Institute of Standards and Technology (NIST), "Federal Information Processing Standards Publications (FIPS PUBS) 180-2: Secure Hash Standard", August 2002.
- [RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, August 1980.
- [RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.
- [RFC2617] Franks, J., Hallam-Baker, P., Hostetler, J., Lawrence, S., Leach, P., Luotonen, A., and L. Stewart, "HTTP Authentication: Basic and Digest Access Authentication", RFC 2617, June 1999.
- [RFC2818] Rescorla, E., "HTTP Over TLS", RFC 2818, May 2000.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.
- [RFC3551] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, July 2003.
- [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, RFC 3629, November 2003.
- [RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)",

RFC 3711, March 2004.

- [RFC3830] Arkko, J., Carrara, E., Lindholm, F., Naslund, M., and K. Norrman, "MIKEY: Multimedia Internet KEYing", RFC 3830, August 2004.
- [RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.
- [RFC3987] Duerst, M. and M. Suignard, "Internationalized Resource Identifiers (IRIs)", RFC 3987, January 2005.
- [RFC4086] Eastlake, D., Schiller, J., and S. Crocker, "Randomness Requirements for Security", BCP 106, RFC 4086, June 2005.
- [RFC4288] Freed, N. and J. Klensin, "Media Type Specifications and Registration Procedures", BCP 13, RFC 4288, December 2005.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.
- [RFC4395] Hansen, T., Hardie, T., and L. Masinter, "Guidelines and Registration Procedures for New URI Schemes", BCP 35, RFC 4395, February 2006.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC4567] Arkko, J., Lindholm, F., Naslund, M., Norrman, K., and E. Carrara, "Key Management Extensions for Session Description Protocol (SDP) and Real Time Streaming Protocol (RTSP)", RFC 4567, July 2006.
- [RFC4571] Lazzaro, J., "Framing Real-time Transport Protocol (RTP) and RTP Control Protocol (RTCP) Packets over Connection-Oriented Transport", RFC 4571, July 2006.
- [RFC4585] Ott, J., Wenger, S., Sato, N., Burmeister, C., and J. Rey, "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)", RFC 4585, July 2006.
- [RFC4648] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", RFC 4648, October 2006.
- [RFC4738] Ignjatic, D., Dondeti, L., Audet, F., and P. Lin, "MIKEY-RSA-R: An Additional Mode of Key Distribution in

Multimedia Internet KEYing (MIKEY)", RFC 4738,
November 2006.

- [RFC5124] Ott, J. and E. Carrara, "Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)", RFC 5124, February 2008.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.
- [RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, August 2008.
- [RFC5280] Cooper, D., Santesson, S., Farrell, S., Boeyen, S., Housley, R., and W. Polk, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", RFC 5280, May 2008.
- [RFC5646] Phillips, A. and M. Davis, "Tags for Identifying Languages", BCP 47, RFC 5646, September 2009.
- [RFC5751] Ramsdell, B. and S. Turner, "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Message Specification", RFC 5751, January 2010.
- [RFC5761] Perkins, C. and M. Westerlund, "Multiplexing RTP Data and Control Packets on a Single Port", RFC 5761, April 2010.

23.2. Informative References

- [I-D.ietf-mmusic-rtsp-nat]
Goldberg, J., Westerlund, M., and T. Zeng, "A Network Address Translator (NAT) Traversal mechanism for media controlled by Real-Time Streaming Protocol (RTSP)", draft-ietf-mmusic-rtsp-nat-11 (work in progress), October 2011.
- [ISO.13818-6.1995]
International Organization for Standardization,
"Information technology - Generic coding of moving pictures and associated audio information - part 6: Extension for digital storage media and control",
ISO Draft Standard 13818-6, November 1995.

- [ISO.8601.2000] International Organization for Standardization, "Data elements and interchange formats - Information interchange - Representation of dates and times", ISO/IEC Standard 8601, December 2000.
- [RFC1123] Braden, R., "Requirements for Internet Hosts - Application and Support", STD 3, RFC 1123, October 1989.
- [RFC1644] Braden, B., "T/TCP -- TCP Extensions for Transactions Functional Specification", RFC 1644, July 1994.
- [RFC2068] Fielding, R., Gettys, J., Mogul, J., Nielsen, H., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2068, January 1997.
- [RFC2326] Schulzrinne, H., Rao, A., and R. Lanphier, "Real Time Streaming Protocol (RTSP)", RFC 2326, April 1998.
- [RFC2663] Srisuresh, P. and M. Holdrege, "IP Network Address Translator (NAT) Terminology and Considerations", RFC 2663, August 1999.
- [RFC2822] Resnick, P., "Internet Message Format", RFC 2822, April 2001.
- [RFC2974] Handley, M., Perkins, C., and E. Whelan, "Session Announcement Protocol", RFC 2974, October 2000.
- [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, June 2002.
- [RFC4145] Yon, D. and G. Camarillo, "TCP-Based Media Transport in the Session Description Protocol (SDP)", RFC 4145, September 2005.
- [RFC5583] Schierl, T. and S. Wenger, "Signaling Media Decoding Dependency in the Session Description Protocol (SDP)", RFC 5583, July 2009.
- [RFC5888] Camarillo, G. and H. Schulzrinne, "The Session Description Protocol (SDP) Grouping Framework", RFC 5888, June 2010.
- [RFC5905] Mills, D., Martin, J., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, June 2010.

[Stevens98]

Stevens, W., "Unix Networking Programming - Volume 1,
second edition", 1998.

Appendix A. Examples

This section contains several different examples trying to illustrate possible ways of using RTSP. The examples can also help with the understanding of how functions of RTSP work. However, remember that these are examples and the normative and syntax description in the other sections takes precedence. Please also note that many of the examples contain syntax illegal line breaks to accommodate the formatting restriction that the RFC series impose.

A.1. Media on Demand (Unicast)

This is an example of media on demand streaming of a media stored in a container file. For purposes of this example, a container file is a storage entity in which multiple continuous media types pertaining to the same end-user presentation are present. In effect, the container file represents an RTSP presentation, with each of its components being RTSP controlled media streams. Container files are a widely used means to store such presentations. While the components are transported as independent streams, it is desirable to maintain a common context for those streams at the server end.

This enables the server to keep a single storage handle open easily. It also allows treating all the streams equally in case of any prioritization of streams by the server.

It is also possible that the presentation author may wish to prevent selective retrieval of the streams by the client in order to preserve the artistic effect of the combined media presentation. Similarly, in such a tightly bound presentation, it is desirable to be able to control all the streams via a single control message using an aggregate URI.

The following is an example of using a single RTSP session to control multiple streams. It also illustrates the use of aggregate URIs. In a container file it is also desirable to not write any URI parts which is not kept, when the container is distributed, like the host and most of the path element. Therefore this example also uses the "*" and relative URI in the delivered SDP.

Also this presentation description (SDP) is not cachable, as the Expires header is set to an equal value with date indicating immediate expiration of its validity.

Client C requests a presentation from media server M. The movie is stored in a container file. The client has obtained an RTSP URI to the container file.

```
C->M: DESCRIBE rtsp://example.com/twister.3gp RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2
```

```
M->C: RTSP/2.0 200 OK
      CSeq: 1
      Server: PhonyServer/1.0
      Date: Thu, 24 Jan 1997 15:35:06 GMT
      Content-Type: application/sdp
      Content-Length: 271
      Content-Base: rtsp://example.com/twister.3gp/
      Expires: 24 Jan 1997 15:35:06 GMT
```

```
v=0
o=- 2890844256 2890842807 IN IP4 198.51.100.5
s=RTSP Session
i=An Example of RTSP Session Usage
e=adm@example.com
c=IN IP4 0.0.0.0
a=control: *
a=range: npt=0-0:10:34.10
t=0 0
m=audio 0 RTP/AVP 0
a=control: trackID=1
m=video 0 RTP/AVP 26
a=control: trackID=4
```

C->M: SETUP rtsp://example.com/twister.3gp/trackID=1 RTSP/2.0
CSeq: 2
User-Agent: PhonyClient/1.2
Require: play.basic
Transport: RTP/AVP;unicast;dest_addr=":8000"/":8001"
Accept-Ranges: NPT, SMPTE, UTC

M->C: RTSP/2.0 200 OK
CSeq: 2
Server: PhonyServer/1.0
Transport: RTP/AVP;unicast; ssrc=93CB001E;
 dest_addr="192.0.2.53:8000"/"192.0.2.53:8001";
 src_addr="198.51.100.5:9000"/"198.51.100.5:9001"
Session: 12345678
Expires: 24 Jan 1997 15:35:12 GMT
Date: 24 Jan 1997 15:35:12 GMT
Accept-Ranges: NPT
Media-Properties: Random-Access=0.02, Immutable, Unlimited

C->M: SETUP rtsp://example.com/twister.3gp/trackID=4 RTSP/2.0
CSeq: 3
User-Agent: PhonyClient/1.2
Require: play.basic
Transport: RTP/AVP;unicast;dest_addr=":8002"/":8003"
Session: 12345678
Accept-Ranges: NPT, SMPTE, UTC

M->C: RTSP/2.0 200 OK
CSeq: 3
Server: PhonyServer/1.0
Transport: RTP/AVP;unicast; ssrc=A813FC13;
 dest_addr="192.0.2.53:8002"/"192.0.2.53:8003";
 src_addr="198.51.100.5:9002"/"198.51.100.5:9003";

Session: 12345678
Expires: 24 Jan 1997 15:35:13 GMT
Date: 24 Jan 1997 15:35:13 GMT
Accept-Range: NPT
Media-Properties: Random-Access=0.8, Immutable, Unlimited

C->M: PLAY rtsp://example.com/twister.3gp/ RTSP/2.0
CSeq: 4
User-Agent: PhonyClient/1.2
Range: npt=30-
Seek-Style: RAP
Session: 12345678

M->C: RTSP/2.0 200 OK
CSeq: 4
Server: PhonyServer/1.0
Date: 24 Jan 1997 15:35:14 GMT
Session: 12345678
Range: npt=30-634.10
Seek-Style: RAP
RTP-Info: url="rtsp://example.com/twister.3gp/trackID=4"
 ssrc=0D12F123:seq=12345;rtptime=3450012,
 url="rtsp://example.com/twister.3gp/trackID=1"
 ssrc=4F312DD8:seq=54321;rtptime=2876889

C->M: PAUSE rtsp://example.com/twister.3gp/ RTSP/2.0
CSeq: 5
User-Agent: PhonyClient/1.2
Session: 12345678

M->C: RTSP/2.0 200 OK
CSeq: 5
Server: PhonyServer/1.0
Date: 24 Jan 1997 15:36:01 GMT
Session: 12345678
Range: npt=30.87-634.10

C->M: PLAY rtsp://example.com/twister.3gp/ RTSP/2.0
CSeq: 6
User-Agent: PhonyClient/1.2
Range: npt=30.87-634.10
Seek-Style: Next
Session: 12345678

M->C: RTSP/2.0 200 OK
CSeq: 6
Server: PhonyServer/1.0
Date: 24 Jan 1997 15:36:01 GMT
Session: 12345678
Range: npt=30.87-634.10
Seek-Style: Next
RTP-Info: url="rtsp://example.com/twister.3gp/trackID=4"
 ssrc=0D12F123:seq=12555;rtptime=6330012,
 url="rtsp://example.com/twister.3gp/trackID=1"
 ssrc=4F312DD8:seq=55021;rtptime=3132889

C->M: TEARDOWN rtsp://example.com/twister.3gp/ RTSP/2.0
CSeq: 7
User-Agent: PhonyClient/1.2
Session: 12345678

```
M->C: RTSP/2.0 200 OK
      CSeq: 7
      Server: PhonyServer/1.0
      Date: 24 Jan 1997 15:49:34 GMT
```

A.2. Media on Demand using Pipelining

This example is basically the example above (Appendix A.1), but now utilizing pipelining to speed up the setup. It requires only two round trip times until the media starts flowing. First of all, the session description is retrieved to determine what media resources need to be setup. In the second step, one sends the necessary SETUP requests and the PLAY request to initiate media delivery.

Client C requests a presentation from media server M. The movie is stored in a container file. The client has obtained an RTSP URI to the container file.

```
C->M: DESCRIBE rtsp://example.com/twister.3gp RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2
```

```
M->C: RTSP/2.0 200 OK
      CSeq: 1
      Server: PhonyServer/1.0
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Content-Type: application/sdp
      Content-Length: 271
      Content-Base: rtsp://example.com/twister.3gp/
      Expires: 24 Jan 1997 15:35:06 GMT
```

```
v=0
o=- 2890844256 2890842807 IN IP4 192.0.2.5
s=RTSP Session
i=An Example of RTSP Session Usage
e=adm@example.com
c=IN IP4 0.0.0.0
a=control: *
a=range: npt=0-0:10:34.10
t=0 0
m=audio 0 RTP/AVP 0
a=control: trackID=1
m=video 0 RTP/AVP 26
a=control: trackID=4
```

```
C->M: SETUP rtsp://example.com/twister.3gp/trackID=1 RTSP/2.0
      CSeq: 2
      User-Agent: PhonyClient/1.2
```

Require: play.basic
Transport: RTP/AVP;unicast;dest_addr=":8000"/":8001"
Accept-Ranges: NPT, SMPTE, UTC
Pipelined-Requests: 7654

C->M: SETUP rtsp://example.com/twister.3gp/trackID=4 RTSP/2.0
CSeq: 3
User-Agent: PhonyClient/1.2
Require: play.basic
Transport: RTP/AVP;unicast;dest_addr=":8002"/":8003"
Accept-Ranges: NPT, SMPTE, UTC
Pipelined-Requests: 7654

C->M: PLAY rtsp://example.com/twister.3gp/ RTSP/2.0
CSeq: 4
User-Agent: PhonyClient/1.2
Range: npt=0-
Seek-Style: RAP
Pipelined-Requests: 7654

M->C: RTSP/2.0 200 OK
CSeq: 2
Server: PhonyServer/1.0
Transport: RTP/AVP;unicast;
 dest_addr="192.0.2.53:8000"/"192.0.2.53:8001";
 src_addr="198.51.100.5:9000"/"198.51.100.5:9001";
 ssrc=93CB001E
Session: 12345678
Expires: 24 Jan 1997 15:35:12 GMT
Date: 23 Jan 1997 15:35:12 GMT
Accept-Ranges: NPT
Pipelined-Requests: 7654
Media-Properties: Random-Access=0.2, Immutable, Unlimited

M->C: RTSP/2.0 200 OK
CSeq: 3
Server: PhonyServer/1.0
Transport: RTP/AVP;unicast;
 dest_addr="192.0.2.53:8002"/"192.0.2.53:8003";
 src_addr="198.51.100.5:9002"/"198.51.100.5:9003";
 ssrc=A813FC13
Session: 12345678
Expires: 24 Jan 1997 15:35:13 GMT
Date: 23 Jan 1997 15:35:13 GMT
Accept-Range: NPT
Pipelined-Requests: 7654
Media-Properties: Random-Access=0.8, Immutable, Unlimited

```
M->C: RTSP/2.0 200 OK
      CSeq: 4
      Server: PhonyServer/1.0
      Date: 23 Jan 1997 15:35:14 GMT
      Session: 12345678
      Range: npt=0-623.10
      Seek-Style: RAP
      RTP-Info: url="rtsp://example.com/twister.3gp/trackID=4"
                ssrc=0D12F123:seq=12345;rtptime=3450012,
                url="rtsp://example.com/twister.3gp/trackID=1"
                ssrc=4F312DD8:seq=54321;rtptime=2876889
      Pipelined-Requests: 7654
```

A.3. Media on Demand (Unicast)

An alternative example of media on demand with a bit more tweaks is the following. Client C requests a movie distributed from two different media servers A (audio.example.com) and V (video.example.com). The media description is stored on a web server W. The media description contains descriptions of the presentation and all its streams, including the codecs that are available, dynamic RTP payload types, the protocol stack, and content information such as language or copyright restrictions. It may also give an indication about the timeline of the movie.

In this example, the client is only interested in the last part of the movie.

C->W: GET /twister.sdp HTTP/1.1
Host: www.example.com
Accept: application/sdp

W->C: HTTP/1.0 200 OK
Date: Thu, 23 Jan 1997 15:35:06 GMT
Content-Type: application/sdp
Content-Length: 278
Expires: 23 Jan 1998 15:35:06 GMT

v=0
o=- 2890844526 2890842807 IN IP4 198.51.100.5
s=RTSP Session
e=adm@example.com
c=IN IP4 0.0.0.0
a=range:npt=0-1:49:34
t=0 0
m=audio 0 RTP/AVP 0
a=control:rtsp://audio.example.com/twister/audio.en
m=video 0 RTP/AVP 31
a=control:rtsp://video.example.com/twister/video

C->A: SETUP rtsp://audio.example.com/twister/audio.en RTSP/2.0
CSeq: 1
User-Agent: PhonyClient/1.2
Transport: RTP/AVP/UDP;unicast;dest_addr=":3056"/":3057",
 RTP/AVP/TCP;unicast;interleaved=0-1
Accept-Ranges: NPT, SMPTE, UTC

A->C: RTSP/2.0 200 OK
CSeq: 1
Session: 12345678
Transport: RTP/AVP/UDP;unicast;
 dest_addr="192.0.2.53:3056"/"192.0.2.53:3057";
 src_addr="198.51.100.5:5000"/"198.51.100.5:5001"
Date: 23 Jan 1997 15:35:12 GMT
Server: PhonyServer/1.0
Expires: 24 Jan 1997 15:35:12 GMT
Cache-Control: public
Accept-Ranges: NPT, SMPTE
Media-Properties: Random-Access=0.02, Immutable, Unlimited

```
C->V: SETUP rtsp://video.example.com/twister/video RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2
      Transport: RTP/AVP/UDP;unicast;
                  dest_addr="192.0.2.53:3058"/"192.0.2.53:3059",
                  RTP/AVP/TCP;unicast;interleaved=0-1
      Accept-Ranges: NPT, SMPTE, UTC

V->C: RTSP/2.0 200 OK
      CSeq: 1
      Session: 23456789
      Transport: RTP/AVP/UDP;unicast;
                  dest_addr="192.0.2.53:3058"/"192.0.2.53:3059";
                  src_addr="198.51.100.5:5002"/"198.51.100.5:5003"
      Date: 23 Jan 1997 15:35:12 GMT
      Server: PhonyServer/1.0
      Cache-Control: public
      Expires: 24 Jan 1997 15:35:12 GMT
      Accept-Ranges: NPT, SMPTE
      Media-Properties: Random-Access=1.2, Immutable, Unlimited

C->V: PLAY rtsp://video.example.com/twister/video RTSP/2.0
      CSeq: 2
      User-Agent: PhonyClient/1.2
      Session: 23456789
      Range: smpte=0:10:00-

V->C: RTSP/2.0 200 OK
      CSeq: 2
      Session: 23456789
      Range: smpte=0:10:00-1:49:23
      Seek-Style: First-Prior
      RTP-Info: url="rtsp://video.example.com/twister/video"
                  ssrc=A17E189D:seq=12312232;rtptime=78712811
      Server: PhonyServer/2.0
      Date: 23 Jan 1997 15:35:13 GMT
```

```
C->A: PLAY rtsp://audio.example.com/twister/audio.en RTSP/2.0
      CSeq: 2
      User-Agent: PhonyClient/1.2
      Session: 12345678
      Range: smpte=0:10:00-

A->C: RTSP/2.0 200 OK
      CSeq: 2
      Session: 12345678
      Range: smpte=0:10:00-1:49:23
      Seek-Style: First-Prior
      RTP-Info: url="rtsp://audio.example.com/twister/audio.en"
                ssrc=3D124F01:seq=876655;rtptime=1032181
      Server: PhonyServer/1.0
      Date: 23 Jan 1997 15:35:13 GMT
```

```
C->A: TEARDOWN rtsp://audio.example.com/twister/audio.en RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Session: 12345678
```

```
A->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/1.0
      Date: 23 Jan 1997 15:36:52 GMT
```

```
C->V: TEARDOWN rtsp://video.example.com/twister/video RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Session: 23456789
```

```
V->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/2.0
      Date: 23 Jan 1997 15:36:52 GMT
```

Even though the audio and video track are on two different servers that may start at slightly different times and may drift with respect to each other over time, the client can perform initial synchronization of the two media using RTP-Info and Range received in the PLAY responses. If the two servers are time synchronized the RTCP packets can also be used to maintain synchronization.

A.4. Single Stream Container Files

Some RTSP servers may treat all files as though they are "container files", yet other servers may not support such a concept. Because of this, clients needs to use the rules set forth in the session description for Request-URIs, rather than assuming that a consistent URI may always be used throughout. Below is an example of how a multi-stream server might expect a single-stream file to be served:

```
C->S: DESCRIBE rtsp://foo.example.com/test.wav RTSP/2.0
      Accept: application/x-rtsp-mh, application/sdp
      CSeq: 1
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 1
      Content-base: rtsp://foo.example.com/test.wav/
      Content-type: application/sdp
      Content-length: 163
      Server: PhonyServer/1.0
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Expires: 23 Jan 1997 17:00:00 GMT
```

```
v=0
o=- 872653257 872653257 IN IP4 192.0.2.5
s=mu-law wave file
i=audio test
c=IN IP4 0.0.0.0
t=0 0
a=control: *
m=audio 0 RTP/AVP 0
a=control:streamid=0
```

```
C->S: SETUP rtsp://foo.example.com/test.wav/streamid=0 RTSP/2.0
      Transport: RTP/AVP/UDP;unicast;
            dest_addr=":6970"/":6971";mode="PLAY"
      CSeq: 2
      User-Agent: PhonyClient/1.2
      Accept-Ranges: NPT, SMPTE, UTC
```

```
S->C: RTSP/2.0 200 OK
      Transport: RTP/AVP/UDP;unicast;
            dest_addr="192.0.2.53:6970"/"192.0.2.53:6971";
            src_addr="198.51.100.5:6970"/"198.51.100.5:6971";
            mode="PLAY";ssrc=EAB98712
      CSeq: 2
      Session: 2034820394
      Expires: 23 Jan 1997 16:00:00 GMT
      Server: PhonyServer/1.0
      Date: 23 Jan 1997 15:35:07 GMT
      Accept-Ranges: NPT
      Media-Properties: Random-Acces=0.5, Immutable, Unlimited
```

```
C->S: PLAY rtsp://foo.example.com/test.wav/ RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Session: 2034820394
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/1.0
      Date: 23 Jan 1997 15:35:08 GMT
      Session: 2034820394
      Range: npt=0-600
      Seek-Style: RAP
      RTP-Info: url="rtsp://foo.example.com/test.wav/streamid=0"
            ssrc=0D12F123:seq=981888;rtptime=3781123
```

Note the different URI in the SETUP command, and then the switch back to the aggregate URI in the PLAY command. This makes complete sense when there are multiple streams with aggregate control, but is less than intuitive in the special case where the number of streams is one. However, the server has declared the aggregated control URI in the SDP and therefore this is legal.

In this case, it is also required that servers accept implementations that use the non-aggregated interpretation and use the individual media URI, like this:

```
C->S: PLAY rtsp://example.com/test.wav/streamid=0 RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Session: 2034820394
```

A.5. Live Media Presentation Using Multicast

The media server M chooses the multicast address and port. Here, it is assumed that the web server only contains a pointer to the full description, while the media server M maintains the full description.

```
C->W: GET /sessions.html HTTP/1.1
      Host: www.example.com
```

```
W->C: HTTP/1.1 200 OK
      Content-Type: text/html
```

```
<html>
...
  <a href "rtsp://live.example.com/concert/audio">
    Streamed Live Music performance </a>
...
</html>
```

```
C->M: DESCRIBE rtsp://live.example.com/concert/audio RTSP/2.0
      CSeq: 1
      Supported: play.basic, play.scale
      User-Agent: PhonyClient/1.2
```

```
M->C: RTSP/2.0 200 OK
      CSeq: 1
      Content-Type: application/sdp
      Content-Length: 183
      Server: PhonyServer/1.0
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Supported: play.basic
```

```
v=0
o=- 2890844526 2890842807 IN IP4 192.0.2.5
s=RTSP Session
t=0 0
m=audio 3456 RTP/AVP 0
c=IN IP4 233.252.0.54/16
a=control: rtsp://live.example.com/concert/audio
a=range:npt=0-
```

```
C->M: SETUP rtsp://live.example.com/concert/audio RTSP/2.0
      CSeq: 2
      Transport: RTP/AVP;multicast;
               dest_addr="233.252.0.54:3456"/"233.252.0.54:3457";ttl=16
      Accept-Ranges: NPT, SMPTE, UTC
      User-Agent: PhonyClient/1.2

M->C: RTSP/2.0 200 OK
      CSeq: 2
      Server: PhonyServer/1.0
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Transport: RTP/AVP;multicast;
               dest_addr="233.252.0.54:3456"/"233.252.0.54:3457";ttl=16
               ;ssrc=4D12AB92/0DF876A3
      Session: 0456804596
      Accept-Ranges: NPT, UTC
      Media-Properties: No-Seeking, Time-Progressing, Time-Duration=0

C->M: PLAY rtsp://live.example.com/concert/audio RTSP/2.0
      CSeq: 3
      Session: 0456804596
      User-Agent: PhonyClient/1.2

M->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/1.0
      Date: 23 Jan 1997 15:35:07 GMT
      Session: 0456804596
      Seek-Style: Next
      Range:npt=1256-
      RTP-Info: url="rtsp://live.example.com/concert/audio"
               ssrc=0D12F123;seq=1473; rtptime=80000
```

A.6. Capability Negotiation

This example illustrates how the client and server determines their capability to support a special feature, in this case "play.scale". The server, through the clients request and the included Supported header, learns the client supports RTSP 2.0, and also supports the playback time scaling feature of RTSP. The server's response contains the following feature related information to the client; it supports the basic media delivery functions (play.basic), the extended functionality of time scaling of content (play.scale), and one "example.com" proprietary feature (com.example.flight). The client also learns the methods supported (Public header) by the server for the indicated resource.

```
C->S: OPTIONS rtsp://media.example.com/movie/twister.3gp RTSP/2.0
      CSeq: 1
      Supported: play.basic, play.scale
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 1
      Public: OPTIONS, SETUP, PLAY, PAUSE, TEARDOWN, DESCRIBE, GET_PARAMETER
      Allow: OPTIONS, SETUP, PLAY, PAUSE, TEARDOWN, DESCRIBE
      Server: PhonyServer/2.0
      Supported: play.basic, play.scale, com.example.flight
```

When the client sends its SETUP request it tells the server that it requires support of the play.scale feature for this session by including the Require header.

```
C->S: SETUP rtsp://media.example.com/twister.3gp/trackID=1 RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Transport: RTP/AVP/UDP;unicast;
                 dest_addr="192.0.2.53:3056"/"192.0.2.53:3057",
                 RTP/AVP/TCP;unicast;interleaved=0-1
      Require: play.scale
      Accept-Ranges: NPT, SMPTE, UTC
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 3
      Session: 12345678
      Transport: RTP/AVP/UDP;unicast;
                 dest_addr="192.0.2.53:3056"/"192.0.2.53:3057";
                 src_addr="198.51.100.5:5000"/"198.51.100.5:5001"
      Server: PhonyServer/2.0
      Accept-Ranges: NPT, SMPTE
      Media-Properties: Random-Access=0.8, Immutable, Unlimited
```


Appendix B. RTSP Protocol State Machine

The RTSP session state machine describes the behavior of the protocol from RTSP session initialization through RTSP session termination.

The State machine is defined on a per session basis which is uniquely identified by the RTSP session identifier. The session may contain one or more media streams depending on state. If a single media stream is part of the session it is in non-aggregated control. If two or more is part of the session it is in aggregated control.

The below state machine is an informative description of the protocols behavior. In case of ambiguity with the earlier parts of this specification, the description in the earlier parts take precedence.

B.1. States

The state machine contains three states, described below. For each state there exist a table which shows which requests and events are allowed and whether they will result in a state change.

Init: Initial state no session exists.

Ready: Session is ready to start playing.

Play: Session is playing, i.e. sending media stream data in the direction S->C.

B.2. State variables

This representation of the state machine needs more than its state to work. A small number of variables is also needed and they are explained below.

NRM: The number of media streams part of this session.

RP: Resume point, the point in the presentation time line at which a request to continue playing will resume from. A time format for the variable is not mandated.

B.3. Abbreviations

To make the state tables more compact a number of abbreviations are used, which are explained below.

IFI: IF Implemented.

md: Media

PP: Pause Point, the point in the presentation time line at which the presentation was paused.

Prs: Presentation, the complete multimedia presentation.

RedP: Redirect Point, the point in the presentation time line at which a REDIRECT was specified to occur.

SES: Session.

B.4. State Tables

This section contains a table for each state. The table contains all the requests and events that this state is allowed to act on. The events which are method names are, unless noted, requests with the given method in the direction client to server (C->S). In some cases there exist one or more requisite. The response column tells what type of response actions should be performed. Possible actions that are requested for an event includes: response codes, e.g. 200, headers that needs to be included in the response, setting of state variables, or setting of other session related parameters. The new state column tells which state the state machine changes to.

The response to a valid request meeting the requisites is normally a 2xx (SUCCESS) unless other noted in the response column. The exceptions need to be given a response according to the response column. If the request does not meet the requisite, is erroneous or some other type of error occur, the appropriate response code is to be sent. If the response code is a 4xx the session state is unchanged. A response code of 3rr will result in that the session is ended and its state is changed to Init. A response code of 304 results in no state change. However, there are restrictions to when a 3rr response may be used. A 5xx response does not result in any change of the session state, except if the error is not possible to recover from. A unrecoverable error results in the ending of the session. As it in the general case can't be determined if it was a unrecoverable error or not the client will be required to test. In the case that the next request after a 5xx is responded with 454 (Session Not Found) the client knows that the session has ended. For any request message that cannot be responded to within the time defined in Section 10.4, a 100 response must be sent.

The server will timeout the session after the period of time specified in the SETUP response, if no activity from the client is

detected. Therefore there exists a timeout event for all states except Init.

In the case that $NRM = 1$ the presentation URI is equal to the media URI or a specified presentation URI. For $NRM > 1$ the presentation URI needs to be other than any of the medias that are part of the session. This applies to all states.

Event	Prerequisite	Response
DESCRIBE	Needs REDIRECT	3rr, Redirect
DESCRIBE		200, Session description
OPTIONS	Session ID	200, Reset session timeout timer
OPTIONS		200
SET_PARAMETER	Valid parameter	200, change value of parameter
GET_PARAMETER	Valid parameter	200, return value of parameter

Table 13: None state-machine changing events

The methods in Table 13 do not have any effect on the state machine or the state variables. However, some methods do change other session related parameters, for example SET_PARAMETER which will set the parameter(s) specified in its body. Also all of these methods that allow Session header will also update the keep-alive timer for the session.

Action	Requisite	New State	Response
SETUP		Ready	$NRM=1$, $RP=0.0$
SETUP	Needs Redirect	Init	3rr Redirect
S -> C: REDIRECT	No Session hdr	Init	Terminate all SES

Table 14: State: Init

The initial state of the state machine, see Table 14 can only be left by processing a correct SETUP request. As seen in the table the two

state variables are also set by a correct request. This table also shows that a correct SETUP can in some cases be redirected to another URI and/or server by a 3rr response.

Action	Requisite	New State	Response
SETUP	New URI	Ready	NRM +=1
SETUP	URI Setup prior	Ready	Change transport param
TEARDOWN	Prs URI,	Init	No session hdr, NRM = 0
TEARDOWN	md URI, NRM=1	Init	No Session hdr, NRM = 0
TEARDOWN	md URI, NRM>1	Ready	Session hdr, NRM -= 1
PLAY	Prs URI, No range	Play	Play from RP
PLAY	Prs URI, Range	Play	According to range
PLAY	md URI, NRM=1, Range	Play	According to range
PLAY	md URI, NRM=1	Play	Play from RP
PAUSE	Prs URI	Ready	Return PP
SC:REDIRECT	Terminate-Reason	Ready	Set RedP
SC:REDIRECT	No Terminate-Reason time parameter	Init	Session is removed
Timeout		Init	
RedP reached		Init	TEARDOWN of session

Table 15: State: Ready

In the Ready state, see Table 15, some of the actions are depending

on the number of media streams (NRM) in the session, i.e., aggregated or non-aggregated control. A SETUP request in the Ready state can either add one more media stream to the session or, if the media stream (same URI) already is part of the session, change the transport parameters. TEARDOWN is depending on both the Request-URI and the number of media stream within the session. If the Request-URI is the presentations URI the whole session is torn down. If a media URI is used in the TEARDOWN request and more than one media exists in the session, the session will remain and a session header is returned in the response. If only a single media stream remains in the session when performing a TEARDOWN with a media URI the session is removed. The number of media streams remaining after tearing down a media stream determines the new state.

Action	Requisite	New State	Response
PAUSE	Prs URI	Ready	Set RP to present point
End of media	All media	Play	Set RP = End of media
End of range		Play	Set RP = End of range
PLAY	Prs URI, No range	Play	Play from present point
PLAY	Prs URI, Range	Play	According to range
SC:PLAY_NOTIFY		Play	200
SETUP	New URI	Play	455
SETUP	Setuped URI	Play	455
SETUP	Setuped URI, IFI	Play	Change transport param.
TEARDOWN	Prs URI	Init	No session hdr
TEARDOWN	md URI,NRM=1	Init	No Session hdr, NRM=0
TEARDOWN	md URI	Play	455
SC:REDIRECT	Terminate Reason with Time parameter	Play	Set RedP
SC:REDIRECT		Init	Session is removed
RedP reached		Init	TEARDOWN of session
Timeout		Init	Stop Media playout

Table 16: State: Play

The Play state table, see Table 16, contains a number of requests that need a presentation URI (labeled as Prs URI) to work on (i.e., the presentation URI has to be used as the Request-URI). This is due to the exclusion of non-aggregated stream control in sessions with more than one media stream.

To avoid inconsistencies between the client and server, automatic state transitions are avoided. This can be seen at for example "End of media" event when all media has finished playing, the session still remains in Play state. An explicit PAUSE request needs to be sent to change the state to Ready. It may appear that there exist automatic transitions in "RedP reached" and "PP reached". However, they are requested and acknowledged before they take place. The time at which the transition will happen is known by looking at the range header. If the client sends a request close in time to these transitions it needs to be prepared for receiving error messages, as the state may or may not have changed.

Appendix C. Media Transport Alternatives

This section defines how certain combinations of protocols, profiles and lower transports are used. This includes the usage of the Transport header's source and destination address parameters "src_addr" and "dest_addr".

C.1. RTP

This section defines the interaction of RTSP with respect to the RTP protocol [RFC3550]. It also defines any necessary media transport signalling with regards to RTP.

The available RTP profiles and lower layer transports are described below along with rules on signalling the available combinations.

C.1.1. AVP

The usage of the "RTP Profile for Audio and Video Conferences with Minimal Control" [RFC3551] when using RTP for media transport over different lower layer transport protocols is defined below in regards to RTSP.

One such case is defined within this document, the use of embedded (interleaved) binary data as defined in Section 14. The usage of this method is indicated by including the "interleaved" parameter.

When using embedded binary data the "src_addr" and "dest_addr" MUST NOT be used. This addressing and multiplexing is used as defined with use of channel numbers and the interleaved parameter.

C.1.2. AVP/UDP

This part describes sending of RTP [RFC3550] over lower transport layer UDP [RFC0768] according to the profile "RTP Profile for Audio and Video Conferences with Minimal Control" defined in RFC 3551 [RFC3551]. This profile requires one or two uni- or bi-directional UDP flows per media stream. The first UDP flow is for RTP and the second is for RTCP. Embedding of RTP data with the RTSP messages, in accordance with Section 14, SHOULD NOT be performed when RTSP messages are transported over unreliable transport protocols, like UDP [RFC0768].

The RTP/UDP and RTCP/UDP flows can be established using the Transport header's "src_addr", and "dest_addr" parameters.

In RTSP PLAY mode, the transmission of RTP packets from client to server is unspecified. The behavior in regards to such RTP packets

MAY be defined in future.

The "src_addr" and "dest_addr" parameters are used in the following way for media delivery and playback mode, i.e. Mode=PLAY:

- o The "src_addr" and "dest_addr" parameters MUST contain either 1 or 2 address specifications.
- o Each address specification for RTP/AVP/UDP or RTP/AVP/TCP MUST contain either:
 - * both an address and a port number, or
 - * a port number without an address.
- o The first address and port pair given in either of the parameters applies to the RTP stream. The second address and port pair if present applies to the RTCP stream.
- o The RTP/UDP packets from the server to the client MUST be sent to the address and port given by the first address and port pair of the "dest_addr" parameter.
- o The RTCP/UDP packets from the server to the client MUST be sent to the address and port given by the second address and port pair of the "dest_addr" parameter. If no second pair is specified RTCP MUST NOT be sent.
- o The RTCP/UDP packets from the client to the server MUST be sent to the address and port given by the second address and port pair of the "src_addr" parameter. If no second pair is given RTCP MUST NOT be sent.
- o The RTP/UDP packets from the client to the server MUST be sent to the address and port given by the first address and port pair of the "src_addr" parameter.
- o RTP and RTCP Packets SHOULD be sent from the corresponding receiver port, i.e. RTCP packets from the server should be sent from the "src_addr" parameters second address port pair.

C.1.3. AVPF/UDP

The RTP profile "Extended RTP Profile for RTCP-based Feedback (RTP/AVPF)" [RFC4585] MAY be used as RTP profiles in sessions using RTP. All that is defined for AVP MUST also apply for AVPF.

The usage of AVPF is indicated by the media initialization protocol

used. In the case of SDP it is indicated by media lines (m=) containing the profile RTP/AVPF. That SDP MAY also contain further AVPF related SDP attributes configuring the AVPF session regarding reporting interval and feedback messages to be used. This configuration MUST be followed.

C.1.4. SAVP/UDP

The RTP profile "The Secure Real-time Transport Protocol (SRTP)" [RFC3711] is an RTP profile (SAVP) that MAY be used in RTSP sessions using RTP. All that is defined for AVP MUST also apply for SAVP.

The usage of SRTP requires that a security context is established. The default key-management unless otherwise signalled shall be MIKEY in RSA-R mode as defined in Appendix C.1.4.1, and not according to the procedure defined in "Key Management Extensions for Session Description Protocol (SDP) and Real Time Streaming Protocol (RTSP)" [RFC4567]. The reason is that RFC 4567 sends the initial MIKEY message in SDP, thus both requiring the usage of the DESCRIBE method and forcing the server to keep state for clients performing DESCRIBE in anticipation that they might require key management.

MIKEY is selected as default method for establishing SRTP cryptographic context within an RTSP session as it can be embedded in the RTSP messages, while still ensuring confidentiality of content of the keying material, even when using hop-by-hop TLS security for the RTSP messages. This method does also support pipelining of the RTSP messages.

C.1.4.1. MIKEY Key Establishment

This method for using MIKEY to establish the SRTP cryptographic context is initiated in the client's SETUP request, and the servers response to the SETUP carries the MIKEY response. Thus ensuring that the crypto context establishment happens simultaneously with the establishment of the media stream being protected. By using MIKEY's RSA-R mode [RFC4738] the client can be the initiator and still allow the server to set the parameters in accordance with the actual media stream.

The SRTP cryptographic context establishment is done according to the following process:

1. The client determines that SAVP or SAVPF shall be used from media description format, e.g. SDP. If no other key management method is explicitly signalled, then MIKEY SHALL be used as defined herein. This specification does not specify an explicit method for indicating this SRTP cryptographic context

establishment method, but future specifications may.

2. The client SHALL establish a TLS connection for RTSP messages, directly or hop by hop with the server. If hop-by-hop TLS security is used, the User method SHALL be indicated in the Accept-Credentials header. We do note that using hop-by-hop does allow the proxy to insert itself as a man in the middle also in the MIKEY exchange by providing one of its certificates, rather than the server's in the Connection-Credentials header. The client SHALL therefore validate the server certificate.
3. The client retrieves the servers certificate from a direct TLS connection, or if hop by hop from Connection-Credentials header. The client then checks that the server certificate is valid and belongs to the server.
4. The client forms the MIKEY Initiator message using RSA-R mode in unicast mode as specified in [RFC4738]. The client SHOULD use the same certificate for TLS and in MIKEY to enable the server to bind the two together. The client's certificate SHALL be included in the MIKEY message. The client SHALL indicate its SRTP capabilities in the message.
5. The MIKEY message from the previous step is base64 [RFC4648] encoded and becomes the value of the MIKEY parameter that is included in the transport specification(s) that specifies a SRTP based profile (SAVP, SAVPF) in the SETUP request.
6. Any proxy encountering the MIKEY parameter SHALL forward it without modification. A proxy requiring to understand transport specification which doesn't support SAVP/SAVPF with MIKEY will discard the whole transport specification. Most types of proxy can easily support SAVP and SAVPF with MIKEY. If possible bypassing the proxy should be tried.
7. The server upon receiving the SETUP request, will need to decide upon the transport specification to use, if multiple are included by the client. In the determination of which transport specifications that are supported and preferred, the server SHOULD decode the MIKEY message to take the embedded SRTP parameters into account. If all transport specs require SRTP but no MIKEY parameter or other supported keying method is included, the server SHALL respond with 403.
8. Upon generating a response the following outcomes can occur:
 - * A transport spec not using SRTP and MIKEY is selected. Thus the response will not contain any MIKEY parameter.

- * A transport spec using SRTP and MIKEY is selected but an error is encountered in the MIKEY processing. In that case an RTSP error response code of 466 "Key Management Error" SHALL be used. A MIKEY message describing the error MAY be included.
 - * A transport spec using SRTP and MIKEY is selected and a MIKEY response message can be created. The server SHOULD use the same certificate for TLS and in MIKEY to enable client to bind the two together. If a different certificate is used it SHALL be included in the MIKEY message. It is RECOMMENDED that the envelope key cache type is set to 'Cache' and that a single envelope key is reused for all MIKEY messages to the client. That message is included in the MIKEY parameter part of the single selected transport specification in the SETUP response. The server will set the SRTP parameters as preferred for this media stream within the supported range by the client.
9. The server transmits the SETUP response back to the client.
10. The client receives the SETUP response and if the response code indicates a successful request it decodes the MIKEY message and establish the SRTP cryptographic context from the parameters in the MIKEY response.

In the above method the client's certificate may be self-signed in cases where the client's identity is not necessary to establish and the security goal is only to ensure that the RTSP signalling client is the same as the one receiving the SRTP security context.

C.1.5. SAVPF/UDP

The RTP profile "Extended Secure RTP Profile for RTCP-based Feedback (RTP/SAVPF)" [RFC5124] is an RTP profile (SAVPF) that MAY be used in RTSP sessions using RTP. All that is defined for AVP MUST also apply for SAVPF.

The usage of SRTP requires that a cryptographic context is established. The default mechanism for establishing that security association is to use MIKEY[RFC3830] with RTSP as defined in Appendix C.1.4.1.

C.1.6. RTCP usage with RTSP

RTCP has several usages when RTP is used for media transport as explained below. Due to that RTCP MUST be supported if an RTSP agent handles RTP.

C.1.6.1. Media synchronization

RTCP provides media synchronization and clock drift compensation. The initial media synchronization is available from RTP-Info header. However, to be able to handle any clock drift between the media streams, RTCP is needed.

C.1.6.2. RTSP Session keep-alive

RTCP traffic from the RTSP client to the RTSP server MUST function as keep-alive. This requires an RTSP server supporting RTP to use the received RTCP packets as indications that the client desires the related RTSP session to be kept alive.

C.1.6.3. Bit-rate adaption

RTCP Receiver reports and any additional feedback from the client MUST be used to adapt the bit-rate used over the transport for all cases when RTP is sent over UDP. An RTP sender without reserved resources MUST NOT use more than its fair share of the available resources. This can be determined by comparing on short to medium term (some seconds) the used bit-rate and adapt it so that the RTP sender sends at a bit-rate comparable to what a TCP sender would achieve on average over the same path.

C.1.6.4. RTP and RTCP Multiplexing

RTSP can be used to negotiate the usage of RTP and RTCP multiplexing as described in [RFC5761]. This allows servers and client to reduce the amount of resources required for the session by only requiring one underlying transport stream per media stream instead of two when using RTP and RTCP. This lessens the server port consumption and also the necessary state and keep-alive work when operating across Network and Address Translators [RFC2663].

Content must be prepared with some consideration for RTP and RTCP multiplexing, mainly ensuring that the RTP payload types used do not collide with the ones used for RTCP packet types. This option likely needs explicit support from the content unless the RTP payload types can be remapped by the server and that is correctly reflected in the session description. Beyond that support of this feature should come at little cost and much gain.

It is recommended that if the content and server support RTP and RTCP multiplexing that this is indicated in the session description, for example using the SDP attribute "a=rtcp-mux". If the SDP message contains the a=rtcp-mux attribute for a media stream, the server MUST support RTP and RTCP multiplexing. If indicated or otherwise desired

by the client it can include the Transport parameter "RTCP-mux" in any transport specification where it desires to use RTCP-mux. The server will indicate if it supports RTCP-mux. Servers and Clients SHOULD support RTP and RTCP multiplexing.

For capability exchange, an RTSP feature tag for RTP and RTCP multiplexing is defined: "setup.rtp.rtcp.mux".

C.2. RTP over TCP

Transport of RTP over TCP can be done in two ways: over independent TCP connections using RFC 4571 [RFC4571] or interleaved in the RTSP control connection. In both cases the protocol MUST be "rtp" and the lower layer MUST be TCP. The profile may be any of the above specified ones; AVP, AVPF, SAVP or SAVPF.

C.2.1. Interleaved RTP over TCP

The use of embedded (interleaved) binary data transported on the RTSP connection is possible as specified in Section 14. When using this declared combination of interleaved binary data the RTSP messages MUST be transported over TCP. TLS may or may not be used.

One should, however, consider that this will result in all media streams go through any proxy. Using independent TCP connections can avoid that issue.

C.2.2. RTP over independent TCP

In this Appendix, we describe the sending of RTP [RFC3550] over lower transport layer TCP [RFC0793] according to "Framing Real-time Transport Protocol (RTP) and RTP Control Protocol (RTCP) Packets over Connection-Oriented Transport" [RFC4571]. This Appendix adapts the guidelines for using RTP over TCP within SIP/SDP [RFC4145] to work with RTSP.

A client codes the support of RTP over independent TCP by specifying an RTP/AVP/TCP transport option without an interleaved parameter in the Transport line of a SETUP request. This transport option MUST include the "unicast" parameter.

If the client wishes to use RTP with RTCP, two ports (or two address/port pairs) are specified by the dest_addr parameter. If the client wishes to use RTP without RTCP, one port (or one address/port pair) is specified by the dest_addr parameter. Ordering rules of dest_addr ports follow the rules for RTP/AVP/UDP.

If the client wishes to play the active role in initiating the TCP

connection, it MAY set the "setup" parameter (See Section 16.52) on the Transport line to be "active", or it MAY omit the setup parameter, as active is the default. If the client signals the active role, the ports for all dest_addr values MUST be set to 9 (the discard port).

If the client wishes to play the passive role in TCP connection initiation, it MUST set the "setup" parameter on the Transport line to be "passive". If the client is able to assume the active or the passive role, it MUST set the "setup" parameter on the Transport line to be "actpass". In either case, the dest_addr port value for RTP MUST be set to the TCP port number on which the client is expecting to receive the RTP stream connection, and the dest_addr port value for RTCP MUST be set to the TCP port number on which the client is expecting to receive the RTCP stream connection.

If upon receipt of a non-interleaved RTP/AVP/TCP SETUP request, a server decides to accept this requested option, the 2xx reply MUST contain a Transport option that specifies RTP/AVP/TCP (without using the interleaved parameter, and with using the unicast parameter). The dest_addr parameter value MUST be echoed from the parameter value in the client request unless the destination address (only port) was not provided in which case the server MAY include the source address of the RTSP TCP connection with the port number unchanged.

In addition, the server reply MUST set the setup parameter on the Transport line, to indicate the role the server will play in the connection setup. Permissible values are "active" (if a client set "setup" to "passive" or "actpass") and "passive" (if a client set "setup" to "active" or "actpass").

If a server sets "setup" to "passive", the "src_addr" in the reply MUST indicate the ports the server is willing to receive an RTP connection and (if the client requested an RTCP connection by specifying two dest_addr ports or address/port pairs) and RTCP connection. If a server sets "setup" to "active", the ports specified in "src_addr" MUST be set to 9. The server MAY use the "ssrc" parameter, following the guidance in Section 16.52. Port ordering for src_addr follows the rules for RTP/AVP/UDP.

Servers MUST support taking the passive role and MAY support taking the active role. Servers with a public IP address take the passive role, thus enabling clients behind NATs and Firewalls a better chance of successful connect to the server by actively connecting outwards. Therefore the clients are RECOMMENDED to take the active role.

After sending (receiving) a 2xx reply for a SETUP method for a non-interleaved RTP/AVP/TCP media stream, the active party SHOULD

initiate the TCP connection as soon as possible. The client **MUST NOT** send a **PLAY** request prior to the establishment of all the TCP connections negotiated using **SETUP** for the session. In case the server receives a **PLAY** request in a session that has not yet established all the TCP connections, it **MUST** respond using the 464 "Data Transport Not Ready Yet" (Section 15.4.29) error code.

Once the **PLAY** request for a media resource transported over non-interleaved RTP/AVP/TCP occurs, media begins to flow from server to client over the RTP TCP connection, and RTCP packets flow bidirectionally over the RTCP TCP connection. As in the RTP/UDP case, client to server traffic on the TCP port is unspecified by this memo. The packets that travel on these connections **MUST** be framed using the protocol defined in [RFC4571], not by the framing defined for interleaving RTP over the RTSP control connection defined in Section 14.

A successful **PAUSE** request for a media being transported over RTP/AVP/TCP pauses the flow of packets over the connections, without closing the connections. A successful **TEARDOWN** request signals that the TCP connections for RTP and RTCP are to be closed as soon as possible.

Subsequent **SETUP** requests on an already-**SETUP** RTP/AVP/TCP URI may be ambiguous in the following way: does the client wish to open up new TCP RTP and RTCP connections for the URI, or does the client wish to continue using the existing TCP RTP and RTCP connections? The client **SHOULD** use the "connection" parameter (defined in Section 16.52) on the Transport line to make its intention clear (by setting "connection" to "new" if new connections are needed, and by setting "connection" to "existing" if the existing connections are to be used). After a 2xx reply for a **SETUP** request for a new connection, parties should close the pre-existing connections, after waiting a suitable period for any stray RTP or RTCP packets to arrive.

The usage of SRTP, i.e. either SAVP or SAVPF profiles requires that a security association is established. The default mechanism for establishing that security association is to use MIKEY[RFC3830] with RTSP as defined Appendix C.1.4.1.

Below, we rewrite part of the example media on demand example shown in Appendix A.1 to use RTP/AVP/TCP non-interleaved:

C->M: DESCRIBE rtsp://example.com/twister.3gp RTSP/2.0
CSeq: 1
User-Agent: PhonyClient/1.2

M->C: RTSP/2.0 200 OK
CSeq: 1
Server: PhonyServer/1.0
Date: Thu, 23 Jan 1997 15:35:06 GMT
Content-Type: application/sdp
Content-Length: 227
Content-Base: rtsp://example.com/twister.3gp/
Expires: 24 Jan 1997 15:35:06 GMT

v=0
o=- 2890844256 2890842807 IN IP4 198.51.100.34
s=RTSP Session
i=An Example of RTSP Session Usage
e=adm@example.com
c=IN IP4 0.0.0.0
a=control: *
a=range: npt=0-0:10:34.10
t=0 0
m=audio 0 RTP/AVP 0
a=control: trackID=1

C->M: SETUP rtsp://example.com/twister.3gp/trackID=1 RTSP/2.0
CSeq: 2
User-Agent: PhonyClient/1.2
Require: play.basic
Transport: RTP/AVP/TCP;unicast;dest_addr=":9"/":9";
 setup=active;connection=new
Accept-Ranges: NPT, SMPTE, UTC

```
M->C: RTSP/2.0 200 OK
      CSeq: 2
      Server: PhonyServer/1.0
      Transport: RTP/AVP/TCP;unicast;
                 dest_addr=":9"/":9";
                 src_addr="198.51.100.5:53478"/"198.51.100:54091";
                 setup=passive;connection=new;ssrc=93CB001E
      Session: 12345678
      Expires: 24 Jan 1997 15:35:12 GMT
      Date: 23 Jan 1997 15:35:12 GMT
      Accept-Ranges: NPT
      Media-Properties: Random-Access=0.8, Immutable, Unlimited
```

```
C->M: TCP Connection Establishment x2
```

```
C->M: PLAY rtsp://example.com/twister.3gp/ RTSP/2.0
      CSeq: 4
      User-Agent: PhonyClient/1.2
      Range: npt=30-
      Session: 12345678
```

```
M->C: RTSP/2.0 200 OK
      CSeq: 4
      Server: PhonyServer/1.0
      Date: 23 Jan 1997 15:35:14 GMT
      Session: 12345678
      Range: npt=30-623.10
      Seek-Style: First-Prior
      RTP-Info: url="rtsp://example.com/twister.3gp/trackID=1"
                 ssrc=4F312DD8;seq=54321;rtptime=2876889
```

C.3. Handling Media Clock Time Jumps in the RTP Media Layer

RTSP allows media clients to control selected, non-contiguous sections of media presentations, rendering those streams with an RTP media layer [RFC3550]. Two cases occur, the first is when a new PLAY request replaces an old ongoing request and the new request results in a jump in the media. This should produce in the RTP layer a continuous media stream. A client may also directly following a completed PLAY request perform a new PLAY request. This will result in some gap in the media layer. The below text will look into both cases.

A PLAY request that replaces an ongoing request allows the media layer rendering the RTP stream without being affected by jumps in media clock time. The RTP timestamps for the new media range is set so that they become continuous with the previous media range in the previous request. The RTP sequence number for the first packet in

the new range will be the next following the last packet in the previous range, i.e. monotonically increasing. The goal is to allow the media rendering layer to work without interruption or reconfiguration across the jumps in media clock. This should be possible in all cases of replaced PLAY requests for media that has random-access properties. In this case care is needed to align frames or similar media dependent structures.

In cases where jumps in media clock time are a result of RTSP signalling operations arriving after a completed PLAY operation, the request timing will result in that media becomes non-continuous. The server becomes unable to send the media so that it arrives timely and still carry timestamps to make the media stream continuous. In these cases the server will produce RTP streams where there are gaps in the RTP timeline for the media. In such cases, if the media has frame structure, aligning the timestamp for the next frame with the previous structure reduces the burden to render this media. The gap should represent the time the server hasn't been serving media, e.g. the time between the end of the media stream or a PAUSE request and the new PLAY request. In these cases the RTP sequence number would normally be monotonically increasing across the gap.

For RTSP sessions with media that lacks random access properties, such as live streams, any media clock jump is commonly the result of a correspondingly long pause of delivery. The RTP timestamp will have increased in direct proportion to the duration of the paused delivery. Note also that in this case the RTP sequence number should be the next packet number. If not, the RTCP packet loss reporting will indicate as loss all packets not received between the point of pausing and later resuming. This may trigger congestion avoidance mechanisms. An allowed exception from the above recommendation on monotonically increasing RTP sequence number is live media streams, likely being relayed. In this case, when the client resumes delivery, it will get the media that is currently being delivered to the server itself. For this type of basic delivery of live streams to multiple users over unicast, individual rewriting of RTP sequence numbers becomes quite a burden. For solutions that anyway caches media, timeshifts, etc, the rewriting should be a minor issue.

The goal when handling jumps in media clock time is that the provided stream is continuous without gaps in RTP timestamp or sequence number. However, when delivery has been halted for some reason the RTP timestamp when resuming MUST represent the duration the delivery was halted. RTP sequence number MUST generally be the next number, i.e. monotonically increasing modulo 65536. For media resources with the properties Time-Progressing and Time-Duration=0.0 the server MAY create RTP media streams with RTP sequence number jumps in them due to the client first halting delivery and later resuming it (PAUSE and

then later PLAY). However, servers utilizing this exception must take into consideration the resulting RTCP receiver reports that likely contains loss reports for all the packets part of the discontinuity. A client cannot rely on that a server will align when resuming playing even if it is RECOMMENDED. The RTP-Info header will provide information on how the server acts in each case.

We cannot assume that the RTSP client can communicate with the RTP media agent, as the two may be independent processes. If the RTP timestamp shows the same gap as the NPT, the media agent will assume that there is a pause in the presentation. If the jump in NPT is large enough, the RTP timestamp may roll over and the media agent may believe later packets to be duplicates of packets just played out. Having the RTP timestamp jump will also affect the RTCP measurements based on this.

As an example, assume an RTP timestamp frequency of 8000 Hz, a packetization interval of 100 ms and an initial sequence number and timestamp of zero.

```
C->S: PLAY rtsp://example.com/fizzle RTSP/2.0
      CSeq: 4
      Session: abcdefgh
      Range: npt=10-15
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 4
      Session: abcdefgh
      Range: npt=10-15
      RTP-Info: url="rtsp://example.com/fizzle/audiotrack"
                ssrc=0D12F123:seq=0;rtptime=0
```

The ensuing RTP data stream is depicted below:

```
S -> C: RTP packet - seq = 0,   rtptime = 0,       NPT time = 10s
S -> C: RTP packet - seq = 1,   rtptime = 800,     NPT time = 10.1s
. . .
S -> C: RTP packet - seq = 49,  rtptime = 39200,   NPT time = 14.9s
```

Upon the completion of the requested delivery the server sends a PLAY_NOTIFY

```
S->C: PLAY_NOTIFY rtsp://example.com/fizzle RTSP/2.0
      CSeq: 5
      Notify-Reason: end-of-stream
      Request-Status: cseq=4 status=200 reason="OK"
      Range: npt=-15
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
               ssrc=0D12F123:seq=49;rtptime=39200
      Session: abcdefgh

C->S: RTSP/2.0 200 OK
      CSeq: 5
      User-Agent: PhonyClient/1.2
```

Upon the completion of the play range, the client follows up with a request to PLAY from a new NPT.

```
C->S: PLAY rtsp://example.com/fizzle RTSP/2.0
      CSeq: 6
      Session: abcdefg
      Range: npt=18-20
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 6
      Session: abcdefg
      Range: npt=18-20
      RTP-Info: url="rtsp://example.com/fizzle/audiotrack"
               ssrc=0D12F123:seq=50;rtptime=40100
```

The ensuing RTP data stream is depicted below:

```
S->C: RTP packet - seq = 50, rtptime = 40100, NPT time = 18s
S->C: RTP packet - seq = 51, rtptime = 40900, NPT time = 18.1s
. . .
S->C: RTP packet - seq = 69, rtptime = 55300, NPT time = 19.9s
```

In this example, first, NPT 10 through 15 is played, then the client requests the server to skip ahead and play NPT 18 through 20. The first segment is presented as RTP packets with sequence numbers 0 through 49 and timestamp 0 through 39,200. The second segment consists of RTP packets with sequence number 50 through 69, with timestamps 40,100 through 55,200. While there is a gap in the NPT, there is no gap in the sequence number space of the RTP data stream.

The RTP timestamp gap is present in the above example due to the time it takes to perform the second play request, in this case 12.5 ms (100/8000).

C.4. Handling RTP Timestamps after PAUSE

During a PAUSE / PLAY interaction in an RTSP session, the duration of time for which the RTP transmission was halted MUST be reflected in the RTP timestamp of each RTP stream. The duration can be calculated for each RTP stream as the time elapsed from when the last RTP packet was sent before the PAUSE request was received and when the first RTP packet was sent after the subsequent PLAY request was received. The duration includes all latency incurred and processing time required to complete the request.

The RTP RFC [RFC3550] states that: The RTP timestamp for each unit [packet] would be related to the wallclock time at which the unit becomes current on the virtual presentation timeline.

In order to satisfy the requirements of [RFC3550], the RTP timestamp space needs to increase continuously with real time. While this is not optimal for stored media, it is required for RTP and RTCP to function as intended. Using a continuous RTP timestamp space allows the same timestamp model for both stored and live media and allows better opportunity to integrate both types of media under a single control.

As an example, assume a clock frequency of 8000 Hz, a packetization interval of 100 ms and an initial sequence number and timestamp of zero.

```
C->S: PLAY rtsp://example.com/fizzle RTSP/2.0
      CSeq: 4
      Session: abcdefg
      Range: npt=10-15
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 4
      Session: abcdefg
      Range: npt=10-15
      RTP-Info: url="rtsp://example.com/fizzle/audiotrack"
                ssrc=0D12F123:seq=0;rtptime=0
```

The ensuing RTP data stream is depicted below:

```
S -> C: RTP packet - seq = 0, rtptime = 0,    NPT time = 10s
S -> C: RTP packet - seq = 1, rtptime = 800,  NPT time = 10.1s
S -> C: RTP packet - seq = 2, rtptime = 1600, NPT time = 10.2s
S -> C: RTP packet - seq = 3, rtptime = 2400, NPT time = 10.3s
```

The client then sends a PAUSE request:

```
C->S: PAUSE rtsp://example.com/fizzle RTSP/2.0
      CSeq: 5
      Session: abcdefg
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 5
      Session: abcdefg
      Range: npt=10.4-15
```

20 seconds elapse and then the client sends a PLAY request. In addition the server requires 15 ms to process the request:

```
C->S: PLAY rtsp://example.com/fizzle RTSP/2.0
      CSeq: 6
      Session: abcdefg
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 6
      Session: abcdefg
      Range: npt=10.4-15
      RTP-Info: url="rtsp://example.com/fizzle/audiotrack"
                 ssrc=0D12F123:seq=4;rtptime=164400
```

The ensuing RTP data stream is depicted below:

```
S -> C: RTP packet - seq = 4, rtptime = 164400, NPT time = 10.4s
S -> C: RTP packet - seq = 5, rtptime = 165200, NPT time = 10.5s
S -> C: RTP packet - seq = 6, rtptime = 166000, NPT time = 10.6s
```

First, NPT 10 through 10.3 is played, then a PAUSE is received by the server. After 20 seconds a PLAY is received by the server which takes 15ms to process. The duration of time for which the session was paused is reflected in the RTP timestamp of the RTP packets sent after this PLAY request.

A client can use the RTSP range header and RTP-Info header to map NPT time of a presentation with the RTP timestamp.

Note: In RFC 2326 [RFC2326], this matter was not clearly defined and was misunderstood commonly. However, for RTSP 2.0 it is expected that this will be handled correctly and no exception handling will be required.

Note Further: To ensure correct media decoding and usually jitter-buffer handling resetting some of the state when issuing a PLAY request is needed.

C.5. RTSP / RTP Integration

For certain datatypes, tight integration between the RTSP layer and the RTP layer will be necessary. This by no means precludes the above restrictions. Combined RTSP/RTP media clients should use the RTP-Info field to determine whether incoming RTP packets were sent before or after a seek or before or after a PAUSE.

C.6. Scaling with RTP

For scaling (see Section 16.44), RTP timestamps should correspond to the rendering timing. For example, when playing video recorded at 30 frames/second at a scale of two and speed (Section 16.48) of one, the server would drop every second frame to maintain and deliver video packets with the normal timestamp spacing of 3,000 per frame, but NPT would increase by 1/15 second for each video frame.

Note: The above scaling puts requirements on the media codec or a media stream to support it. For example motion JPEG or other non-predictive video coding can easier handle the above example.

C.7. Maintaining NPT synchronization with RTP timestamps

The client can maintain a correct display of NPT (Normal Play Time) by noting the RTP timestamp value of the first packet arriving after repositioning. The sequence parameter of the RTP-Info (Section 16.43) header provides the first sequence number of the next segment.

C.8. Continuous Audio

For continuous audio, the server SHOULD set the RTP marker bit at the beginning of serving a new PLAY request or at jumps in timeline. This allows the client to perform playout delay adaptation.

C.9. Multiple Sources in an RTP Session

Note that more than one SSRC MAY be sent in the media stream. If it happens all sources are expected to be rendered simultaneously.

C.10. Usage of SSRCS and the RTCP BYE Message During an RTSP Session

The RTCP BYE message indicates the end of use of a given SSRC. If all sources leave an RTP session, it can, in most cases, be assumed to have ended. Therefore, a client or server MUST NOT send an RTCP BYE message until it has finished using a SSRC. A server SHOULD keep using a SSRC until the RTP session is terminated. Prolonging the use of a SSRC allows the established synchronization context associated

with that SSRC to be used to synchronize subsequent PLAY requests even if the PLAY response is late.

An SSRC collision with the SSRC that transmits media does also have consequences, as it will normally force the media sender to change its SSRC in accordance with the RTP specification[RFC3550]. However, an RTSP server may wait and see if the client changes and thus resolve the conflict to minimize the impact. As media sender SSRC change will result in a loss of synchronization context, and require any receiver to wait for RTCP sender reports for all media requiring synchronization before being able to play out synchronized. Due to these reasons a client joining a session should take care to not select the same SSRC(s) as the server indicates in the ssrc Transport header parameter. Any SSRC signalled in the Transport header MUST be avoided. A client detecting a collision prior to sending any RTP or RTCP messages SHALL also select a new SSRC.

C.11. Future Additions

It is the intention that any future protocol or profile regarding media delivery and lower transport should be easy to add to RTSP. This section provides the necessary steps that needs to be meet.

The following things needs to be considered when adding a new protocol or profile for use with RTSP:

- o The protocol or profile needs to define a name tag representing it. This tag is required to be an ABNF "token" to be possible to use in the Transport header specification.
- o The useful combinations of protocol, profiles and lower layer transport for this extension needs to be defined. For each combination declare the necessary parameters to use in the Transport header.
- o For new media protocols the interaction with RTSP needs to be addressed. One important factor will be the media synchronization. May need new headers similar to RTP info to carry information.
- o Discuss congestion control for media, especially if transport without built in congestion control is used.

See the IANA section (Section 22) for information how to register new attributes.

Appendix D. Use of SDP for RTSP Session Descriptions

The Session Description Protocol (SDP, [RFC4566]) may be used to describe streams or presentations in RTSP. This description is typically returned in reply to a DESCRIBE request on an URI from a server to a client, or received via HTTP from a server to a client.

This appendix describes how an SDP file determines the operation of an RTSP session. SDP as is provides no mechanism by which a client can distinguish, without human guidance, between several media streams to be rendered simultaneously and a set of alternatives (e.g., two audio streams spoken in different languages). The SDP extension "Grouping of Media Lines in the Session Description Protocol (SDP)" [RFC5888] provides such functionality to some degree. Appendix D.4 describes the usage of SDP media line grouping for RTSP.

D.1. Definitions

The terms "session-level", "media-level" and other key/attribute names and values used in this appendix are to be used as defined in SDP[RFC4566]:

D.1.1. Control URI

The "a=control:" attribute is used to convey the control URI. This attribute is used both for the session and media descriptions. If used for individual media, it indicates the URI to be used for controlling that particular media stream. If found at the session level, the attribute indicates the URI for aggregate control (presentation URI). The session level URI MUST be different from any media level URI. The presence of a session level control attribute MUST be interpreted as support for aggregated control. The control attribute MUST be present on media level unless the presentation only contains a single media stream, in which case the attribute MAY only be present on the session level and then also apply to that single media level.

ABNF for the attribute is defined in Section 20.3.

Example:

```
a=control:rtsp://example.com/foo
```

This attribute MAY contain either relative or absolute URIs, following the rules and conventions set out in RFC 3986 [RFC3986]. Implementations MUST look for a base URI in the following order:

1. the RTSP Content-Base field;

2. the RTSP Content-Location field;
3. the RTSP Request-URI.

If this attribute contains only an asterisk (*), then the URI MUST be treated as if it were an empty embedded URI, and thus inherit the entire base URI.

Note, RFC 2326 was very unclear on the processing of relative URI and several RTSP 1.0 implementations at the point of publishing this document did not perform RFC 3986 processing to determine the resulting URI, instead simple concatenation is common. To avoid this issue completely it is recommended to use absolute URI in the SDP.

The URI handling for SDPs from container files need special consideration. For example lets assume that a container file has the URI: "rtsp://example.com/container.mp4". Lets further assume this URI is the base URI, and that there is an absolute media level URI: "rtsp://example.com/container.mp4/trackID=2". A relative media level URI that resolves in accordance with RFC 3986 [RFC3986] to the above given media URI is: "container.mp4/trackID=2". It is usually not desirable to need to include in or modify the SDP stored within the container file with the server local name of the container file. To avoid this, one can modify the base URI used to include a trailing slash, e.g. "rtsp://example.com/container.mp4/". In this case the relative URI for the media will only need to be: "trackID=2". However, this will also mean that using "*" in the SDP will result in control URI including the trailing slash, i.e. "rtsp://example.com/container.mp4/".

Note: The usage of TrackID in the above is not a standardized form, but one example out of several similar strings such as TrackID, Track_ID, StreamID that is used by different server vendors to indicate a particular piece of media inside a container file.

D.1.2. Media Streams

The "m=" field is used to enumerate the streams. It is expected that all the specified streams will be rendered with appropriate synchronization. If the session is over multicast, the port number indicated SHOULD be used for reception. The client MAY try to override the destination port, through the Transport header. The servers MAY allow this, the response will indicate if allowed or not. If the session is unicast, the port numbers are the ones RECOMMENDED by the server to the client, about which receiver ports to use; the client MUST still include its receiver ports in its SETUP request.

The client MAY ignore this recommendation. If the server has no preference, it SHOULD set the port number value to zero.

The "m=" lines contain information about which transport protocol, profile, and possibly lower-layer is to be used for the media stream. The combination of transport, profile and lower layer, like RTP/AVP/UDP needs to be defined for how to be used with RTSP. The currently defined combinations are defined in Appendix C, further combinations MAY be specified.

Example:

m=audio 0 RTP/AVP 31

D.1.3. Payload Type(s)

The payload type(s) are specified in the "m=" line. In case the payload type is a static payload type from RFC 3551 [RFC3551], no other information may be required. In case it is a dynamic payload type, the media attribute "rtpmap" is used to specify what the media is. The "encoding name" within the "rtpmap" attribute may be one of those specified in RFC 3551 (Sections 5 and 6), or media type registered with IANA [RFC4288], or an experimental encoding as specified in SDP (RFC 4566 [RFC4566]). Codec-specific parameters are not specified in this field, but rather in the "fmtp" attribute described below.

The selection of the RTP payload type numbers used may be required to consider RTP and RTCP Multiplexing [RFC5761] if that is to be supported by the server.

D.1.4. Format-Specific Parameters

Format-specific parameters are conveyed using the "fmtp" media attribute. The syntax of the "fmtp" attribute is specific to the encoding(s) that the attribute refers to. Note that some of the format specific parameters may be specified outside of the fmtp parameters, like for example the "ptime" attribute for most audio encodings.

D.1.5. Directionality of media stream

The SDP attributes "a=sendrecv", "a=recvonly" and "a=sendonly" provide instructions about the direction the media streams flow within a session. When using RTSP the SDP can be delivered to a client using either RTSP DESCRIBE or a number of RTSP external methods, like HTTP, FTP, and email. Based on this the SDP applies to how the RTSP client will see the complete session. Thus media streams delivered from the RTSP server to the client, would be given

the "a=recvonly" attribute.

The direction attributes are not commonly used in SDPs for RTSP, but may occur. "a=recvonly" in a SDP provided to the RTSP client MUST indicate that media delivery will only occur in the direction from the RTSP server to the client. In SDP provided to the RTSP client that lacks any of the directionality attributes (a=recvonly, a=sendonly, a=sendrecv) MUST behave as if the "a=recvonly" attribute was received. Note that this overrides the normal default rule defined in SDP[RFC4566]. The usage of "a=sendonly" or "a=sendrecv" is not defined, nor is the interpretation of SDP by other entities than the RTSP client.

D.1.6. Range of Presentation

The "a=range" attribute defines the total time range of the stored session or an individual media. Non-seekable live sessions can be indicated as specified below, while the length of live sessions can be deduced from the "t" and "r" SDP parameters.

The attribute is both a session and a media level attribute. For presentations that contain media streams of the same durations, the range attribute SHOULD only be used at session-level. In case of different length the range attribute MUST be given at media level for all media, and SHOULD NOT be given at session level. If the attribute is present at both media level and session level the media level values MUST be used.

Note: Usually one will specify the same length for all media, even if there isn't media available for the full duration on all media. However, that requires that the server accepts PLAY requests within that range.

Servers MUST take care to provide RTSP Range (see Section 16.38) values that are consistent with what is presented in the SDP for the content. There is no reason for non dynamic content, like media clips provided on demand to have inconsistent values. Inconsistent values between the SDP and the actual values for the content handled by the server is likely to generate some failure, like 457 "Invalid Range", in case the client uses PLAY requests with a Range header. In case the content is dynamic in length and it is infeasible to provide a correct value in the SDP the server is recommended to describe this as non-seekable content (see below). The server MAY override that property in the response to a PLAY request using the correct values in the Range header.

The unit is specified first, followed by the value range. The units and their values are as defined in Section 4.4, Section 4.5 and

Section 4.6 and MAY be extended with further formats. Any open ended range (start-), i.e. without stop range, is of unspecified duration and MUST be considered as non-seekable content unless this property is overridden. Multiple instances carrying different clock formats MAY be included at either session or media level.

ABNF for the attribute is defined in Section 20.3.

Examples:

```
a=range:npt=0-34.4368
a=range:clock=19971113T211503Z-19971113T220300Z
Non seekable stream of unknown duration:
a=range:npt=0-
```

D.1.7. Time of Availability

The "t=" field defines when the SDP is valid. For on-demand content the server SHOULD indicate a stop time value for which it guarantees the description to be valid, and a start time that is equal to or before the time at which the DESCRIBE request was received. It MAY also indicate start and stop times of 0, meaning that the session is always available.

For sessions that are of live type, i.e. specific start time, unknown stop time, likely unseekable, the "t=" and "r=" field SHOULD be used to indicate the start time of the event. The stop time SHOULD be given so that the live event will have ended at that time, while still not be unnecessary long into the future.

D.1.8. Connection Information

In SDP, the "c=" field contains the destination address for the media stream. If a multicast address is specified the client SHOULD use this address in any SETUP request as destination address, including any additional parameters, such as TTL. For on-demand unicast streams and some multicast streams, the destination address MAY be specified by the client via the SETUP request, thus overriding any specified address. To identify streams without a fixed destination address, where the client is required to specify a destination address, the "c=" field SHOULD be set to a null value. For addresses of type "IP4", this value MUST be "0.0.0.0", and for type "IP6", this value MUST be "0:0:0:0:0:0:0:0" (can also be written as ":::"), i.e. the unspecified address according to RFC 4291 [RFC4291].

D.1.9. Message Body Tag

The optional "a=mtag" attribute identifies a version of the session description. It is opaque to the client. SETUP requests may include

this identifier in the If-Match field (see Section 16.23) to only allow session establishment if this attribute value still corresponds to that of the current description. The attribute value is opaque and may contain any character allowed within SDP attribute values.

ABNF for the attribute is defined in Section 20.3.

Example:

```
a=mtag:"158bb3e7c7fd62ce67f12b533f06b83a"
```

One could argue that the "o=" field provides identical functionality. However, it does so in a manner that would put constraints on servers that need to support multiple session description types other than SDP for the same piece of media content.

D.2. Aggregate Control Not Available

If a presentation does not support aggregate control no session level "a=control:" attribute is specified. For a SDP with multiple media sections specified, each section will have its own control URI specified via the "a=control:" attribute.

Example:

```
v=0
o=- 2890844256 2890842807 IN IP4 192.0.2.56
s=I came from a web page
e=adm@example.com
c=IN IP4 0.0.0.0
t=0 0
m=video 8002 RTP/AVP 31
a=control:rtsp://audio.example.com/movie.aud
m=audio 8004 RTP/AVP 3
a=control:rtsp://video.example.com/movie.vid
```

Note that the position of the control URI in the description implies that the client establishes separate RTSP control sessions to the servers audio.example.com and video.example.com.

It is recommended that an SDP file contains the complete media initialization information even if it is delivered to the media client through non-RTSP means. This is necessary as there is no mechanism to indicate that the client should request more detailed media stream information via DESCRIBE.

D.3. Aggregate Control Available

In this scenario, the server has multiple streams that can be controlled as a whole. In this case, there are both a media-level "a=control:" attributes, which are used to specify the stream URIs, and a session-level "a=control:" attribute which is used as the Request-URI for aggregate control. If the media-level URI is relative, it is resolved to absolute URIs according to Appendix D.1.1 above.

Example:

```
C->M: DESCRIBE rtsp://example.com/movie RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2
```

```
M->C: RTSP/2.0 200 OK
      CSeq: 1
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Expires: Thu, 23 Jan 1997 16:35:06 GMT
      Content-Type: application/sdp
      Content-Base: rtsp://example.com/movie/
      Content-Length: 227

      v=0
      o=- 2890844256 2890842807 IN IP4 192.0.2.211
      s=I contain
      i=<more info>
      e=adm@example.com
      c=IN IP4 0.0.0.0
      a=control:*
      t=0 0
      m=video 8002 RTP/AVP 31
      a=control:trackID=1
      m=audio 8004 RTP/AVP 3
      a=control:trackID=2
```

In this example, the client is recommended to establish a single RTSP session to the server, and uses the URIs `rtsp://example.com/movie/trackID=1` and `rtsp://example.com/movie/trackID=2` to set up the video and audio streams, respectively. The URI `rtsp://example.com/movie/`, which is resolved from the "*", controls the whole presentation (movie).

A client is not required to issue SETUP requests for all streams within an aggregate object. Servers should allow the client to ask for only a subset of the streams.

D.4. Grouping of Media Lines in SDP

For some types of media it is desirable to express a relationship between various media components, for instance, for lip synchronization or Scalable Video Codec (SVC) [RFC5583]. This relationship is expressed on the SDP level by grouping of media lines, as described in [RFC5888] and can be exposed to RTSP.

For RTSP it is mainly important to know how to handle grouped medias received by means of SDP, i.e., if the media are under aggregate control (see Appendix D.3) or if aggregate control is not available (see Appendix D.2).

It is RECOMMENDED that grouped medias are handled by aggregate control, to give the client the ability to control either the whole presentation or single medias.

D.5. RTSP external SDP delivery

There are some considerations that need to be made when the session description is delivered to the client outside of RTSP, for example via HTTP or email.

First of all, the SDP needs to contain absolute URIs, since relative will in most cases not work as the delivery will not correctly forward the base URI.

The writing of the SDP session availability information, i.e. "t=" and "r=", needs to be carefully considered. When the SDP is fetched by the DESCRIBE method, the probability that it is valid is very high. However, the same is much less certain for SDPs distributed using other methods. Therefore the publisher of the SDP should take care to follow the recommendations about availability in the SDP specification [RFC4566].

Appendix E. RTSP Use Cases

This Appendix describes the most important and considered use cases for RTSP. They are listed in descending order of importance in regards to ensuring that all necessary functionality is present. This specification only fully supports usage of the two first. Also in these first two cases, there are special cases or exceptions that are not supported without extensions, e.g. the redirection of media delivery to another address than the controlling agent's (client's).

E.1. On-demand Playback of Stored Content

An RTSP capable server stores content suitable for being streamed to a client. A client desiring playback of any of the stored content uses RTSP to set up the media transport required to deliver the desired content. RTSP is then used to initiate, halt and manipulate the actual transmission (playout) of the content. RTSP is also required to provide necessary description and synchronization information for the content.

The above high level description can be broken down into a number of functions that RTSP needs to be capable of.

Presentation Description: Provide initialization information about the presentation (content); for example, which media codecs are needed for the content. Other information that is important includes the number of media streams the presentation contains, the transport protocols used for the media streams, and identifiers for these media streams. This information is required before setup of the content is possible and to determine if the client is even capable of using the content.

This information need not be sent using RTSP; other external protocols can be used to transmit the transport presentation descriptions. Two good examples are the use of HTTP [RFC2616] or email to fetch or receive presentation descriptions like SDP [RFC4566]

Setup: Set up some or all of the media streams in a presentation. The setup itself consists of selecting the protocol for media transport and the necessary parameters for the protocol, like addresses and ports.

Control of Transmission: After the necessary media streams have been established the client can request the server to start transmitting the content. The client must be allowed to start or stop the transmission of the content at arbitrary times. The client must also be able to start the transmission at any

point in the timeline of the presentation.

Synchronization: For media transport protocols like RTP [RFC3550] it might be beneficial to carry synchronization information within RTSP. This may be due to either the lack of inter-media synchronization within the protocol itself, or the potential delay before the synchronization is established (which is the case for RTP when using RTCP).

Termination: Terminate the established contexts.

For this use case there are a number of assumptions about how it works. These are:

On-Demand content: The content is stored at the server and can be accessed at any time during a time period when it is intended to be available.

Independent sessions: A server is capable of serving a number of clients simultaneously, including from the same piece of content at different points in that presentations time-line.

Unicast Transport: Content for each individual client is transmitted to them using unicast traffic.

It is also possible to redirect the media traffic to a different destination than that of the agent controlling the traffic. However, allowing this without appropriate mechanisms for checking that the destination approves of this allows for distributed denial of service attacks (DDoS).

E.2. Unicast Distribution of Live Content

This use case is similar to the above on-demand content case (see Appendix E.1) the difference is the nature of the content itself. Live content is continuously distributed as it becomes available from a source; i.e., the main difference from on-demand is that one starts distributing content before the end of it has become available to the server.

In many cases the consumer of live content is only interested in consuming what actually happens "now"; i.e., very similar to broadcast TV. However, in this case it is assumed that there exist no broadcast or multicast channel to the users, and instead the server functions as a distribution node, sending the same content to multiple receivers, using unicast traffic between server and client. This unicast traffic and the transport parameters are individually negotiated for each receiving client.

Another aspect of live content is that it often has a very limited time of availability, as it is only available for the duration of the event the content covers. An example of such a live content could be a music concert which lasts 2 hour and starts at a predetermined time. Thus there is a need to announce when and for how long the live content is available.

In some cases, the server providing live content may be saving some or all of the content to allow clients to pause the stream and resume it from the paused point, or to "rewind" and play continuously from a point earlier than the live point. Hence, this use case does not necessarily exclude playing from other than the live point of the stream, playing with scales other than 1.0, etc.

E.3. On-demand Playback using Multicast

It is possible to use RTSP to request that media be delivered to a multicast group. The entity setting up the session (the controller) will then control when and what media is delivered to the group. This use case has some potential for denial of service attacks by flooding a multicast group. Therefore, a mechanism is needed to indicate that the group actually accepts the traffic from the RTSP server.

An open issue in this use case is how one ensures that all receivers listening to the multicast or broadcast receives the session presentation configuring the receivers. This specification has to rely on an external solution to solve this issue.

E.4. Inviting an RTSP server into a conference

If one has an established conference or group session, it is possible to have an RTSP server distribute media to the whole group. Transmission to the group is simplest when controlled by a single participant or leader of the conference. Shared control might be possible, but would require further investigation and possibly extensions.

This use case assumes that there exists either multicast or a conference focus that redistribute media to all participants.

This use case is intended to be able to handle the following scenario: A conference leader or participant (hereafter called the controller) has some pre-stored content on an RTSP server that he wants to share with the group. The controller sets up an RTSP session at the streaming server for this content and retrieves the session description for the content. The destination for the media content is set to the shared multicast group or conference focus.

When desired by the controller, he/she can start and stop the transmission of the media to the conference group.

There are several issues with this use case that are not solved by this core specification for RTSP:

Denial of service: To avoid an RTSP server from being an unknowing participant in a denial of service attack the server needs to be able to verify the destination's acceptance of the media. Such a mechanism to verify the approval of received media does not yet exist; instead, only policies can be used, which can be made to work in controlled environments.

Distributing the presentation description to all participants in the group: To enable a media receiver to correctly decode the content the media configuration information needs to be distributed reliably to all participants. This will most likely require support from an external protocol.

Passing control of the session: If it is desired to pass control of the RTSP session between the participants, some support will be required by an external protocol to exchange state information and possibly floor control of who is controlling the RTSP session.

E.5. Live Content using Multicast

This use case in its simplest form does not require any use of RTSP at all; this is what multicast conferences being announced with SAP [RFC2974] and SDP are intended to handle. However, in use cases where more advanced features like access control to the multicast session are desired, RTSP could be used for session establishment.

A client desiring to join a live multicasted media session with cryptographic (encryption) access control could use RTSP in the following way. The source of the session announces the session and gives all interested an RTSP URI. The client connects to the server and requests the presentation description, allowing configuration for reception of the media. In this step it is possible for the client to use secured transport and any desired level of authentication; for example, for billing or access control. An RTSP link also allows for load balancing between multiple servers.

If these were the only goals, they could be achieved by simply using HTTP. However, for cases where the sender likes to keep track of each individual receiver of a session, and possibly use the session as a side channel for distributing key-updates or other information on a per-receiver basis, and the full set of receivers is not known

prior to the session start, the state establishment that RTSP provides can be beneficial. In this case a client would establish an RTSP session for this multicast group with the RTSP server. The RTSP server will not transmit any media, but instead will point to the multicast group. The client and server will be able to keep the session alive for as long as the receiver participates in the session thus enabling, for example, the server to push updates to the client.

This use case will most likely not be able to be implemented without some extensions to the server-to-client push mechanism. Here the PLAY_NOTIFY method (see Section 13.5) with a suitable extension could provide clear benefits.

Appendix F. Text format for Parameters

A resource of type "text/parameters" consists of either 1) a list of parameters (for a query) or 2) a list of parameters and associated values (for an response or setting of the parameter). Each entry of the list is a single line of text. Parameters are separated from values by a colon. The parameter name **MUST** only use US-ASCII visible characters while the values are UTF-8 text strings. The media type registration form is in Section 22.16.

There is a potential interoperability issue for this format. It was named in RFC 2326 but never defined, even if used in examples that hint at the syntax. This format matches the purpose and its syntax supports the examples provided. However, it goes further by allowing UTF-8 in the value part, thus usage of UTF-8 strings may not be supported. However, as individual parameters are not defined, the using application anyway needs to have out-of-band agreement or using feature-tag to determine if the end-point supports the parameters.

The ABNF [RFC5234] grammar for "text/parameters" content is:

```

file           = *((parameter / parameter-value) CRLF)
parameter      = 1*visible-except-colon
parameter-value = parameter *WSP ":" value
visible-except-colon = %x21-39 / %x3B-7E      ; VCHAR - ":"
value           = *(TEXT-UTF8char / WSP)
TEXT-UTF8char   = %x21-7E / UTF8-NONASCII
UTF8-NONASCII   = %xC0-DF 1UTF8-CONT
                  / %xE0-EF 2UTF8-CONT
                  / %xF0-F7 3UTF8-CONT
                  / %xF8-FB 4UTF8-CONT
                  / %xFC-FD 5UTF8-CONT
UTF8-CONT       = %x80-BF
WSP              = <See RFC 5234> ; Space or HTAB
VCHAR           = <See RFC 5234>
CRLF            = <See RFC 5234>

```

Appendix G. Requirements for Unreliable Transport of RTSP

This section provides anyone intending to define how to transport of RTSP messages over a unreliable transport protocol with some information learned by the attempt in RFC 2326 [RFC2326]. RFC 2326 defined both an URI scheme and some basic functionality for transport of RTSP messages over UDP, however, it was not sufficient for reliable usage and successful interoperability.

The RTSP scheme defined for unreliable transport of RTSP messages was "rtspu". It has been reserved by this specification as at least one commercial implementation exists, thus avoiding any collisions in the name space.

The following considerations should exist for operation of RTSP over an unreliable transport protocol:

- o Request shall be acknowledged by the receiver. If there is no acknowledgement, the sender may resend the same message after a timeout of one round-trip time (RTT). Any retransmissions due to lack of acknowledgement must carry the same sequence number as the original request.
- o The round-trip time can be estimated as in TCP (RFC 1123) [RFC1123], with an initial round-trip value of 500 ms. An implementation may cache the last RTT measurement as the initial value for future connections.
- o If RTSP is used over a small-RTT LAN, standard procedures for optimizing initial TCP round trip estimates, such as those used in T/TCP (RFC 1644) [RFC1644], can be beneficial.
- o The Timestamp header (Section 16.51) is used to avoid the retransmission ambiguity problem [Stevens98].
- o The registered default port for RTSP over UDP for the server is 554.
- o RTSP messages can be carried over any lower-layer transport protocol that is 8-bit clean.
- o RTSP messages are vulnerable to bit errors and should not be subjected to them.
- o Source authentication, or at least validation that RTSP messages comes from the same entity becomes extremely important, as session hijacking may be substantially easier for RTSP message transport using an unreliable protocol like UDP than for TCP.

There are two RTSP headers that are primarily intended for being used by the unreliable handling of RTSP messages and which will be maintained:

- o CSeq: See Section 16.19
- o Timestamp: See Section 16.51

Appendix H. Backwards Compatibility Considerations

This section contains notes on issues about backwards compatibility with clients or servers being implemented according to RFC 2326 [RFC2326]. Note that there exists no requirement to implement RTSP 1.0; in fact we recommend against it as it is difficult to do in an interoperable way.

A server implementing RTSP/2.0 MUST include an RTSP-Version of RTSP/2.0 in all responses to requests containing RTSP-Version RTSP/2.0. If a server receives an RTSP/1.0 request, it MAY respond with an RTSP/1.0 response if it chooses to support RFC 2326. If the server chooses not to support RFC 2326, it MUST respond with a 505 (RTSP Version not supported) status code. A server MUST NOT respond to an RTSP-Version RTSP/1.0 request with an RTSP-Version RTSP/2.0 response.

Clients implementing RTSP/2.0 MAY use an OPTIONS request with a RTSP-Version of 2.0 to determine whether a server supports RTSP/2.0. If the server responds with either an RTSP-Version of 1.0 or a status code of 505 (RTSP Version not supported), the client will have to use RTSP/1.0 requests if it chooses to support RFC 2326.

H.1. Play Request in Play State

The behavior in the server when a Play is received in Play state has changed (Section 13.4). In RFC 2326, the new PLAY request would be queued until the current Play completed. Any new PLAY request now takes effect immediately replacing the previous request.

H.2. Using Persistent Connections

Some server implementations of RFC 2326 maintain a one-to-one relationship between a connection and an RTSP session. Such implementations require clients to use a persistent connection to communicate with the server and when a client closes its connection, the server may remove the RTSP session. This is worth noting if a RTSP 2.0 client also supporting 1.0 connects to a 1.0 server.

Appendix I. Changes

This appendix briefly lists the differences between RTSP 1.0 [RFC2326] and RTSP 2.0 for an informational purpose. For implementers of RTSP 2.0 it is recommended to read carefully through this memo and not to rely on the list of changes below to adapt from RTSP 1.0 to RTSP 2.0, as RTSP 2.0 is not intended to be backwards compatible with RTSP 1.0 [RFC2326] other than the version negotiation mechanism.

I.1. Brief Overview

The following protocol elements were removed in RTSP 2.0 compared to RTSP 1.0:

- o there is no section on minimal implementation anymore, but more the definition of RTSP 2.0 core;
- o the RECORD and ANNOUNCE methods and all related functionality (including 201 (Created) and 250 (Low On Storage Space) status codes);
- o the use of UDP for RTSP message transport was removed due to missing interest and to broken specification;
- o the use of PLAY method for keep-alive in Play state.

The following protocol elements were added or changed in RTSP 2.0 compared to RTSP 1.0:

- o RTSP session TEARDOWN from the server to the client;
- o IPv6 support;
- o extended IANA registries (e.g., transport headers parameters, transport-protocol, profile, lower-transport, and mode);
- o request pipelining for quick session start-up;
- o fully reworked state-machine;
- o RTSP messages now use URIs rather than URLs;
- o incorporated much of related HTTP text ([RFC2616]) in this memo, compared to just referencing the sections in HTTP, to avoid ambiguities;

- o the REDIRECT method was expanded and diversified for different situations;
- o Includes a new section about how to setup different media transport alternatives and their profiles, and lower layer protocols. This caused the appendix on RTP interaction to be moved there instead of being in the part which describes RTP. The section also includes guidelines what to consider when writing usage guidelines for new protocols and profiles;
- o Added an asynchronous notification method PLAY_NOTIFY. This method is used by the RTSP server to asynchronously notify clients about session changes while in Play state. To a limited extent this is comparable with some implementations of ANNOUNCE in RTSP 1.0 not intended for Recording.

I.2. Detailed List of Changes

Compared to RTSP 1.0 (RFC 2326), the below changes has been made when defining RTSP 2.0. Note that this list does not reflect minor changes in wording or correction of typographical errors.

- o The section on minimal implementation was deleted without substitution.
- o The Transport header has been changed in the following way:
 - * The ABNF has been changed to define that extensions are possible, and that unknown parameters result in that servers ignore the transport specification.
 - * To prevent backwards compatibility issues, any extension or new parameter requires the usage of a feature-tag combined with the Require header.
 - * Syntax unclarities with the Mode parameter has been resolved.
 - * Syntax error with ";" for multicast and unicast has been resolved.
 - * Two new addressing parameters has been defined, src_addr and dest_addr. These replaces the parameters "port", "client_port", "server_port", "destination", "source".
 - * Support for IPv6 explicit addresses in all address fields has been included.

- * To handle URI definitions that contain ";" or "," a quoted URI format has been introduced and is required.
 - * Defined IANA registries for the transport headers parameters, transport-protocol, profile, lower-transport, and mode.
 - * The transport headers interleaved parameter's text was made more strict and uses formal requirements levels. It was also clarified that the interleaved channels are symmetric and that it is the server that sets the channel numbers.
 - * It has been clarified that the client can't request of the server to use a certain RTP SSRC, using a request with the transport parameter SSRC.
 - * Syntax definition for SSRC has been clarified to require 8HEX. It has also been extended to allow multiple values for clients supporting this version.
 - * Clarified the text on the transport headers "dest_addr" parameters regarding what security precautions the server is required to perform.
- o The Range formats has been changed in the following way:
 - * The NPT format has been given an initial NPT identifier that must now be used.
 - * All formats now support initial open ended formats of type "npt=-10" and also format only "Range: smpte" ranges for usage with GET_PARAMETER requests.
 - o RTSP message handling has been changed in the following way:
 - * RTSP messages now use URIs rather than URLs.
 - * It has been clarified that a 4xx message due to missing CSeq header shall be returned without a CSeq header.
 - * The 300 (Multiple Choices) response code has been removed.
 - * Rules for how to handle timing out RTSP messages has been added.
 - * Extended Pipelining rules allowing for quick session startup.

- o The HTTP references have been updated to RFC 2616 and RFC 2617. Most of the text has been copied and then altered to fit RTSP into this specification. Public, and the Content-Base header has also been imported from RFC 2068 so that they are defined in the RTSP specification. Known effects on RTSP due to HTTP clarifications:
 - * Content-Encoding header can include encoding of type "identity".
- o The state machine section has completely been rewritten. It includes now more details and is also more clear about the model used.
- o An IANA section has been included with contains a number of registries and their rules. This will allow us to use IANA to keep track of RTSP extensions.
- o The transport of RTSP messages has seen the following changes:
 - * The use of UDP for RTSP message transport has been deprecated due to missing interest and to broken specification.
 - * The rules for how TCP connections are to be handled has been clarified. Now it is made clear that servers should not close the TCP connection unless they have been unused for significant time.
 - * Strong recommendations why server and clients should use persistent connections have also been added.
 - * There is now a requirement on the servers to handle non-persistent connections as this provides fault tolerance.
 - * Added wording on the usage of Connection:Close for RTSP.
 - * specified usage of TLS for RTSP messages, including a scheme to approve a proxy's TLS connection to the next hop.
- o The following header related changes have been made:
 - * Accept-Ranges response header is added. This header clarifies which range formats that can be used for a resource.
 - * Fixed the missing definitions for the Cache-Control header. Also added to the syntax definition the missing delta-seconds for max-stale and min-fresh parameters.

- * Put requirement on CSeq header that the value is increased by one for each new RTSP request. A Recommendation to start at 0 has also been added.
- * Added requirement that the Date header must be used for all messages with message body and the Server should always include it.
- * Removed possibility of using Range header with Scale header to indicate when it is to be activated, since it can't work as defined. Also added rule that lack of Scale header in response indicates lack of support for the header. Feature-tags for scaled playback has been defined.
- * The Speed header must now be responded to indicate support and the actual speed going to be used. A feature-tag is defined. Notes on congestion control were also added.
- * The Supported header was borrowed from SIP [RFC3261] to help with the feature negotiation in RTSP.
- * Clarified that the Timestamp header can be used to resolve retransmission ambiguities.
- * The Session header text has been expanded with an explanation on keep alive and which methods to use. SET_PARAMETER is now recommended to use if only keep-alive within RTSP is desired.
- * It has been clarified how the Range header formats are used to indicate pause points in the PAUSE response.
- * Clarified that RTP-Info URIs that are relative, use the Request-URI as base URI. Also clarified that the used URI must be the one that was used in the SETUP request. The URIs are now also required to be quoted. The header also expresses the SSRC for the provided RTP timestamp and sequence number values.
- * Added text that requires the Range to always be present in PLAY responses. Clarified what should be sent in case of live streams.
- * The headers table has been updated using a structure borrowed from SIP. Those tables carries much more information and should provide a good overview of the available headers.
- * It has been clarified that any message with a message body is required to have a Content-Length header. This was the case in RFC 2326, but could be misinterpreted.

- * ETag has changed name to MTag.
 - * To resolve functionality around MTag. The MTag and If-None-Match header have been added from HTTP with necessary clarification in regards to RTSP operation.
 - * Imported the Public header from HTTP RFC 2068 [RFC2068] since it has been removed from HTTP due to lack of use. Public is used quite frequently in RTSP.
 - * Clarified rules for populating the Public header so that it is an intersection of the capabilities of all the RTSP agents in a chain.
 - * Added the Media-Range header for listing the current availability of the media range.
 - * Added the Notify-Reason header for giving the reason when sending PLAY_NOTIFY requests.
 - * A new header Seek-Style has been defined to direct and inform how any seek operation should/have been performed.
- o The Protocol Syntax has been changed in the following way:
- * All ABNF definitions are updated according to the rules defined in RFC 5234 [RFC5234] and have been gathered in a separate Section 20.
 - * The ABNF for the User-Agent and Server headers have been corrected.
 - * Some definitions in the introduction regarding the RTSP session have been changed.
 - * The protocol has been made fully IPv6 capable.
 - * Added a fragment part to the RTSP URI. This seemed to be indicated by the note below the definition, however, it was not part of the ABNF.
 - * The CHAR rule has been changed to exclude NULL.
- o The Status codes have been changed in the following way:
- * The use of status code 303 "See Other" has been deprecated as it does not make sense to use in RTSP.

- * When sending response 451 and 458 the response body should contain the offending parameters.
 - * Clarification on when a 3rr redirect status code can be received has been added. This includes receiving 3rr as a result of a request within a established session. This provides clarification to a previous unspecified behavior.
 - * Removed the 201 (Created) and 250 (Low On Storage Space) status codes as they are only relevant to recording, which is deprecated.
 - * Several new Status codes have been defined: 464 "Data Transport Not Ready Yet", 465 "Notification Reason Unknown", 470 "Connection Authorization Required", 471 "Connection Credentials not accepted", 472 "Failure to establish secure connection".
- o The following functionality has been deprecated from the protocol:
 - * The use of Queued Play.
 - * The use of PLAY method for keep-alive in Play state.
 - * The RECORD and ANNOUNCE methods and all related functionality. Some of the syntax has been removed.
 - * The possibility to use timed execution of methods with the time parameter in the Range header.
 - * The description on how rtspu works is not part of the core specification and will require external description. Only that it exist is defined here and some requirements for the transport is provided.
 - o The following changes have been made in relation to methods:
 - * The OPTIONS method has been clarified with regards to the use of the Public and Allow headers.
 - * Added text clarifying the usage of SET_PARAMETER for keep-alive and usage without any body.
 - * PLAY method is now allowed to be pipelined with the pipelining of one or more SETUP requests following the initial that generates the session for aggregated control.

- * REDIRECT has been expanded and diversified for different situations.
- * Added a new method PLAY_NOTIFY. This method is used by the RTSP server to asynchronously notify clients about session changes.
- o Wrote a new section about how to setup different media transport alternatives and their profiles, and lower layer protocols. This caused the appendix on RTP interaction to be moved there instead of being in the part which describes RTP. The section also includes guidelines what to consider when writing usage guidelines for new protocols and profiles.
- o Setup and usage of independent TCP connections for transport of RTP has been specified.
- o Added a new section describing the available mechanisms to determine if functionality is supported, called "Capability Handling". Renamed option-tags to feature-tags.
- o Added a contributors section with people who have contributed actual text to the specification.
- o Added a section Use Cases that describes the major use cases for RTSP.
- o Clarified the usage of a=range and how to indicate live content that are not seekable with this header.
- o Text specifying the special behavior of PLAY for live content.

Appendix J. Acknowledgements

This memorandum defines RTSP version 2.0 which is a revision of the Proposed Standard RTSP version 1.0 which is defined in [RFC2326]. The authors of RFC 2326 are Henning Schulzrinne, Anup Rao, and Robert Lanphier.

Both RTSP version 1.0 and RTSP version 2.0 borrow format and descriptions from HTTP/1.1.

This document has benefited greatly from the comments of all those participating in the MMUSIC-WG. In addition to those already mentioned, the following individuals have contributed to this specification:

Rahul Agarwal, Jeff Ayars, Milko Boic, Torsten Braun, Brent Browning, Bruce Butterfield, Steve Casner, Francisco Cortes, Kelly Djahandari, Martin Dunsmuir, Eric Fleischman, Jay Geagan, Andy Grignon, V. Guruprasad, Peter Haight, Mark Handley, Brad Hefta-Gaub, Volker Hilt, John K. Ho, Go Hori, Philipp Hoschka, Anne Jones, Ingemar Johansson, Anders Klemets, Ruth Lang, Stephanie Leif, Jonathan Lennox, Eduardo F. Llach, Thomas Marshall, Rob McCool, David Oran, Joerg Ott, Maria Papadopouli, Sujal Patel, Ema Patki, Alagu Periyannan, Colin Perkins, Igor Plotnikov, Jonathan Sergeant, Pinaki Shah, David Singer, Lior Sion, Jeff Smith, Alexander Sokolsky, Dale Stammen, John Francis Stracke, Maureen Chesire, David Walker, Geetha Srikantan, Stephan Wenger, Pekka Pessi, Jae-Hwan Kim, Holger Schmidt, Stephen Farrell, Xavier Marjou, Joe Pallas, Martti Mela, Byungjo Yoon and Patrick Hoffman, Jinhang Choi, Ross Finlayson, and especially to Flemming Andreassen.

J.1. Contributors

The following people have made written contributions that were included in the specification:

- o Tom Marshall contributed text on the usage of 3rr status codes.
- o Thomas Zheng contributed text on the usage of the Range in PLAY responses and proposed an earlier version of the PLAY_NOTIFY method.
- o Sean Sheedy contributed text on the timeout behavior of RTSP messages and connections, the 463 status code, and proposed an earlier version of the PLAY_NOTIFY method.
- o Greg Sherwood proposed an earlier version of the PLAY_NOTIFY method.

- o Fredrik Lindholm contributed text about the RTSP security framework.
- o John Lazzaro contributed the text for RTP over Independent TCP.
- o Aravind Narasimhan contributed by rewriting Media Transport Alternatives (Appendix C) and editorial improvements on a number of places in the specification.
- o Torbjorn Einarsson has done some editorial improvements of the text.

Appendix K. RFC Editor Consideration

Please replace RFC XXXX with the RFC number this specification receives.

Authors' Addresses

Henning Schulzrinne
Columbia University
1214 Amsterdam Avenue
New York, NY 10027
USA

Email: schulzrinne@cs.columbia.edu

Anup Rao
Cisco
USA

Email: anrao@cisco.com

Rob Lanphier
Seattle, WA
USA

Email: robla@robla.net

Magnus Westerlund
Ericsson AB
Faeroegatan 6
STOCKHOLM, SE-164 80
SWEDEN

Email: magnus.westerlund@ericsson.com

Martin Stiemerling
NEC Laboratories Europe, NEC Europe Ltd.
Kurfuersten-Anlage 36
Heidelberg 69115
Germany

Phone: +49 (0) 6221 4342 113
Email: martin.stiemerling@neclab.eu
URI: <http://ietf.stiemerling.org>

MMUSIC Working Group
Internet-Draft
Obsoletes: 2326 (if approved)
Intended status: Standards Track
Expires: August 14, 2014

H. Schulzrinne
Columbia University
A. Rao
Cisco
R. Lanphier

M. Westerlund
Ericsson AB
M. Stiemerling (Ed.)
NEC
February 10, 2014

Real Time Streaming Protocol 2.0 (RTSP)
draft-ietf-mmusic-rfc2326bis-40

Abstract

This memorandum defines RTSP version 2.0 which obsoletes RTSP version 1.0 defined in RFC 2326.

The Real Time Streaming Protocol, or RTSP, is an application-layer protocol for setup and control of the delivery of data with real-time properties. RTSP provides an extensible framework to enable controlled, on-demand delivery of real-time data, such as audio and video. Sources of data can include both live data feeds and stored clips. This protocol is intended to control multiple data delivery sessions, provide a means for choosing delivery channels such as UDP, multicast UDP and TCP, and provide a means for choosing delivery mechanisms based upon RTP (RFC 3550).

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 14, 2014.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	10
2. Protocol Overview	11
2.1. Presentation Description	11
2.2. Session Establishment	12
2.3. Media Delivery Control	13
2.4. Session Parameter Manipulations	15
2.5. Media Delivery	16
2.5.1. Media Delivery Manipulations	16
2.6. Session Maintenance and Termination	19
2.7. Extending RTSP	20
3. Document Conventions	21
3.1. Notational Conventions	21
3.2. Terminology	21
4. Protocol Parameters	24
4.1. RTSP Version	24
4.2. RTSP IRI and URI	25
4.3. Session Identifiers	27
4.4. Media Time Formats	27
4.4.1. SMPTE Relative Timestamps	28

4.4.2.	Normal Play Time	28
4.4.3.	Absolute Time	30
4.5.	Feature-Tags	30
4.6.	Message Body Tags	31
4.7.	Media Properties	31
4.7.1.	Random Access and Seeking	32
4.7.2.	Retention	33
4.7.3.	Content Modifications	33
4.7.4.	Supported Scale Factors	33
4.7.5.	Mapping to the Attributes	34
5.	RTSP Message	34
5.1.	Message Types	34
5.2.	Message Headers	35
5.3.	Message Body	36
5.4.	Message Length	36
6.	General Header Fields	36
7.	Request	38
7.1.	Request Line	38
7.2.	Request Header Fields	40
8.	Response	42
8.1.	Status-Line	42
8.1.1.	Status Code and Reason Phrase	42
8.2.	Response Headers	46
9.	Message Body	46
9.1.	Message-Body Header Fields	47
9.2.	Message Body	48
9.3.	Message Body Format Negotiation	48
10.	Connections	49
10.1.	Reliability and Acknowledgements	49
10.2.	Using Connections	50
10.3.	Closing Connections	53
10.4.	Timing Out Connections and RTSP Messages	54
10.5.	Showing Liveness	55
10.6.	Use of IPv6	57
10.7.	Overload Control	57
11.	Capability Handling	58
11.1.	Feature Tag: play.basic	60
12.	Pipelining Support	61
13.	Method Definitions	61
13.1.	OPTIONS	63
13.2.	DESCRIBE	64
13.3.	SETUP	66
13.3.1.	Changing Transport Parameters	69
13.4.	PLAY	70
13.4.1.	General Usage	70
13.4.2.	Aggregated Sessions	75
13.4.3.	Updating current PLAY Requests	76
13.4.4.	Playing On-Demand Media	79

13.4.5.	Playing Dynamic On-Demand Media	79
13.4.6.	Playing Live Media	79
13.4.7.	Playing Live with Recording	80
13.4.8.	Playing Live with Time-Shift	81
13.5.	PLAY_NOTIFY	81
13.5.1.	End-of-Stream	82
13.5.2.	Media-Properties-Update	84
13.5.3.	Scale-Change	85
13.6.	PAUSE	86
13.7.	TEARDOWN	89
13.7.1.	Client to Server	89
13.7.2.	Server to Client	90
13.8.	GET_PARAMETER	91
13.9.	SET_PARAMETER	93
13.10.	REDIRECT	95
14.	Embedded (Interleaved) Binary Data	97
15.	Proxies	99
15.1.	Proxies and Protocol Extensions	101
15.2.	Multiplexing and Demultiplexing of Messages	102
16.	Caching	102
16.1.	Validation Model	103
16.1.1.	Last-Modified Dates	104
16.1.2.	Message Body Tag Cache Validators	104
16.1.3.	Weak and Strong Validators	104
16.1.4.	Rules for When to Use Message Body Tags and Last-Modified Dates	107
16.1.5.	Non-validating Conditionals	108
16.2.	Invalidation After Updates or Deletions	108
17.	Status Code Definitions	109
17.1.	Informational 1xx	109
17.1.1.	100 Continue	109
17.2.	Success 2xx	110
17.2.1.	200 OK	110
17.3.	Redirection 3xx	110
17.3.1.	300	111
17.3.2.	301 Moved Permanently	111
17.3.3.	302 Found	111
17.3.4.	303 See Other	111
17.3.5.	304 Not Modified	111
17.3.6.	305 Use Proxy	112
17.4.	Client Error 4xx	112
17.4.1.	400 Bad Request	112
17.4.2.	401 Unauthorized	112
17.4.3.	402 Payment Required	113
17.4.4.	403 Forbidden	113
17.4.5.	404 Not Found	113
17.4.6.	405 Method Not Allowed	113
17.4.7.	406 Not Acceptable	113

17.4.8.	407 Proxy Authentication Required	114
17.4.9.	408 Request Timeout	114
17.4.10.	410 Gone	114
17.4.11.	411 Length Required	114
17.4.12.	412 Precondition Failed	114
17.4.13.	413 Request Message Body Too Large	115
17.4.14.	414 Request-URI Too Long	115
17.4.15.	415 Unsupported Media Type	115
17.4.16.	451 Parameter Not Understood	115
17.4.17.	452 reserved	115
17.4.18.	453 Not Enough Bandwidth	116
17.4.19.	454 Session Not Found	116
17.4.20.	455 Method Not Valid in This State	116
17.4.21.	456 Header Field Not Valid for Resource	116
17.4.22.	457 Invalid Range	116
17.4.23.	458 Parameter Is Read-Only	116
17.4.24.	459 Aggregate Operation Not Allowed	116
17.4.25.	460 Only Aggregate Operation Allowed	116
17.4.26.	461 Unsupported Transport	117
17.4.27.	462 Destination Unreachable	117
17.4.28.	463 Destination Prohibited	117
17.4.29.	464 Data Transport Not Ready Yet	117
17.4.30.	465 Notification Reason Unknown	117
17.4.31.	466 Key Management Error	117
17.4.32.	470 Connection Authorization Required	118
17.4.33.	471 Connection Credentials not accepted	118
17.4.34.	472 Failure to establish secure connection	118
17.5.	Server Error 5xx	118
17.5.1.	500 Internal Server Error	118
17.5.2.	501 Not Implemented	118
17.5.3.	502 Bad Gateway	118
17.5.4.	503 Service Unavailable	119
17.5.5.	504 Gateway Timeout	119
17.5.6.	505 RTSP Version Not Supported	119
17.5.7.	551 Option not supported	119
17.5.8.	553 Proxy Unavailable	119
18.	Header Field Definitions	120
18.1.	Accept	131
18.2.	Accept-Credentials	131
18.3.	Accept-Encoding	132
18.4.	Accept-Language	133
18.5.	Accept-Ranges	134
18.6.	Allow	134
18.7.	Authentication-Info	135
18.8.	Authorization	135
18.9.	Bandwidth	136
18.10.	Blocksize	136
18.11.	Cache-Control	137

18.12.	Connection	139
18.13.	Connection-Credentials	140
18.14.	Content-Base	141
18.15.	Content-Encoding	141
18.16.	Content-Language	142
18.17.	Content-Length	143
18.18.	Content-Location	143
18.19.	Content-Type	144
18.20.	CSeq	144
18.21.	Date	146
18.22.	Expires	147
18.23.	From	148
18.24.	If-Match	148
18.25.	If-Modified-Since	149
18.26.	If-None-Match	149
18.27.	Last-Modified	150
18.28.	Location	150
18.29.	Media-Properties	151
18.30.	Media-Range	153
18.31.	MTag	153
18.32.	Notify-Reason	154
18.33.	Pipelined-Requests	154
18.34.	Proxy-Authenticate	155
18.35.	Proxy-Authentication-Info	155
18.36.	Proxy-Authorization	156
18.37.	Proxy-Require	156
18.38.	Proxy-Supported	156
18.39.	Public	157
18.40.	Range	158
18.41.	Referrer	160
18.42.	Request-Status	160
18.43.	Require	161
18.44.	Retry-After	162
18.45.	RTP-Info	162
18.46.	Scale	164
18.47.	Seek-Style	165
18.48.	Server	167
18.49.	Session	167
18.50.	Speed	168
18.51.	Supported	169
18.52.	Terminate-Reason	170
18.53.	Timestamp	170
18.54.	Transport	171
18.55.	Unsupported	178
18.56.	User-Agent	178
18.57.	Via	179
18.58.	WWW-Authenticate	179
19.	Security Framework	180

19.1.	RTSP and HTTP Authentication	180
19.1.1.	Digest Authentication	181
19.2.	RTSP over TLS	182
19.3.	Security and Proxies	183
19.3.1.	Accept-Credentials	184
19.3.2.	User approved TLS procedure	185
20.	Syntax	187
20.1.	Base Syntax	187
20.2.	RTSP Protocol Definition	189
20.2.1.	Generic Protocol elements	190
20.2.2.	Message Syntax	192
20.2.3.	Header Syntax	196
20.3.	SDP extension Syntax	205
21.	Security Considerations	205
21.1.	Signaling Protocol Threats	206
21.2.	Media Stream Delivery Threats	209
21.2.1.	Remote Denial of Service Attack	210
21.2.2.	RTP Security analysis	211
22.	IANA Considerations	212
22.1.	Feature-tags	213
22.1.1.	Description	214
22.1.2.	Registering New Feature-tags with IANA	214
22.1.3.	Registered entries	214
22.2.	RTSP Methods	215
22.2.1.	Description	215
22.2.2.	Registering New Methods with IANA	215
22.2.3.	Registered Entries	216
22.3.	RTSP Status Codes	216
22.3.1.	Description	216
22.3.2.	Registering New Status Codes with IANA	216
22.3.3.	Registered Entries	217
22.4.	RTSP Headers	217
22.4.1.	Description	217
22.4.2.	Registering New Headers with IANA	217
22.4.3.	Registered entries	217
22.5.	Accept-Credentials	219
22.5.1.	Accept-Credentials policies	219
22.5.2.	Accept-Credentials hash algorithms	219
22.6.	Cache-Control Cache Directive Extensions	220
22.7.	Media Properties	221
22.7.1.	Description	221
22.7.2.	Registration Rules	221
22.7.3.	Registered Values	221
22.8.	Notify-Reason header	222
22.8.1.	Description	222
22.8.2.	Registration Rules	222
22.8.3.	Registered Values	222
22.9.	Range Header Formats	223

22.9.1.	Description	223
22.9.2.	Registration Rules	223
22.9.3.	Registered Values	223
22.10.	Terminate-Reason Header	223
22.10.1.	Redirect Reasons	224
22.10.2.	Terminate-Reason Header Parameters	224
22.11.	RTP-Info header parameters	225
22.11.1.	Description	225
22.11.2.	Registration Rules	225
22.11.3.	Registered Values	225
22.12.	Seek-Style Policies	225
22.12.1.	Description	226
22.12.2.	Registration Rules	226
22.12.3.	Registered Values	226
22.13.	Transport Header Registries	227
22.13.1.	Transport Protocol Identifier	227
22.13.2.	Transport modes	228
22.13.3.	Transport Parameters	229
22.14.	URI Schemes	230
22.14.1.	The rtsp URI Scheme	230
22.14.2.	The rtspS URI Scheme	231
22.14.3.	The rtspu URI Scheme	232
22.15.	SDP attributes	233
22.16.	Media Type Registration for text/parameters	234
23.	References	235
23.1.	Normative References	235
23.2.	Informative References	239
Appendix A.	Examples	241
A.1.	Media on Demand (Unicast)	241
A.2.	Media on Demand using Pipelining	245
A.3.	Secured Media Session for on Demand Content	247
A.4.	Media on Demand (Unicast)	250
A.5.	Single Stream Container Files	254
A.6.	Live Media Presentation Using Multicast	256
A.7.	Capability Negotiation	257
Appendix B.	RTSP Protocol State Machine	258
B.1.	States	259
B.2.	State variables	259
B.3.	Abbreviations	259
B.4.	State Tables	260
Appendix C.	Media Transport Alternatives	264
C.1.	RTP	264
C.1.1.	AVP	265
C.1.2.	AVP/UDP	265
C.1.3.	AVPF/UDP	266
C.1.4.	SAVP/UDP	267
C.1.5.	SAVPF/UDP	269
C.1.6.	RTCP usage with RTSP	269

C.2.	RTP over TCP	271
C.2.1.	Interleaved RTP over TCP	271
C.2.2.	RTP over independent TCP	272
C.3.	Handling Media Clock Time Jumps in the RTP Media Layer	276
C.4.	Handling RTP Timestamps after PAUSE	280
C.5.	RTSP / RTP Integration	282
C.6.	Scaling with RTP	282
C.7.	Maintaining NPT synchronization with RTP timestamps	282
C.8.	Continuous Audio	282
C.9.	Multiple Sources in an RTP Session	282
C.10.	Usage of SSRCs and the RTCP BYE Message During an RTSP Session	283
C.11.	Future Additions	283
Appendix D.	Use of SDP for RTSP Session Descriptions	284
D.1.	Definitions	284
D.1.1.	Control URI	284
D.1.2.	Media Streams	286
D.1.3.	Payload Type(s)	286
D.1.4.	Format-Specific Parameters	286
D.1.5.	Directionality of media stream	287
D.1.6.	Range of Presentation	287
D.1.7.	Time of Availability	288
D.1.8.	Connection Information	288
D.1.9.	Message Body Tag	289
D.2.	Aggregate Control Not Available	289
D.3.	Aggregate Control Available	290
D.4.	Grouping of Media Lines in SDP	291
D.5.	RTSP external SDP delivery	292
Appendix E.	RTSP Use Cases	292
E.1.	On-demand Playback of Stored Content	292
E.2.	Unicast Distribution of Live Content	294
E.3.	On-demand Playback using Multicast	294
E.4.	Inviting an RTSP server into a conference	295
E.5.	Live Content using Multicast	296
Appendix F.	Text format for Parameters	296
Appendix G.	Requirements for Unreliable Transport of RTSP	297
Appendix H.	Backwards Compatibility Considerations	298
H.1.	Play Request in Play State	299
H.2.	Using Persistent Connections	299
Appendix I.	Changes	299
I.1.	Brief Overview	299
I.2.	Detailed List of Changes	300
Appendix J.	Acknowledgements	307
J.1.	Contributors	308
Appendix K.	RFC Editor Consideration	308
Authors' Addresses		308

1. Introduction

This memo defines version 2.0 of the Real Time Streaming Protocol (RTSP 2.0). RTSP 2.0 is an application-layer protocol for setup and control over the delivery of data with real-time properties, typically streaming media. Streaming media is, for instance, video on demand or audio live streaming. Put simply, RTSP acts as a "network remote control" for multimedia servers.

The protocol operates between RTSP 2.0 clients and servers, but also supports the usage of proxies placed between clients and servers. Clients can request information about streaming media from servers by asking for a description of the media or use media description provided externally. The media delivery protocol is used to establish the media streams described by the media description. Clients can then request to play out the media, pause it, or stop it completely. The requested media can consist of multiple audio and video streams that are delivered as time-synchronized streams from servers to clients.

RTSP 2.0 is a replacement of RTSP 1.0 [RFC2326] and obsoletes that specification. This protocol is based on RTSP 1.0 but is not backwards compatible other than in the basic version negotiation mechanism. The changes are documented in Appendix I. There are many reasons why RTSP 2.0 can't be backwards compatible with RTSP 1.0 but some of the main ones are:

- o Most headers that needed to be extensible did not define the allowed syntax, preventing safe deployment of extensions;
- o The changed behavior of the PLAY method when received in Play state;
- o Changed behavior of the extensibility model and its mechanism;
- o The change of syntax for some headers.

In summary, there are so many small details that changing version became necessary to enable clarification and consistent behavior. Anyone implementing RTSP for a new usage where they have no installed based of RTSP 1.0 should only implement RTSP 2.0 to avoid having to deal with RTSP 1.0 inconsistencies.

This document is structured as follows. It begins with an overview of the protocol operations and its functions in an informal way. Then a set of definitions of terms used and document conventions is introduced. It is followed by the actual RTSP 2.0 core protocol specification. The appendixes describe and define some

functionalities that are not part of the core RTSP specification, but which are still important to enable some usages. Among them, the RTP usage is defined in Appendix C, the Session Description Protocol (SDP) usage with RTSP is defined in Appendix D, and the text/parameters file format Appendix F, are three normative specification appendices. Others include a number of informational parts discussing the changes, use cases, different considerations or motivations.

2. Protocol Overview

This section provides an informative overview of the different mechanisms in the RTSP 2.0 protocol, to give the reader a high level understanding before getting into all the different details. In case of conflict with this description and the later sections, the later sections take precedence. For more information about use cases considered for RTSP see Appendix E.

RTSP 2.0 is a bi-directional request and response protocol that first establishes a context including content resources (the media) and then controls the delivery of these content resources from the provider to the consumer. RTSP has three fundamental parts: Session Establishment, Media Delivery Control, and an extensibility model described below. The protocol is based on some assumptions about existing functionality to provide a complete solution for client controlled real-time media delivery.

RTSP uses text-based messages, requests and responses, that may contain a binary message body. An RTSP request starts with a method line that identifies the method, the protocol and version and the resource to act on. The resource is identified by a URI and the hostname part of the URI is used by RTSP client to resolve the IPv4 or IPv6 address of the RTSP server. Following the method line are a number of RTSP headers. This part is ended by two consecutive carriage return line feed (CRLF) character pairs. The message body if present follows the two CRLF and the body's length is described by a message header. RTSP responses are similar, but start with a response line with the protocol and version, followed by a status code and a reason phrase. RTSP messages are sent over a reliable transport protocol between the client and server. RTSP 2.0 requires clients and servers to implement TCP, and TLS over TCP, as mandatory transports for RTSP messages.

2.1. Presentation Description

RTSP exists to provide access to multi-media presentations and content, but tries to be agnostic about the media type or the actual media delivery protocol that is used. To enable a client to

implement a complete system, an RTSP-external mechanism for describing the presentation and the delivery protocol(s) is used. RTSP assumes that this description is either delivered completely out of band or as a data object in the response to a client's request using the DESCRIBE method (Section 13.2).

Parameters that commonly have to be included in the Presentation Description are the following:

- o Number of media streams;
- o The resource identifier for each media stream/resource that is to be controlled by RTSP;
- o The protocol that each media stream is to be delivered over;
- o Transport protocol parameters that are not negotiated or vary with each client;
- o Media encoding information enabling a client to correctly decode the media upon reception;
- o An aggregate control resource identifier.

RTSP uses its own URI schemes ("rtsp" and "rtsps") to reference media resources and aggregates under common control (See Section 4.2).

This specification describes in Appendix D how one uses SDP [RFC4566] for Presentation Description

2.2. Session Establishment

The RTSP client can request the establishment of an RTSP session after having used the presentation description to determine which media streams are available, which media delivery protocol is used and the resource identifiers of the media streams. The RTSP session is a common context between the client and the server that consists of one or more media resources that are to be under common media delivery control.

The client creates an RTSP session by sending a request using the SETUP method (Section 13.3) to the server. In the "Transport" header (Section 18.54) of the SETUP request, the client also includes all the transport parameters necessary to enable the media delivery protocol to function. This includes parameters that are pre-established by the presentation description but necessary for any middlebox to correctly handle the media delivery protocols. The Transport header in a request may contain multiple alternatives for

media delivery in a prioritized list, which the server can select from. These alternatives are typically based on information in the presentation description.

The server determines if the media resource is available upon receiving a SETUP request and if any of the transport parameter specifications are acceptable. If that is successful, an RTSP session context is created and the relevant parameters and state is stored. An identifier is created for the RTSP session and included in the response in the Session header (Section 18.49). The SETUP response includes a Transport header that specifies which of the alternatives has been selected and relevant parameters.

A SETUP request that references an existing RTSP session but identifies a new media resource is a request to add that media resource under common control with the already present media resources in an aggregated session. A client can expect this to work for all media resources under RTSP control within a multi-media content. However, aggregating resources from different content are likely to be refused by the server. Even if a RTSP session contains only a single media, the RTSP session can be referenced by the aggregate control URI.

To avoid an extra round trip in the session establishment of aggregated RTSP sessions, RTSP 2.0 supports pipelined requests; i.e., the client can send multiple requests back-to-back without waiting first for the completion of any of them. The client uses a client-selected identifier in the Pipelined-Requests header (Section 18.33) to instruct the server to bind multiple requests together as if they included the session identifier.

The SETUP response also provides additional information about the established sessions in a couple of different headers. The Media-Properties header (Section 18.29) includes a number of properties that apply for the aggregate that is valuable when doing media delivery control and configuring user interface. The Accept-Ranges header (Section 18.5) informs the client about which range formats that the server supports with these media resources. The Media-Range header (Section 18.30) informs the client about the time range of the media currently available.

2.3. Media Delivery Control

After having established an RTSP session, the client can start controlling the media delivery. The basic operations are Start by using the PLAY method (Section 13.4) and Halt by using the PAUSE method (Section 13.6). PLAY also allows for choosing the starting media position from which the server should deliver the media. The

positioning is done by using the Range header (Section 18.40) that supports several different time formats: Normal Play Time (NPT) (Section 4.4.2), Society of Motion Picture and Television Engineers (SMPTE) Timestamps (Section 4.4.1) and absolute time (Section 4.4.3). The Range header does further allow the client to specify a position where delivery should end, thus allowing a specific interval to be delivered.

The support for positioning/searching within a media content depends on the content's media properties. Content exists in a number of different types, such as: on-demand, live, and live with simultaneous recording. Even within these categories there are differences in how the content is generated and distributed, which affect how it can be accessed for playback. The properties applicable for the RTSP session are provided by the server in the SETUP response using the Media-Properties header (Section 18.29). These are expressed using one or several independent attributes. A first attribute is Random Access, which expresses if positioning can be done, and with what granularity. Another aspect is whether the content will change during the lifetime of the session. While on-demand content will be provided in full from the beginning, a live stream being recorded results in the length of the accessible content growing as the session goes on. There also exists content that is dynamically built by another protocol than RTSP and thus also changes in steps during the session, but maybe not continuously. Furthermore, when content is recorded, there are cases where not the complete content is maintained, but, for example, only the last hour. All these properties result in the need for mechanisms that will be discussed below.

When the client accesses on-demand content that allows random access, the client can issue the PLAY request for any point in the content between the start and the end. The server will deliver media from the closest random access point prior to the requested point and indicate that in its PLAY response. If the client issues a PAUSE, the delivery will be halted and the point at which the server stopped will be reported back in the response. The client can later resume by sending a PLAY request without a range header. When the server is about to complete the PLAY request by delivering the end of the content or the requested range, the server will send a PLAY_NOTIFY request (Section 13.5) indicating this.

When playing live content with no extra functions, such as recording, the client will receive the live media from the server after having sent a PLAY request. Seeking in such content is not possible as the server does not store it, but only forwards it from the source of the session. Thus delivery continues until the client sends a PAUSE request, tears down the session, or the content ends.

For live sessions that are being recorded the client will need to keep track of how the recording progresses. Upon session establishment the client will learn the current duration of the recording from the Media-Range header. As the recording is ongoing the content grows in direct relation to the passed time. Therefore, each server's response to a PLAY request will contain the current Media-Range header. The server should also regularly send approximately every 5 minutes the current media range in a PLAY_NOTIFY request (Section 13.5.2). If the live transmission ends, the server must send a PLAY_NOTIFY request with the updated Media-Properties indicating that the content stopped being a recorded live session and instead became on-demand content; the request also contains the final media range. While the live delivery continues the client can request to play the current live point by using the NPT timescale symbol "now", or it can request a specific point in the available content by an explicit range request for that point. If the requested point is outside of the available interval the server will adjust the position to the closest available point, i.e., either at the beginning or the end.

A special case of recording is that where the recording is not retained longer than a specific time period, thus as the live delivery continues the client can access any media within a moving window that covers, for example, "now" to "now" minus 1 hour. A client that pauses on a specific point within the content may not be able to retrieve the content anymore. If the client waits too long before resuming the pause point, the content may no longer be available. In this case the pause point will be adjusted to the closest point in the available media.

2.4. Session Parameter Manipulations

A session may have additional state or functionality that affects how the server or client treats the session, content, how it functions, or feedback on how well the session works. Such extensions are not defined in this specification, but may be done in various extensions. RTSP has two methods for retrieving and setting parameter values on either the client or the server: GET_PARAMETER (Section 13.8) and SET_PARAMETER (Section 13.9). These methods carry the parameters in a message body of the appropriate format. One can also use headers to query state with the GET_PARAMETER method. As an example, clients needing to know the current media-range for a time-progressing session can use the GET_PARAMETER method and include the media-range. Furthermore, synchronization information can be requested by using a combination of RTP-Info (Section 18.45) and Range (Section 18.40).

RTSP 2.0 does not have a strong mechanism for providing negotiation of the headers, or parameters and their formats, that can be used.

However, responses will indicate request-headers or parameters that are not supported. A priori determination of what features are available needs to be done through out-of-band mechanisms, like the session description, or through the usage of feature tags (Section 4.5).

2.5. Media Delivery

This document specifies how media is delivered with RTP [RFC3550] over UDP [RFC0768], TCP [RFC0793] or the RTSP connection. Additional protocols may be specified in the future based on demand.

The usage of RTP as media delivery protocol requires some additional information to function well. The PLAY response contains information to enable reliable and timely delivery of how a client should synchronize different sources in the different RTP sessions. It also provides a mapping between RTP timestamps and the content time scale. When the server wants to notify the client about the completion of the media delivery, it sends a PLAY_NOTIFY request to the client. The PLAY_NOTIFY request includes information about the stream end, including the last RTP sequence number for each stream, thus enabling the client to empty the buffer smoothly.

2.5.1. Media Delivery Manipulations

The basic playback functionality of RTSP enables delivery of a range of requested content to the client at the pace intended by the content's creator. However, RTSP can also manipulate the delivery to the client in two ways.

Scale: The ratio of media content time delivered per unit playback time.

Speed: The ratio of playback time delivered per unit of wallclock time.

Both affect the media delivery per time unit. However, they manipulate two independent time scales and the effects are possible to combine.

Scale (Section 18.46) is used for fast forward or slow motion control as it changes the amount of content timescale that should be played back per time unit. Scale > 1.0, means fast forward, e.g., Scale=2.0 results in that 2 seconds of content is played back every second of playback. Scale = 1.0 is the default value that is used if no Scale is specified, i.e., playback at the content's original rate. Scale values between 0 and 1.0 is providing for slow motion. Scale can be negative to allow for reverse playback in either regular pace (Scale

= -1.0) or fast backwards (Scale < -1.0) or slow motion backwards (-1.0 < Scale < 0). Scale = 0 would be equal to pause and is not allowed.

In most cases the realization of scale means server side manipulation of the media to ensure that the client can actually play it back. The nature of these media manipulations and when they are needed is highly media-type dependent. Let's consider an example with two common media types audio and video.

It is very difficult to modify the playback rate of audio. A maximum of 10-30% is possible by changing the pitch-rate of speech. Music goes out of tune if one tries to manipulate the playback rate by resampling it. This is a well known problem and audio is commonly muted or played back in short segments with skips to keep up with the current playback point.

For video it is possible to manipulate the frame rate, although the rendering capabilities are often limited to certain frame rates. Also the allowed bitrates in decoding, the structure used in the encoding and the dependency between frames and other capabilities of the rendering device limits the possible manipulations. Therefore, the basic fast forward capabilities often are implemented by selecting certain subsets of frames.

Due to the media restrictions, the possible scale values are commonly restricted to the set of realizable scale ratios. To enable the clients to select from the possible scale values, RTSP can signal the supported Scale ratios for the content. To support aggregated or dynamic content, where this may change during the ongoing session and dependent on the location within the content, a mechanism for updating the media properties and the scale factor currently in use, exists.

Speed (Section 18.50) affects how much of the playback timeline is delivered in a given wallclock period. The default is Speed = 1 which means to deliver at the same rate the media is consumed. Speed > 1 means that the receiver will get content faster than it regularly would consume it. Speed < 1 means that delivery is slower than the regular media rate. Speed values of 0 or lower have no meaning and are not allowed. This mechanism enables two general functionalities. One is client side scale operations, i.e., the client receives all the frames and makes the adjustment to the playback locally. The second is delivery control for buffering of media. By specifying a speed over 1.0 the client can build up the amount of playback time it has present in its buffers to a level that is sufficient for its needs.

A naive implementation of Speed would only affect the transmission schedule of the media and has a clear impact on the needed bandwidth. This would result in the data rate being proportional to the speed factor. Speed = 1.5, i.e., 50% faster than normal delivery, would result in a 50% increase in the data transport rate. If that can be supported or not depends solely on the underlying network path. Scale may also have some impact on the required bandwidth due to the manipulation of the content in the new playback schedule. An example is fast forward where only the independently decodable intra frames are included in the media stream. This usage of solely intra frames increases the data rate significantly compared to a normal sequence with the same number of frames, where most frames are encoded using prediction.

This potential increase of the data rate needs to be handled by the media sender. The client has requested that the media will be delivered in a specific way, which should be honored. However, the media sender cannot ignore if the network path between the sender and the receiver can't handle the resulting media stream. In that case the media stream needs to be adapted to fit the available resources of the path. This can result in a reduced media quality.

The need for bitrate adaptation becomes especially problematic in connection with the Speed semantics. If the goal is to fill up the buffer, the client may not want to do that at the cost of reduced quality. If the client wants to make local playout changes then it may actually require that the requested speed be honored. To resolve this issue, Speed uses a range so that both cases can be supported. The server is requested to use the highest possible speed value within the range which is compatible with the available bandwidth. As long as the server can maintain a speed value within the range it shall not change the media quality, but instead modify the actual delivery rate in response to available bandwidth and reflect this in the Speed value in the response. However, if this is not possible, the server should instead modify the media quality to respect the lowest speed value and the available bandwidth.

This functionality enables the local scaling implementation to use a tight range, or even a range where the lower bound equals the upper bound, to identify that it requires the server to deliver the requested amount of media time per delivery time independent of how much it needs to adapt the media quality to fit within the available path bandwidth. For buffer filling, it is suitable to use a range with a reasonable span and with a lower bound at the nominal media rate 1.0, such as 1.0 - 2.5. If the client wants to reduce the buffer, it can specify an upper bound that is below 1.0 to force the server to deliver slower than the nominal media rate.

2.6. Session Maintenance and Termination

The session context that has been established is kept alive by having the client show liveness. This is done in two main ways:

- o Media transport protocol keep-alive. RTP Control Protocol (RTCP) may be used when using RTP.
- o Any RTSP request referencing the session context.

Section 10.5 discusses the methods for showing liveness in more depth. If the client fails to show liveness for more than the established session timeout value (normally 60 seconds), the server may terminate the context. Other values may be selected by the server through the inclusion of the timeout parameter in the session header.

The session context is normally terminated by the client sending a TEARDOWN request (Section 13.7) to the server referencing the aggregated control URI. An individual mediaresource can be removed from a session context by a TEARDOWN request referencing that particular media resource. If all media resources are removed from a session context, the session context is terminated.

A client may keep the session alive indefinitely if allowed by the server; however, a client is recommended to release the session context when an extended period of time without media delivery activity has passed. The client can re-establish the session context if required later. What constitutes an extended period of time is dependent on the client, server and their usage. It is recommended that the client terminates the session before ten times the session timeout value has passed. A server may terminate the session after one session timeout period without any client activity beyond keep-alive. When a server terminates the session context, it does that by sending a TEARDOWN request indicating the reason.

A server can also request that the client tear down the session and re-establish it at an alternative server, as may be needed for maintenance. This is done by using the REDIRECT method (Section 13.10). The Terminate-Reason header (Section 18.52) is used to indicate when and why. The Location header indicates where it should connect if there is an alternative server available. When the deadline expires, the server simply stops providing the service. To achieve a clean closure, the client needs to initiate session termination prior to the deadline. In case the server has no other server to redirect to, and wants to close the session for maintenance, it shall use the TEARDOWN method with a Terminate-Reason header.

2.7. Extending RTSP

RTSP is quite a versatile protocol which supports extensions in many different directions. Even this core specification contains several blocks of functionality that are optional to implement. The use case and need for the protocol deployment should determine what parts are implemented. Allowing for extensions makes it possible for RTSP to reach out to additional use cases. However, extensions will affect the interoperability of the protocol and therefore it is important that they can be added in a structured way.

The client can learn the capability of a server by using the OPTIONS method (Section 13.1) and the Supported header (Section 18.51). It can also try and possibly fail using new methods, or require that particular features are supported using the Require (Section 18.43) or Proxy-Require (Section 18.37) header.

The RTSP protocol in itself can be extended in three ways, listed here in increasing order of the magnitude of changes supported:

- o Existing methods can be extended with new parameters, for example, headers, as long as these parameters can be safely ignored by the recipient. If the client needs negative acknowledgment when a method extension is not supported, a tag corresponding to the extension may be added in the field of the Require or Proxy-Require headers.
- o New methods can be added. If the recipient of the message does not understand the request, it must respond with error code 501 (Not Implemented) so that the sender can avoid using this method again. A client may also use the OPTIONS method to inquire about methods supported by the server. The server must list the methods it supports using the Public response-header.
- o A new version of the protocol can be defined, allowing almost all aspects (except the position of the protocol version number) to change. A new version of the protocol must be registered through an IETF standards track document.

The basic capability discovery mechanism can be used to both discover support for a certain feature and to ensure that a feature is available when performing a request. For a detailed explanation of this see Section 11.

New media delivery protocols may be added and negotiated at session establishment, in addition to extensions to the core protocol. Certain types of protocol manipulations can be done through parameter formats using SET_PARAMETER and GET_PARAMETER.

3. Document Conventions

3.1. Notational Conventions

Since a few of the definitions are identical to HTTP/1.1, this specification only points to the section where they are defined rather than copying it. For brevity, [HX.Y] is to be taken to refer to Section X.Y of the HTTP/1.1 specification ([RFC2616]).

All the mechanisms specified in this document are described in both prose and the Augmented Backus-Naur form (ABNF) described in detail in [RFC5234].

Indented paragraphs are used to provide informative background and motivation. This is intended to give readers who were not involved with the formulation of the specification an understanding of why things are the way they are in RTSP.

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

The word, "unspecified" is used to indicate functionality or features that are not defined in this specification. Such functionality cannot be used in a standardized manner without further definition in an extension specification to RTSP.

3.2. Terminology

Aggregate control: The concept of controlling multiple streams using a single timeline, generally maintained by the server. A client, for example, uses aggregate control when it issues a single play or pause message to simultaneously control both the audio and video in a movie. A session which is under aggregate control is referred to as an aggregated session.

Aggregate control URI: The URI used in an RTSP request to refer to and control an aggregated session. It normally, but not always, corresponds to the presentation URI specified in the session description. See Section 13.3 for more information.

Client: The client requests media service from the media server.

Connection: A transport layer virtual circuit established between two programs for the purpose of communication.

Container file: A file which may contain multiple media streams which often constitutes a presentation when played together. The concept of a container file is not embedded in the protocol. However, RTSP servers may offer aggregate control on the media streams within these files.

Continuous media: Data where there is a timing relationship between source and sink; that is, the sink needs to reproduce the timing relationship that existed at the source. The most common examples of continuous media are audio and motion video. Continuous media can be real-time (interactive or conversational), where there is a "tight" timing relationship between source and sink, or streaming where the relationship is less strict.

Feature-tag: A tag representing a certain set of functionality, i.e., a feature.

IRI: Internationalized Resource Identifier, is similar to a URI, but allows characters from the whole Universal Character Set (Unicode/ISO 10646), rather than the US-ASCII only. See [RFC3987] for more information.

Live: Normally used to describe a presentation or session with media coming from an ongoing event. This generally results in the session having an unbound or only loosely defined duration, and sometimes no seek operations are possible.

Media initialization: Datatype/codecs specific initialization. This includes such things as clock rates, color tables, etc. Any transport-independent information which is required by a client for playback of a media stream occurs in the media initialization phase of stream setup.

Media parameter: Parameter specific to a media type that may be changed before or during stream delivery.

Media server: The server providing media delivery services for one or more media streams. Different media streams within a presentation may originate from different media servers. A media server may reside on the same host or on a different host from which the presentation is invoked.

(Media) stream: A single media instance, e.g., an audio stream or a video stream as well as a single whiteboard or shared application group. When using RTP, a stream consists of all RTP and RTCP packets created by a source within an RTP session.

Message: The basic unit of RTSP communication, consisting of a structured sequence of octets matching the syntax defined in Section 20 and transmitted over a connection-based transport. A message is either a Request or a Response.

Message Body: The information transferred as the payload of a message (Request or response). A message body consists of meta-information in the form of message-body headers and content in the form of a message-body, as described in Section 9.

Non-Aggregated Control: Control of a single media stream.

Presentation: A set of one or more streams presented to the client as a complete media feed and described by a presentation description as defined below. Presentations with more than one media stream are often handled in RTSP under aggregate control.

Presentation description: A presentation description contains information about one or more media streams within a presentation, such as the set of encodings, network addresses and information about the content. Other IETF protocols such as SDP ([RFC4566]) use the term "session" for a presentation. The presentation description may take several different formats, including but not limited to the session description protocol format, SDP.

Response: An RTSP response to a Request. One type of RTSP message. If an HTTP response is meant, it is indicated explicitly.

Request: An RTSP request. One type of RTSP message. If an HTTP request is meant, it is indicated explicitly.

Request-URI: The URI used in a request to indicate the resource on which the request is to be performed.

RTSP agent: Refers to either an RTSP client, an RTSP server, or an RTSP proxy. In this specification, there are many capabilities that are common to these three entities such as the capability to send requests or receive responses. This term will be used when describing functionality that is applicable to all three of these entities.

RTSP session: A stateful abstraction upon which the main control methods of RTSP operate. An RTSP session is a common context; it is created and maintained on client's request and can be destroyed by either the client or server. It is established by an RTSP server upon the completion of a successful SETUP request (when a 200 OK response is sent) and is labeled with a session identifier at that time. The session exists until timed out by the server or

explicitly removed by a TEARDOWN request. An RTSP session is a stateful entity; an RTSP server maintains an explicit session state machine (see Appendix B) where most state transitions are triggered by client requests. The existence of a session implies the existence of state about the session's media streams and their respective transport mechanisms. A given session can have one or more media streams associated with it. An RTSP server uses the session to aggregate control over multiple media streams.

Origin Server: The server on which a given resource resides.

Transport initialization: The negotiation of transport information (e.g., port numbers, transport protocols) between the client and the server.

URI: Universal Resource Identifier, see [RFC3986]. The URIs used in RTSP are generally URLs as they give a location for the resource. As URLs are a subset of URIs, they will be referred to as URIs to cover also the cases when an RTSP URI would not be an URL.

URL: Universal Resource Locator, is a URI which identifies the resource through its primary access mechanism, rather than identifying the resource by name or by some other attribute(s) of that resource.

4. Protocol Parameters

4.1. RTSP Version

This specification defines version 2.0 of RTSP.

RTSP uses a "<major>.<minor>" numbering scheme to indicate versions of the protocol. The protocol versioning policy is intended to allow the sender to indicate the format of a message and its capacity for understanding further RTSP communication, rather than the features obtained via that communication. No change is made to the version number for the addition of message components which do not affect communication behavior or which only add to extensible field values.

The <minor> number is incremented when the changes made to the protocol add features which do not change the general message parsing algorithm, but which may add to the message semantics and imply additional capabilities of the sender. The <major> number is incremented when the format of a message within the protocol is changed. The version of an RTSP message is indicated by an RTSP-Version field in the first line of the message. Note that the major and minor numbers MUST be treated as separate integers and that each MAY be incremented higher than a single digit. Thus, RTSP/2.4 is a

lower version than RTSP/2.13, which in turn is lower than RTSP/12.3. Leading zeros SHALL NOT be sent and MUST be ignored by recipients.

4.2. RTSP IRI and URI

RTSP 2.0 defines and registers or updates three URI schemes "rtsp", "rtsp" and "rtspu". The usage of the last, "rtspu", is unspecified in RTSP 2.0, and is defined here to register the URI scheme that was defined in RTSP 1.0. The "rtspu" scheme indicates unspecified transport of the RTSP messages over unreliable transport (UDP in RTSP 1.0). An RTSP server MUST respond with an error code indicating the "rtspu" scheme is not implemented (501) to a request that carries a "rtspu" URI scheme.

The details of the syntax of "rtsp" and "rtsp" URIs has been changed from RTSP 1.0. These changes are:

- o Support for IPV6 literal in host part and future IP literals through RFC 3986 defined mechanism.
- o A new relative format to use in the RTSP protocol elements that is not required to start with "/".

Neither should have any significant impact on interoperability. If one is required to use IPV6 literals to reach an RTSP server, then that RTSP server must be IPV6 capable, and RTSP 1.0 is not a fully IPV6 capable protocol. If an RTSP 1.0 client attempts to process the URI it will not match the allowed syntax and be considered invalid and processing will be stopped. This is clearly a failure to reach the resource, however it is not a signification issue as RTSP 2.0 support was needed anyway in both server and client. Thus failure will only occur in a later step when there is a RTSP version mismatch between client and server. The second change will only occur inside RTSP message headers, as the request URI must be an absolute URI. Thus such usages will only occur after an agent has accepted and started processing RTSP 2.0 messages, and an RTSP 1.0 only agent will not be required to parse such types of relative URIs.

This specification also defines the format of the RTSP IRI [RFC3987] that can be used as RTSP resource identifiers and locators, in web pages, user interfaces, on paper, etc. However, the RTSP request message format only allows usage of the absolute URI format. The RTSP IRI format MUST use the rules and transformation for IRIs to URIs, as defined in [RFC3987]. This allows a URI that matches the RTSP 2.0 specification, and so is suitable for use in a request, to be created from an RTSP IRI.

The RTSP IRI and URI are both syntax restricted compared to the generic syntax defined in [RFC3986] and [RFC3987]:

- o An absolute URI requires the authority part; i.e., a host identity MUST be provided.
- o Parameters in the path element are prefixed with the reserved separator ";".

The "scheme" and "host" parts of all URIs [RFC3986] and IRIs [RFC3987] are case-insensitive. All other parts of RTSP URIs and IRIs are case-sensitive, and MUST NOT be case-mapped.

The fragment identifier is used as defined in sections 3.5 and 4.3 of [RFC3986], i.e., the fragment is to be stripped from the IRI by the requester and not included in the request URI. The user agent needs to interpret the value of the fragment based on the media type the request relates to; i.e., the media type indicated in Content-Type header in the response to DESCRIBE.

The syntax of any URI query string is unspecified and responder (usually the server) specific. The query is, from the requester's perspective, an opaque string and needs to be handled as such. Please note that relative URI with queries are difficult to handle due to the RFC 3986 relative URI handling rules. Any change of the path element using a relative URI results in the stripping of the query, which means the relative part needs to contain the query.

The URI scheme "rtsp" requires that commands are issued via a reliable protocol (within the Internet, TCP), while the scheme "rtsp" identifies a reliable transport using secure transport (TLS [RFC5246], see (Section 19).

For the scheme "rtsp", if no port number is provided in the authority part of the URI, the port number 554 MUST be used. For the scheme "rtsp", if no port number is provided in the authority part of the URI port number, the TCP port 322 MUST be used.

A presentation or a stream is identified by a textual media identifier, using the character set and escape conventions of URIs [RFC3986]. URIs may refer to a stream or an aggregate of streams; i.e., a presentation. Accordingly, requests described in (Section 13) can apply to either the whole presentation or an individual stream within the presentation. Note that some request methods can only be applied to streams, not presentations, and vice versa.

For example, the RTSP URI:

```
rtsp://media.example.com:554/twister/audiotrack
```

may identify the audio stream within the presentation "twister", which can be controlled via RTSP requests issued over a TCP connection to port 554 of host media.example.com.

Also, the RTSP URI:

```
rtsp://media.example.com:554/twister
```

identifies the presentation "twister", which may be composed of audio and video streams, but could also be something else like a random media redirector.

This does not imply a standard way to reference streams in URIs. The presentation description defines the hierarchical relationships in the presentation and the URIs for the individual streams. A presentation description may name a stream "a.mov" and the whole presentation "b.mov".

The path components of the RTSP URI are opaque to the client and do not imply any particular file system structure for the server.

This decoupling also allows presentation descriptions to be used with non-RTSP media control protocols simply by replacing the scheme in the URI.

4.3. Session Identifiers

Session identifiers are strings of length 8-128 characters. A session identifier MUST be generated cryptographically random (see [RFC4086]). It is RECOMMENDED that it contains 128 bits of entropy, i.e., approximately 22 characters from a high quality generator (see Section 21). However, note that the session identifier does not provide any security against session hijacking unless it is kept confidential by the client, server and trusted proxies.

4.4. Media Time Formats

RTSP currently supports three different media time formats defined below. Additional time formats may be specified in the future. These time formats can be used with the Range header (Section 18.40) to request playback and specify at which media position protocol requests actually will or have taken place. They are also used in description of the media's properties using the Media-Range header (Section 18.30). The unqualified format identifier is used on its own in Accept-Ranges header (Section 18.5) to declare supported time

formats and also in the Range header (Section 18.40) to request the time format used in the response.

4.4.1. SMPTE Relative Timestamps

A Society of Motion Picture and Television Engineers (SMPTE) relative timestamp expresses time relative to the start of the clip. Relative timestamps are expressed as SMPTE time codes [SMPTE_TC] for frame-level access accuracy. The time code has the format

hours:minutes:seconds:frames.subframes,

with the origin at the start of the clip. The default SMPTE format is "SMPTE 30 drop" format, with frame rate is 29.97 frames per second. Other SMPTE codes MAY be supported (such as "SMPTE 25") through the use of "smpte-type". For SMPTE 30, the "frames" field in the time value can assume the values 0 through 29. The difference between 30 and 29.97 frames per second is handled by dropping the first two frame indices (values 00 and 01) of every minute, except every tenth minute. If the frame and the subframe values are zero, they may be omitted. Subframes are measured in one-hundredth of a frame.

Examples:

```
smpte=10:12:33:20-  
smpte=10:07:33-  
smpte=10:07:00-10:07:33:05.01  
smpte-25=10:07:00-10:07:33:05.01
```

4.4.2. Normal Play Time

Normal play time (NPT) indicates the stream absolute position relative to the beginning of the presentation. The timestamp consists of two parts: the mandatory first part may be expressed in either seconds or hours, minutes, and seconds. The optional second part consists of a decimal point and decimal figures and indicates fractions of a second.

The beginning of a presentation corresponds to 0.0 seconds. Negative values are not defined.

The special constant "now" is defined as the current instant of a live event. It MAY only be used for live events, and MUST NOT be used for on-demand (i.e., non-live) content.

NPT is defined as in DSM-CC [ISO.13818-6.1995]: "Intuitively, NPT is the clock the viewer associates with a program. It is often

digitally displayed on a VCR. NPT advances normally when in normal play mode (scale = 1), advances at a faster rate when in fast scan forward (high positive scale ratio), decrements when in scan reverse (negative scale ratio) and is fixed in pause mode. NPT is (logically) equivalent to SMPTE time codes."

Examples:

```
npt=123.45-125
npt=12:05:35.3-
npt=now-
```

The syntax is based on ISO 8601 [ISO.8601.2000] and expresses the time elapsed since presentation start, with two different notations allowed:

- o The npt-hhmmss notation uses an ISO 8601 extended complete representation of the time of the day format (Section 5.3.1.1 of [ISO.8601.2000]) using colon (":") as separators between hours, minutes and seconds (hh:mm:ss). The hour counter is not limited to 0-24 hours; up to nineteen (19) digits of hours are allowed.
- o In accordance with the requirements of the ISO 8601 time format, the hours, minutes, and seconds MUST all be present, with two digits used for minutes and for seconds, and with a least two digits for hours. An NPT of 7 minutes and 0 seconds is represented as "00:07:00", and an NPT of 392 hours, 0 minutes, and 6 seconds is represented as "392:00:06".
- o RTSP 1.0 allowed NPT in the npt-hhmmss notation without any leading zeros, to ensure that implementations doesn't fail if any implementation follows the RTSP 1.0 format, all implementations are REQUIRED to support receiving NPT values, hours, minutes or seconds, without leading zeros.
- o The npt-sec notation expresses the time in seconds, using between one and nineteen (19) digits.

Both notations allow decimal fractions of seconds as specified in Section 5.3.1.3 of [ISO.8601.2000], using at most 9 digits, and allowing only "." (full stop) as the decimal separator.

The npt-sec notation is optimized for automatic generation, the npt-hhmmss notation for consumption by human readers. The "now" constant allows clients to request to receive the live feed rather than the stored or time-delayed version. This is needed since neither absolute time nor zero time are appropriate for this case.

4.4.3. Absolute Time

Absolute time is expressed following a specific types of ISO 8601 [ISO.8601.2000] based timestamps. The date is complete representation calendar date in basic format (YYYYMMDD) without separators (per Section 5.2.1.1 of [ISO.8601.2000]). The time of day is provided in the complete representation basic format (hhmmss) as specified in Section 5.3.1.1 of [ISO.8601.2000], allowing decimal fractions of seconds following Section 5.3.1.3 requiring "." (full stop) as decimal separator and limiting the number of digits to no more than nine (9). The time expressed MUST be using UTC (GMT), i.e. no timezone offsets allowed. The full date and time specification is the eight digit date followed by a "T" followed by the six digits time value, optionally followed by a full stop followed by one to nine fractions of a second and ended by "Z", e.g. YYYYMMDDThhmmss.ssZ.

The reason for this time format rather than using "Date and Time on the Internet: Timestamps" [RFC3339] are historic and using the format specified in RTSP 1.0. The motivations raised in RFC 3339 applies to why a selection from ISO 8601 was done, but a different and even more restrictive selection was applied in this case.

Example for clock format range request for a starting time of November 8, 1996 at 14h 37 min and 20 and a quarter seconds UTC playing for 10 min and 5 seconds, a Media-Properties header's "Time-Limited UTC property for 24th of December 2014 at 15 hours and 00 mins, and a Terminate-Readon headers "time" property for 18th of June 2013 at 16 hours, 12 minutes and 56 seconds:

```
clock=19961108T143720.25Z-19961108T144725.25Z
Time-Limited=20141224T1500Z
time=20130618T161256Z
```

4.5. Feature-Tags

Feature-tags are unique identifiers used to designate features in RTSP. These tags are used in Require (Section 18.43), Proxy-Require (Section 18.37), Proxy-Supported (Section 18.38), Supported (Section 18.51) and Unsupported (Section 18.55) header fields.

A feature-tag definition MUST indicate which combination of clients, servers or proxies it applies to.

The creator of a new RTSP feature-tag should either prefix the feature-tag with a reverse domain name (e.g., "com.example.mynewfeature" is an apt name for a feature whose inventor can be reached at "example.com"), or register the new

feature-tag with the Internet Assigned Numbers Authority (IANA) (see IANA Section 22).

The usage of feature-tags is further described in Section 11 that deals with capability handling.

4.6. Message Body Tags

Message body tags are opaque strings that are used to compare two message bodies from the same resource, for example in caches or to optimize setup after a redirect. Message body tags can be carried in the MTag header (see Section 18.31) or in SDP (see Appendix D.1.9). MTag is similar to ETag in HTTP/1.1 (see Section 3.11 of [RFC2068]).

A message body tag MUST be unique across all versions of all message bodies associated with a particular resource. A given message body tag value MAY be used for message bodies obtained by requests on different URIs. The use of the same message body tag value in conjunction with message bodies obtained by requests on different URIs does not imply the equivalence of those message bodies

Message body tags are used in RTSP to make some methods conditional. The methods are made conditional through the inclusion of headers; see "If-Match" (Section 18.24) and "If-None-Match" (Section 18.26). Note that RTSP message body tags apply to the complete presentation; i.e., both the presentation description and the individual media streams. Thus message body tags can be used to verify at setup time after a redirect that the same session description applies to the media at the new location using the If-Match header.

4.7. Media Properties

When an RTSP server handles media, it is important to consider the different properties a media instance for delivery and playback can have. This specification considers the media properties listed below in its protocol operations. They are derived from the differences between a number of supported usages.

On-demand: Media that has a fixed (given) duration that doesn't change during the life time of the RTSP session and is known at the time of the creation of the session. It is expected that the content of the media will not change, even if the representation, i.e., encoding, quality, etc, may change. Generally one can seek, i.e., request any range, within the media.

Dynamic On-demand: This is a variation of the on-demand case where external methods are used to manipulate the actual content of the

media setup for the RTSP session. The main example is a content defined by a playlist.

Live: Live media represents a progressing content stream (such as broadcast TV) where the duration may or may not be known. It is not seekable, only the content presently being delivered can be accessed.

Live with Recording: A Live stream that is combined with a server-side capability to store and retain the content of the live session, and allow for random access delivery within the part of the already recorded content. The actual behavior of the media stream is very much dependent on the retention policy for the media stream; either the server will be able to capture the complete media stream, or it will have a limitation in how much will be retained. The media range will dynamically change as the session progress. For servers with a limited amount of storage available for recording, there will typically be a sliding window that moves forward while new data is made available and older data is discarded.

To cover the above usages, the following media properties with appropriate values are specified:

4.7.1. Random Access and Seeking

Random Access is the ability to specify and get media delivered starting from any time instant within the content, an operation called seeking. The Media-Properties header will indicate the general capability for a media resource to perform random access:

Random-Access: The media is seekable to any out of a large number of points within the media. Due to media encoding limitations, a particular point may not be reachable, but seeking to a point close by is enabled. A floating point number of seconds may be provided to express the worst case distance between random access points.

Beginning-Only: Seeking is only possible to the beginning of the content.

No-seeking: Seeking is not possible at all.

If random access is possible, as indicated by the Media-Properties header, the actual behavior policy when seeking can be controlled using the Seek-Style header (Section 18.47).

4.7.2. Retention

Media may have different retention policies in place that affect the operation on media. The following different media retention policies are defined:

Unlimited: The media will not be removed as long as the RTSP session is in existence.

Time-Limited: The media will not be removed before the given wallclock time. After that time it may or may not be available any more.

Time-Duration: The media (on fragment or unit basis) will be retained for the specified duration.

4.7.3. Content Modifications

There is also the question of how the content may change over time for a given media resource:

Immutable: The content of the media will not change, even if the representation, i.e., encoding, quality, etc., may change.

Dynamic: The content can change due to external methods or triggers, such as playlists, but this will be announced by explicit updates.

Time-Progressing: As time progresses new content will become available. If the content also is retained it will become longer as everything between the start point and the point currently being made available can be accessed. If the media server uses a sliding window policy for retention, the start point will also change as time progresses.

4.7.4. Supported Scale Factors

Content often supports only a limited set or range of scales when delivering the media. To enable the client to know what values or ranges of scale operations that the whole content or the current position supports, a media properties attribute for this is defined which contains a list with the values and/or ranges that are supported. The attribute is named "Scales". The "Scales" attribute may be updated at any point in the content due to content consisting of spliced pieces or content being dynamically updated by out-of-band mechanisms.

4.7.5. Mapping to the Attributes

This section shows examples of how one would map the above usages to the properties and their values.

Example of On-demand:

Random Access: Random-Access=5.0, Content Modifications:
Immutable, Retention: Unlimited or Time-Limited.

Example of Dynamic On-demand:

Random Access: Random-Access=3.0, Content Modifications: Dynamic,
Retention: Unlimited or Time-Limited.

Example of Live:

Random Access: No-seeking, Content Modifications: Time-
Progressing, Retention: Time-Duration=0.0

Example of Live with Recording:

Random Access: Random-Access=3.0, Content Modifications: Time-
Progressing, Retention: Time-Duration=7200.0

5. RTSP Message

RTSP is a text-based protocol and uses the ISO 10646 character set in UTF-8 encoding RFC 3629 [RFC3629]. Lines MUST be terminated by CRLF.

Text-based protocols make it easier to add optional parameters in a self-describing manner. Since the number of parameters and the frequency of commands is low, processing efficiency is not a concern. Text-based protocols, if done carefully, also allow easy implementation of research prototypes in scripting languages such as TCL, Visual Basic and Perl.

The ISO 10646 character set avoids character set switching, but is invisible to the application as long as US-ASCII is being used. This is also the encoding used for RTCP [RFC3550].

A request contains a method, the object the method is operating upon, and parameters to further describe the method. Methods are idempotent unless otherwise noted. Methods are also designed to require little or no state maintenance at the media server.

5.1. Message Types

RTSP messages are either requests from client to server, or server to client, and responses in the reverse direction. Request (Section 7) and Response (Section 8) messages use a format based on the generic message format of RFC 5322 [RFC5322] for transferring bodies (the

payload of the message). Both types of messages consist of a start-line, zero or more header fields (also known as "headers"), an empty line (i.e., a line with nothing preceding the CRLF) indicating the end of the headers, and possibly the data of the message body. The below ABNF [RFC5234] is for information and the formal message specification is present in Section 20.2.2.

```
generic-message = start-line
                  *(rtsp-header CRLF)
                  CRLF
                  [ message-body-data ]
start-line = Request-Line | Status-Line
```

In the interest of robustness, agents MUST ignore any empty line(s) received where a Request-Line or Status-Line is expected. In other words, if the agent is reading the protocol stream at the beginning of a message and receives any number of CRLFs first, it MUST ignore any of the CRLFs.

5.2. Message Headers

RTSP header fields (see Section 18) include general-header, request-header, response-header, and message-body header fields.

The order in which header fields with differing field names are received is not significant. However, it is "good practice" to send general-header fields first, followed by request-header or response-header fields, and ending with the Message-body header fields.

Multiple header fields with the same field-name MAY be present in a message if and only if the entire field-value for that header field is defined as a comma-separated list. It MUST be possible to combine the multiple header fields into one "field-name: field-value" pair, without changing the semantics of the message, by appending each subsequent field-value to the first, each separated by a comma. The order in which header fields with the same field-name are received is therefore significant to the interpretation of the combined field value, and thus a proxy MUST NOT change the order of these field values when a message is forwarded.

Unknown message headers MUST be ignored (skipping over the header to the next protocol element, and not causing an error) by a RTSP server or client. An RTSP Proxy MUST forward unknown message headers. Message headers defined outside of this specification that are required to be interpreted by the RTSP agent will need to use feature tags (Section 4.5) and include them in the appropriate Require (Section 18.43) or Proxy-Require (Section 18.37) header.

5.3. Message Body

The message body (if any) of an RTSP message is used to carry further information for a particular resource associated with the request or response. An example of a message body is a Session Description Protocol (SDP) message.

The presence of a message body in either a request or a response **MUST** be signaled by the inclusion of a Content-Length header (see Section 18.17) and Content-Type (see Section 18.19). A message body **MUST NOT** be included in a request or response if the specification of the particular method (see Method Definitions (Section 13)) does not allow sending a message body. In case a message body is received in a message when not expected the message body data **SHOULD** be discarded. This is to allow future extensions to define optional use of a message body.

5.4. Message Length

An RTSP Message that does not contain any message body is terminated by the first empty line after the header fields (Note: An empty line is a line with nothing preceding the CRLF.). In RTSP messages that contain message bodies the empty line is followed by the message body. The length of that body is determined by the value of the Content-Length header (Section 18.17). The value in the header represents the length of the message-body in octets. If this header field is not present, a value of zero is assumed, i.e., no message body present in the message. Unlike an HTTP message, an RTSP message **MUST** contain a Content-Length header whenever it contains a message body. Note that RTSP does not support the HTTP/1.1 "chunked" transfer coding (see [H3.6.1]).

Given the moderate length of presentation descriptions returned, the server should always be able to determine its length, even if it is generated dynamically, making the chunked transfer encoding unnecessary.

6. General Header Fields

General headers are headers that may be used in both requests and responses. The general-headers are listed in Table 1:

Header Name	Defined in Section
Accept-Ranges	Section 18.5
Cache-Control	Section 18.11
Connection	Section 18.12
CSeq	Section 18.20
Date	Section 18.21
Media-Properties	Section 18.29
Media-Range	Section 18.30
Pipelined-Requests	Section 18.33
Proxy-Supported	Section 18.38
Range	Section 18.40
RTP-Info	Section 18.45
Scale	Section 18.46
Seek-Style	Section 18.47
Server	Section 18.48
Session	Section 18.49
Speed	Section 18.50
Supported	Section 18.51
Timestamp	Section 18.53
Transport	Section 18.54
User-Agent	Section 18.56
Via	Section 18.57

Table 1: The general headers used in RTSP

7. Request

A request message uses the format outlined below regardless of the direction of a request, client to server or server to client:

- o Request line, containing the method to be applied to the resource, the identifier of the resource, and the protocol version in use;
- o Zero or more Header lines, that can be of the following types: general-headers (Section 6), request-headers (Section 7.2), or message body headers (Section 9.1);
- o One empty line (CRLF) to indicate the end of the header section;
- o Optionally a message-body, consisting of one or more lines. The length of the message body in octets is indicated by the Content-Length message header.

7.1. Request Line

The request line provides the key information about the request: what method, on what resources and using which RTSP version. The methods that are defined by this specification are listed in Table 2.

Method	Defined in Section
DESCRIBE	Section 13.2
GET_PARAMETER	Section 13.8
OPTIONS	Section 13.1
PAUSE	Section 13.6
PLAY	Section 13.4
PLAY_NOTIFY	Section 13.5
REDIRECT	Section 13.10
SETUP	Section 13.3
SET_PARAMETER	Section 13.9
TEARDOWN	Section 13.7

Table 2: The RTSP Methods

The syntax of the RTSP request line is the following:

```
<Method> SP <Request-URI> SP <RTSP-Version> CRLF
```

Note: This syntax cannot be freely changed in future versions of RTSP. This line needs to remain parsable by older RTSP implementations since it indicates the RTSP version of the message.

In contrast to HTTP/1.1 [RFC2616], RTSP requests identify the resource through an absolute RTSP URI (including scheme, host, and port) (see Section 4.2) rather than just the absolute path.

HTTP/1.1 requires servers to understand the absolute URI, but clients are supposed to use the Host request-header. This is purely needed for backward-compatibility with HTTP/1.0 servers, a consideration that does not apply to RTSP.

An asterisk "*" can be used instead of an absolute URI in the Request-URI part to indicate that the request does not apply to a particular resource, but to the server or proxy itself, and is only allowed when the request method does not necessarily apply to a resource.

For example:

```
OPTIONS * RTSP/2.0
```

An OPTIONS in this form will determine the capabilities of the server or the proxy that first receives the request. If the capability of the specific server needs to be determined, without regard to the capability of an intervening proxy, the server should be addressed explicitly with an absolute URI that contains the server's address.

For example:

```
OPTIONS rtsp://example.com RTSP/2.0
```

7.2. Request Header Fields

The RTSP headers in Table 3 can be included in a request, as request-headers, to modify the specifics of the request.

Header	Defined in Section
Accept	Section 18.1
Accept-Credentials	Section 18.2
Accept-Encoding	Section 18.3
Accept-Language	Section 18.4
Authorization	Section 18.8
Bandwidth	Section 18.9
Blocksize	Section 18.10
From	Section 18.23
If-Match	Section 18.24
If-Modified-Since	Section 18.25
If-None-Match	Section 18.26
Notify-Reason	Section 18.32
Proxy-Authorization	Section 18.36
Proxy-Require	Section 18.37
Referrer	Section 18.41
Request-Status	Section 18.42
Require	Section 18.43
Terminate-Reason	Section 18.52

Table 3: The RTSP request headers

Detailed header definitions are provided in Section 18.

New request-headers may be defined. If the receiver of the request is required to understand the request-header, the request **MUST** include a corresponding feature tag in a Require or Proxy-Require header to ensure the processing of the header.

8. Response

After receiving and interpreting a request message, the recipient responds with an RTSP response message. Normally, there is only one, final, response. Only responses using the response code class 1xx, are allowed to send one or more 1xx response messages prior to the final response message.

The valid response codes and the methods they can be used with are listed in Table 4.

8.1. Status-Line

The first line of a Response message is the Status-Line, consisting of the protocol version followed by a numeric status code and the textual phrase associated with the status code, with each element separated by SP characters. No CR or LF is allowed except in the final CRLF sequence.

<RTSP-Version> SP <Status-Code> SP <Reason-Phrase> CRLF

8.1.1. Status Code and Reason Phrase

The Status-Code element is a 3-digit integer result code of the attempt to understand and satisfy the request. These codes are fully defined in Section 17. The Reason-Phrase is intended to give a short textual description of the Status-Code. The Status-Code is intended for use by automata and the Reason-Phrase is intended for the human user. The client is not required to examine or display the Reason-Phrase.

The first digit of the Status-Code defines the class of response. The last two digits do not have any categorization role. There are 5 values for the first digit:

- 1xx: Informational - Request received, continuing process
- 2xx: Success - The action was successfully received, understood, and accepted
- 3xx: Redirection - Further action needs to be taken in order to complete the request (3xx rather than 3xx is used as 304 is excluded, see Section 17.3)
- 4xx: Client Error - The request contains bad syntax or cannot be fulfilled

5xx: Server Error - The server failed to fulfill an apparently valid request

The individual values of the numeric status codes defined for RTSP/2.0, and an example set of corresponding Reason-Phrases, are presented in Table 4. The reason phrases listed here are only recommended; they may be replaced by local equivalents without affecting the protocol. Note that RTSP adopts most HTTP/1.1 [RFC2616] status codes and adds RTSP-specific status codes starting at x50 to avoid conflicts with future HTTP status codes that are desirable to import into RTSP. All these codes are RTSP specific and RTSP has its own registry separate from HTTP for status codes.

RTSP status codes are extensible. RTSP applications are not required to understand the meaning of all registered status codes, though such understanding is obviously desirable. However, applications MUST understand the class of any status code, as indicated by the first digit, and treat any unrecognized response as being equivalent to the x00 status code of that class, with an exception for unknown 3xx codes, which MUST be treated as a 302 (Found). The reason being that no 300 (Multiple Choices in HTTP) is defined for RTSP. An response with an unrecognized status code MUST NOT be cached. For example, if an unrecognized status code of 431 is received by the client, it can safely assume that there was something wrong with its request and treat the response as if it had received a 400 status code. In such cases, user agents SHOULD present to the user the message body returned with the response, since that message body is likely to include human-readable information which will explain the unusual status.

Code	Reason	Method
100	Continue	all
200	OK	all
301	Moved Permanently	all
302	Found	all
303	reserved	n/a
304	Not Modified	all

305	Use Proxy	all
400	Bad Request	all
401	Unauthorized	all
402	Payment Required	all
403	Forbidden	all
404	Not Found	all
405	Method Not Allowed	all
406	Not Acceptable	all
407	Proxy Authentication Required	all
408	Request Timeout	all
410	Gone	all
412	Precondition Failed	DESCRIBE, SETUP
413	Request Message Body Too Large	all
414	Request-URI Too Long	all
415	Unsupported Media Type	all
451	Parameter Not Understood	SET_PARAMETER, GET_PARAMETER
452	reserved	n/a
453	Not Enough Bandwidth	SETUP
454	Session Not Found	all
455	Method Not Valid In This State	all
456	Header Field Not Valid	all
457	Invalid Range	PLAY, PAUSE

458	Parameter Is Read-Only	SET_PARAMETER
459	Aggregate Operation Not Allowed	all
460	Only Aggregate Operation Allowed	all
461	Unsupported Transport	all
462	Destination Unreachable	all
463	Destination Prohibited	SETUP
464	Data Transport Not Ready Yet	PLAY
465	Notification Reason Unknown	PLAY_NOTIFY
466	Key Management Error	all
470	Connection Authorization Required	all
471	Connection Credentials not accepted	all
472	Failure to establish secure connection	all
500	Internal Server Error	all
501	Not Implemented	all
502	Bad Gateway	all
503	Service Unavailable	all
504	Gateway Timeout	all
505	RTSP Version Not Supported	all
551	Option Not Supported	all
553	Proxy Unavailable	all

Table 4: Status codes and their usage with RTSP methods

8.2. Response Headers

The response-headers allow the request recipient to pass additional information about the response which cannot be placed in the Status-Line. This header gives information about the server and about further access to the resource identified by the Request-URI. All headers currently classified as response-headers are listed in Table 5.

Header	Defined in Section
Authentication-Info	Section 18.7
Connection-Credentials	Section 18.13
Location	Section 18.28
MTag	Section 18.31
Proxy-Authenticate	Section 18.34
Public	Section 18.39
Retry-After	Section 18.44
Unsupported	Section 18.55
WWW-Authenticate	Section 18.58

Table 5: The RTSP response headers

Response-header names can be extended reliably only in combination with a change in the protocol version. However, the usage of feature-tags in the request allows the responding party to learn the capability of the receiver of the response. A new or experimental header can be given the semantics of response-header if all parties in the communication recognize them to be a response-header. Unrecognized headers in responses MUST be ignored.

9. Message Body

Some Request and Response messages include a message body, if not otherwise restricted by the request method or response status code. The message body consists of the content data itself (see also Section 5.3).

The SET_PARAMETER and GET_PARAMETER requests and responses, and the DESCRIBE response as defined by this specification can have a message body; the purpose of the message body is defined in each case. All 4xx and 5xx responses MAY also have a message body to carry additional response information. Generally a message body MAY be attached to any RTSP 2.0 request or response, but the content of the message body MAY be ignored by the receiver. Extensions to this specification can specify the purpose and content of message bodies, including requiring their inclusion.

In this section, both sender and recipient refer to either the client or the server, depending on who sends and who receives the message body.

9.1. Message-Body Header Fields

Message-body header fields define meta-information about the content data in the message body. The message-body header fields are listed in Table 6.

Header	Defined in Section
Allow	Section 18.6
Content-Base	Section 18.14
Content-Encoding	Section 18.15
Content-Language	Section 18.16
Content-Length	Section 18.17
Content-Location	Section 18.18
Content-Type	Section 18.19
Expires	Section 18.22
Last-Modified	Section 18.27

Table 6: The RTSP message-body headers

The extension-header mechanism allows additional message-body header fields to be defined without changing the protocol, but these fields cannot be assumed to be recognizable by the recipient. Unrecognized

header fields MUST be ignored by the recipient and forwarded by proxies.

9.2. Message Body

An RTSP message with a message body MUST include the Content-Type and Content-Length headers. When a message body is included with a message, the data type of that content data is determined via the header fields Content-Type and Content-Encoding.

Content-Type specifies the media type of the underlying data. There is no default media format and the actual format used in the body is required to be explicitly stated in the Content-Type header. By being explicit and always require inclusion of the Content-Type header with accurate information one avoids the many pitfalls in heuristic based interpretation of the body content. These are issue HTTP and email have suffered from.

Content-Encoding may be used to indicate any additional content codings applied to the data, usually for the purpose of data compression, that are a property of the requested resource. The default encoding is 'identity', i.e. no transformation of the message body.

The Content-Length of a message is the length of the content, measured in octets.

9.3. Message Body Format Negotiation

The content format of the message body is provided using the Content-Type header (Section 18.19). To enable the responder of a request to determine which media type it should use, the requestor may include the Accept header (Section 18.1) in a request to identify supported media types or media type ranges suitable to the response. In case the responder is not supporting any of the specified formats, then the request response will be a 406 (Not Acceptable) error code.

The media types that may be used on requests with message bodies need to be determined through the use of feature-tags, specification requirement or trial and error. Trial and error works because when the responder does not support the media type of the message body it will respond with a 415 (Unsupported Media Type).

The formats supported and their negotiation is done individually on a per method and direction (request or response body) direction. Requirements on supporting particular media types for use as message bodies in requests and response SHALL also be specified on per method and direction basis.

10. Connections

RTSP Messages are transferred between RTSP agents and proxies using a transport connection. This transport connection uses TCP or TCP/TLS. This transport connection is referred to as the 'connection' or 'RTSP connection' within this document.

RTSP requests can be transmitted using the two different connection scenarios listed below:

- o persistent - a transport connection is used for several request/response transactions;
- o transient - a transport connection is used for each single request/response transaction.

RFC 2326 attempted to specify an optional mechanism for transmitting RTSP messages in connectionless mode over a transport protocol such as UDP. However, it was not specified in sufficient detail to allow for interoperable implementations. In an attempt to reduce complexity and scope, and due to lack of interest, RTSP 2.0 does not attempt to define a mechanism for supporting RTSP over UDP or other connectionless transport protocols. A side-effect of this is that RTSP requests **MUST NOT** be sent to multicast groups since no connection can be established with a specific receiver in multicast environments.

Certain RTSP headers, such as the CSeq header (Section 18.20), which may appear to be relevant only to connectionless transport scenarios are still retained and **MUST** be implemented according to the specification. In the case of CSeq, it is quite useful for matching responses to requests if the requests are pipelined (see Section 12). It is also useful in proxies for keeping track of the different requests when aggregating several client requests on a single TCP connection.

10.1. Reliability and Acknowledgements

Since RTSP messages are transmitted using reliable transport protocols, they **MUST NOT** be retransmitted at the RTSP protocol level. Instead, the implementation must rely on the underlying transport to provide reliability. The RTSP implementation may use any indication of reception acknowledgment of the message from the underlying transport protocols to optimize the RTSP behavior.

If both the underlying reliable transport such as TCP and the RTSP application retransmit requests, each packet loss or message loss may result in two retransmissions. The receiver typically cannot

take advantage of the application-layer retransmission since the transport stack will not deliver the application-layer retransmission before the first attempt has reached the receiver. If the packet loss is caused by congestion, multiple retransmissions at different layers will exacerbate the congestion.

Lack of acknowledgment of an RTSP request should be handled within the constraints of the connection timeout considerations described below (Section 10.4).

10.2. Using Connections

A TCP transport can be used for both persistent connections (for several message exchanges) and transient connections (for a single message exchange). Implementations of this specification **MUST** support RTSP over TCP. The scheme of the RTSP URI (Section 4.2) allows the client to specify the port it will contact the server on, and defines the default port to use if one is not explicitly given.

In addition to the registered default ports, i.e., 554 (rtsp) and 322 (rtsp), there is an alternative port 8554 registered. This port may provide some benefits over non-registered ports if a RTSP server is unable to use the default ports. The benefits may include pre-configured security policies as well as classifiers in network monitoring tools.

A RTSP client opening a TCP connection for accessing a particular resource as identified by a URI uses the IP address and port derived from the host and port parts of the URI. The IP address is either the explicit address provided in the URI or any of the addresses provided when performing A and AAAA record DNS lookups of the host name in the URI.

A server **MUST** handle both persistent and transient connections.

Transient connections facilitate mechanisms for fault tolerance. They also allow for application layer mobility. A server and client pair that support transient connections can survive the loss of a TCP connection; e.g., due to a NAT timeout. When the client has discovered that the TCP connection has been lost, it can set up a new one when there is need to communicate again.

A persistent connection is **RECOMMENDED** to be used for all transactions between the server and client, including messages for multiple RTSP sessions. However, a persistent connection **MAY** be closed after a few message exchanges. For example, a client may use a persistent connection for the initial SETUP and PLAY message

exchanges in a session and then close the connection. Later, when the client wishes to send a new request, such as a PAUSE for the session, a new connection would be opened. This connection may either be transient or persistent.

An RTSP agent MAY use one connection to handle multiple RTSP sessions on the same server. The RTSP agent SHALL NOT use more than one connection per RTSP session at any given point.

Having only one connection in use at any time avoids confusion on which connection any server to client requests shall be sent on. Using a single connection for multiple RTSP session also saves complexity by enabling the server to maintain less state about its connection resources on the server. Not using more than one connection at a time for a particular RTSP session avoids wasting connection resources and allows the server to track only the most recently used client to server connection for each RTSP session as being the currently valid server to client connection.

RTSP allows a server to send requests to a client. However, this can be supported only if a client establishes a persistent connection with the server. In cases where a persistent connection does not exist between a server and its client, due to the lack of a signaling channel the server may be forced to silently discard RTSP messages, and may even drop an RTSP session without notifying the client. An example of such a case is when the server desires to send a REDIRECT request for an RTSP session to the client but is not able to do so because it cannot reach the client. A server that attempts to send a request to a client that has no connection currently to the server SHALL discard the request.

Without a persistent connection between the client and the server, the media server has no reliable way of reaching the client. Because the likely failure of server to client established connections the server will not even attempt establishing any connection.

Queuing of server to client requests has been considered. However a security issue exists as to how it might be possible to authorize a client establishing a new connection as being a legitimate receiver of a request related to a particular RTSP session, without the client first issuing requests related to the pending request. Thus, it would be likely to make any such requests even more delayed and less useful.

The sending of client and server requests can be asynchronous events. To avoid deadlock situations both client and server MUST be able to send and receive requests simultaneously. As an RTSP response may be

queued up for transmission, reception or processing behind the peer RTSP agent's own requests, all RTSP agents are required to have a certain capability of handling outstanding messages. A potential issue is that outstanding requests may timeout despite them being processed by the peer due to the response being caught in the queue behind a number of requests that the RTSP agent is processing but that take some time to complete. To avoid this problem an RTSP agent is recommended to buffer incoming messages locally so that any response messages can be processed immediately upon reception. If responses are separated from requests and directly forwarded for processing, not only can the result be used immediately, the state associated with that outstanding request can also be released. However, buffering a number of requests on the receiving RTSP agent consumes resources and enables a resource exhaustion attack on the agent. Therefore this buffer should be limited so that an unreasonable number of requests or total message size is not allowed to consume the receiving agent's resources. In most APIs, having the receiving agent stop reading from the TCP socket will result in TCP's window being clamped. Thus forcing the buffering onto the sending agent when the load is larger than expected. However, as both RTSP message sizes and frequency may be changed in the future by protocol extensions, an agent should be careful against taking harsher measurements against a potential attack. When under attack an RTSP agent can close TCP connections and release state associated with that TCP connection.

To provide some guidance on what is reasonable the following guidelines are given. It is RECOMMENDED that:

- o an RTSP agent should not have more than 10 outstanding requests per RTSP session;
- o an RTSP agent should not have more than 10 outstanding requests that are not related to an RTSP session or that are requesting to create an RTSP session.

In light of the above, it is RECOMMENDED that clients use persistent connections whenever possible. A client that supports persistent connections MAY "pipeline" its requests (see Section 12).

RTSP Agents can send requests to multiple different destinations, either servers or client contexts over the same connection to a proxy. Then the proxy forks the message to the different destinations over proxy to agent connections. In these cases when multiple requests are outstanding the requesting agent MUST be ready to receive the responses out of order compared to the order they were sent on the connection. The order between multiple messages

for each destination will be maintained, however, the order between response from different destinations can be different.

The reason for this is to avoid a head-of-line blocking situation. In a sequence of requests an early outstanding request may take time to be processed at one destination. Simultaneously, a response from any other destination that was later in the sequence of requests, may have arrived at the proxy. Thus allowing out-of-order responses avoids forcing the proxy to buffer this response and instead deliver it as soon as possible. Note, this will not affect the order in which the messages sent to each separate destination were processed at request destination.

This scenario can occur in two cases involving proxies. The first is a client issuing requests for sessions on different servers using a common client to proxy connection. The second is for server to client requests, like REDIRECT being sent by the server over a common transport connection the proxy created for its different connecting clients.

10.3. Closing Connections

The client MAY close a connection at any point when no outstanding request/response transactions exist for any RTSP session being managed through the connection. The server, however, SHOULD NOT close a connection until all RTSP sessions being managed through the connection have been timed out (Section 18.49). A server SHOULD NOT close a connection immediately after responding to a session-level TEARDOWN request for the last RTSP session being controlled through the connection. Instead, the server should wait for a reasonable amount of time for the client to receive and act upon the TEARDOWN response, and initiate the connection closing. The server SHOULD wait at least 10 seconds after sending the TEARDOWN response before closing the connection.

This is to ensure that the client has time to issue a SETUP for a new session on the existing connection after having torn the last one down. 10 seconds should give the client ample opportunity to get its message to the server.

A server SHOULD NOT close the connection directly as a result of responding to a request with an error code.

Certain error responses such as "460 Only Aggregate Operation Allowed" (Section 17.4.25) are used for negotiating capabilities of a server with respect to content or other factors. In such cases, it is inefficient for the server to close a connection on an error response. Also, such behavior would prevent

implementation of advanced/special types of requests or result in extra overhead for the client when testing for new features. On the other hand, keeping connections open after sending an error response poses a Denial of Service security risk (Section 21).

The server MAY close a connection if it receives an incomplete message and if the message is not completed within a reasonable amount of time. It is RECOMMENDED that the server waits at least 10 seconds for the completion of a message or for the next part of the message to arrive (which is an indication that the transport and the client are still alive). Servers believing they are under attack or otherwise starved for resources during that event MAY consider using a shorter timeout.

If a server closes a connection while the client is attempting to send a new request, the client will have to close its current connection, establish a new connection and send its request over the new connection.

An RTSP message SHOULD NOT be terminated by closing the connection. Such a message MAY be considered to be incomplete by the receiver and discarded. An RTSP message is properly terminated as defined in Section 5.

10.4. Timing Out Connections and RTSP Messages

Receivers of a request (responder) SHOULD respond to requests in a timely manner even when a reliable transport such as TCP is used. Similarly, the sender of a request (requester) SHOULD wait for a sufficient time for a response before concluding that the responder will not be acting upon its request.

A responder SHOULD respond to all requests within 5 seconds. If the responder recognizes that processing of a request will take longer than 5 seconds, it SHOULD send a 100 (Continue) response as soon as possible. It SHOULD continue sending a 100 response every 5 seconds thereafter until it is ready to send the final response to the requester. After sending a 100 response, the responder MUST send a final response indicating the success or failure of the request.

A requester SHOULD wait at least 10 seconds for a response before concluding that the responder will not be responding to its request. After receiving a 100 response, the requester SHOULD continue waiting for further responses. If more than 10 seconds elapses without receiving any response, the requester MAY assume that the responder is unresponsive and abort the connection by closing the TCP connection.

In cases multiple RTSP sessions share the same transport connection, abandoning a request and closing the connection may have significant impact on those other sessions. First of all, other RTSP requests may have become queued up due to the request taking long time to process. Secondly also those sessions loose the possibility to receive server to client requests. To mitigate that situation the RTSP client or server SHOULD establish a new connection and send any queued up and non-responded requests on this new connection. Secondly, to ensure that the RTSP server knows which connection that is valid for a particular RTSP session, the RTSP agent SHOULD send a keep-alive request, if no other request will be sent immediately for that RTSP session, for each RTSP session on the old connection. The keep-alive request will normally be a SET_PARAMETER with a session header to inform the server that this agent cares about this RTSP session.

A requester SHOULD wait longer than 10 seconds for a response if it is experiencing significant transport delays on its connection to the responder. The requester is capable of determining the round trip time (RTT) of the request/response cycle using the Timestamp header (Section 18.53) in any RTSP request.

10 seconds was chosen for the following reasons. It gives TCP time to perform a couple of retransmissions, even if operating on default values. It is short enough that users may not abandon the process themselves. However, it should be noted that 10 seconds can be aggressive on certain type of networks. The 5 seconds value for lxx messages is half the timeout giving a reasonable chance of successful delivery before timeout happens on the requester side.

10.5. Showing Liveness

RTSP requires the client to periodically show its liveness to the server or the server may terminate any session state. Several different protocol mechanism includes in their usage a liveness proof from the client. These mechanisms are; RTSP requests with a Session header to the server; if RTP & RTCP is used for media data transport and the transport is established the RTCP message proves liveness; or through any other used media transport protocol capable of indicating liveness of the RTSP client. It is RECOMMENDED that a client does not wait to the last second of the timeout before trying to send a liveness message. The RTSP message may take some time to arrive safely at the receiver, due to packet loss and TCP retransmissions. To show liveness between RTSP requests being issued to accomplish other things, the following mechanisms can be used, in descending order of preference:

RTCP: If RTP is used for media transport RTCP SHOULD be used. If RTCP is used to report transport statistics, it will necessarily also function as a keep-alive. The server can determine the client by network address and port together with the fact that the client is reporting on the server's RTP sender sources (SSRCs). A downside of using RTCP is that it only gives statistical guarantees of reaching the server. However, the probability of a false client timeout is so low that it can be ignored in most cases. For example, assume a session with 60 seconds timeout and enough bitrate assigned to RTCP messages to send a message from client to server on average every 5 seconds. That client has, for a network with 5% packet loss, a probability of failing to confirm liveness within the timeout interval for that session of 2.4×10^{-16} . Sessions with shorter timeouts, or much higher packet loss, or small RTCP bandwidths SHOULD also implement one or more of the mechanisms below.

SET_PARAMETER: When using SET_PARAMETER for keep-alive, a body SHOULD NOT be included. This method is the RECOMMENDED RTSP method to use for a request intended only to perform keep-alive. Support of SET_PARAMETER is mandatory for RTSP Servers to ensure clients can use this method.

GET_PARAMETER: When using GET_PARAMETER for keep-alive, a body SHOULD NOT be included. Dependent on implementation support in server. Use OPTIONS method to determine if there are method support or simply try.

OPTIONS: This method is also usable, but it causes the server to perform more unnecessary processing and results in bigger responses than necessary for the task. The reason is that the server needs to determine the capabilities associated with the media resource to correctly populate the Public and Allow headers.

The timeout parameter of the Session header (Section 18.49) MAY be included in a SETUP response, and MUST NOT be included in requests. The server uses it to indicate to the client how long the server is prepared to wait between RTSP commands or other signs of life before closing the session due to lack of activity (see Appendix B). The timeout is measured in seconds, with a default of 60 seconds. The length of the session timeout MUST NOT be changed in an established session.

10.6. Use of IPv6

Explicit IPv6 [RFC2460] support was not present in RTSP 1.0 (RFC 2326). RTSP 2.0 has been updated for explicit IPv6 support. Implementations of RTSP 2.0 MUST understand literal IPv6 addresses in URIs and RTSP headers. Although the general URI format envisages potential future new versions of the literal IP address, usage of any such new version would require other modifications to the RTSP specification (e.g. address fields in the Transport header (Section 18.54)).

10.7. Overload Control

Overload in RTSP can occur when servers and proxies have insufficient resources to complete the processing of a request. An improper handling of such an overload situation at proxies and servers can impact the operation of the RTSP deployment, and probably worsen the situation. RTSP defines the 503 (Service Unavailable) response (Section 17.5.4) to let servers and proxies notify requesting proxies and RTSP clients about an overload situation. In conjunction with the Retry-After header (Section 18.44) the server or proxy can indicate the time after which the requesting entity can send another request to the proxy or server.

There are two scopes of such 503 answers, one for established RTSP sessions, where the request resulting in the 503 response as well as the response carries a Session header identifying the session which is suffering overload. This response only applies to this particular session. The other scope is the general RTSP server as identified by the host in the request URL. Such a 503 answer with any Retry-After header applies to all non-session specific requests to that server, including SETUP request intended to create a new RTSP session.

Another scope for overload situation exists, which is the RTSP proxy. To enable an RTSP proxy to signal that it is overloaded, or otherwise unavailable and can't handle the request, a 553 response code has been defined with the meaning "Proxy Unavailable". As with servers, there is a separation in response scopes between requests associated with existing RTSP sessions, and requests to create new sessions or general proxy requests.

Simply implementing and using the 503 (Service Unavailable) and 553 (Proxy Unavailable) is not sufficient for properly handling overload situations. For instance, a simplistic approach would be to send the 503 response with a Retry-After header set to a fixed value. However, this can cause the situation where multiple RTSP clients again send requests to a proxy or server at roughly the same time which may again cause an overload situation, or if the "old" overload

situation is not yet solved, i.e., the length indicated in the Retry-After header was too short.

An RTSP server or proxy in an overload situation must select the value of the Retry-After header carefully and bearing in mind its current load situation. It is REQUIRED to increase the timeout period in proportion to the current load on the server, i.e., an increasing workload should result in an increased length of the indicated unavailability. It is REQUIRED to not send the same value in the Retry-After header to all requesting proxies and clients, but to add a variation to the mean value of the Retry-After header.

A more complex case may arise when a load balancing RTSP proxy is in use. This is the case when an RTSP proxy is used to select amongst a set of RTSP servers to handle the requests or when multiple server addresses are available for a given server name. The proxy or client may receive a 503 (Service Unavailable) or 553 (Proxy Unavailable) from one of its RTSP servers or proxies, or a TCP timeout (if the server is even unable to handle the request message). The proxy or client simply retries the other addresses or configured proxies, but may also receive a 503 (Service Unavailable) or 553 (Proxy Unavailable) response or TCP timeouts from those addresses. In such a situation, where none of the RTSP servers/proxies/addresses can handle the request, the RTSP agent has to wait before it can send any new requests to the RTSP server. Any additional request to a specific address MUST be delayed according to the Retry-After headers received. For addresses where no response was received or TCP timeout occurred, an initial wait timer SHOULD be set to 5 seconds. That timer MUST be doubled for each additional failure to connect or receive response until the value exceeds 30 minutes when the timers mean value may be set to 30 minutes. It is REQUIRED to not set the same value in the timer for each scheduling, but instead to add a variation to the mean value, resulting in picking a random value within the range from 0.5 to 1.5 times the mean value.

11. Capability Handling

This section describes the available capability handling mechanism which allows RTSP to be extended. Extensions to this version of the protocol are basically done in two ways. First, new headers can be added. Secondly, new methods can be added. The capability handling mechanism is designed to handle both cases.

When a method is added, the involved parties can use the OPTIONS method to discover whether it is supported. This is done by issuing an OPTIONS request to the other party. Depending on the URI it will either apply in regards to a certain media resource, the whole server in general, or simply the next hop. The OPTIONS response MUST

contain a Public header which declares all methods supported for the indicated resource.

It is not necessary to use OPTIONS to discover support of a method, as the client could simply try the method. If the receiver of the request does not support the method it will respond with an error code indicating the method is either not implemented (501) or does not apply for the resource (405). The choice between the two discovery methods depends on the requirements of the service.

Feature-tags are defined to handle functionality additions that are not new methods. Each feature-tag represents a certain block of functionality. The amount of functionality that a feature-tag represents can vary significantly. A feature-tag can for example represent the functionality a single RTSP header provides. Another feature-tag can represent much more functionality, such as the "play.basic" feature-tag (Section 11.1) which represents the minimal media delivery for playback implementation.

Feature-tags are used to determine whether the client, server or proxy supports the functionality that is necessary to achieve the desired service. To determine support of a feature-tag, several different headers can be used, each explained below:

Supported: This header is used to determine the complete set of functionality that both client and server have in general and is not dependent on a specific resource. The intended usage is to determine before one needs to use a functionality that it is supported. It can be used in any method, but OPTIONS is the most suitable one as it at the same time determines all methods that are implemented. When sending a request the requester declares all its capabilities by including all supported feature-tags. This results in the receiver learning the requester's feature support. The receiver then includes its set of features in the response.

Proxy-Supported: This header is used similarly to the Supported header, but instead of giving the supported functionality of the client or server it provides both the requester and the responder a view of the common functionality supported in general by all members of the proxy chain between the two supports and not dependent on the resource. Proxies are required to add this header whenever the Supported header is present, but proxies may also add it independently of the requester.

Require: This header can be included in any request where the endpoint, i.e., the client or server, is required to understand

the feature to correctly perform the request. This can, for example, be a SETUP request where the server is required to understand a certain parameter to be able to set up the media delivery correctly. Ignoring this parameter would not have the desired effect and is not acceptable. Therefore the end-point receiving a request containing a Require MUST negatively acknowledge any feature that it does not understand and not perform the request. The response in cases where features are not supported are 551 (Option Not Supported). Also the features that are not supported are given in the Unsupported header in the response.

Proxy-Require: This header has the same purpose and workings as Require except that it only applies to proxies and not the end-point. Features that need to be supported by both proxies and end-points need to be included in both the Require and Proxy-Require header.

Unsupported: This header is used in a 551 error response, to indicate which features were not supported. Such a response is only the result of the usage of the Require and/or Proxy-Require header where one or more features were not supported. This information allows the requester to make the best of situations as it knows which features are not supported.

11.1. Feature Tag: play.basic

An implementation supporting all normative parts of this specification for the setup and control of playback of media uses the feature tag "play.basic" to indicate this support. The appendices (starting with letters), are not part of the functionality include in the feature tag unless the appendix is explicitly specified in a main section as being a required appendix.

Note: This feature tag does not mandate any media delivery protocol, such as RTP.

In RTSP 1.0 there was a minimal implementation section. However, that was not consistent with the rest of the specification. So rather than making an attempt to explicitly enumerate the features for play.basic this specification has to be taken as a whole and the necessary features normatively defined as being required are included.

12. Pipelining Support

Pipelining is a general method to improve performance of request response protocols by allowing the requesting agent to have more than one request outstanding and send them over the same persistent connection. For RTSP, where the relative order of requests will matter, it is important to maintain the order of the requests. Because of this, the responding agent **MUST** process the incoming requests in their sending order. The sending order can be determined by the CSeq header and its sequence number. For TCP the delivery order will be the same between two agents, as the sending order. The processing of the request **MUST** also have been finished before processing the next request from the same agent. The responses **MUST** be sent in the order the requests were processed.

RTSP 2.0 has extended support for pipelining compared to RTSP 1.0. The major improvement is to allow all requests involved in setting up and initiating media delivery to be pipelined after each other. This is accomplished by the utilization of the Pipelined-Requests header (see Section 18.33). This header allows a client to request that two or more requests are processed in the same RTSP session context which the first request creates. In other words, a client can request that two or more media streams are set-up and then played without needing to wait for a single response. This speeds up the initial startup time for an RTSP session by at least one RTT.

If a pipelined request builds on the successful completion of one or more prior requests the requester must verify that all requests were executed as expected. A common example will be two SETUP requests and a PLAY request. In case one of the SETUP fails unexpectedly, the PLAY request can still be successfully executed. However, the resulting presentation will not be as expected by the requesting client, as only a single media instead of two will be played. In this case the client can send a PAUSE request, correct the failing SETUP request and then request it to be played.

13. Method Definitions

The method indicates what is to be performed on the resource identified by the Request-URI. The method name is case-sensitive. New methods may be defined in the future. Method names **MUST NOT** start with a \$ character (decimal 36) and **MUST** be a token as defined by the ABNF [RFC5234] in the syntax chapter Section 20. The methods are summarized in Table 7.

method	direction	object	Server req.	Client req.
DESCRIBE	C -> S	P,S	recommended	recommended
GET_PARAMETER	C -> S	P,S	optional	optional
	S -> C	P,S	optional	optional
OPTIONS	C -> S	P,S	required	required
	S -> C	P,S	optional	optional
PAUSE	C -> S	P,S	required	required
PLAY	C -> S	P,S	required	required
PLAY_NOTIFY	S -> C	P,S	required	required
REDIRECT	S -> C	P,S	optional	required
SETUP	C -> S	S	required	required
SET_PARAMETER	C -> S	P,S	required	optional
	S -> C	P,S	optional	optional
TEARDOWN	C -> S	P,S	required	required
	S -> C	P	required	required

Table 7: Overview of RTSP methods, their direction, and what objects (P: presentation, S: stream) they operate on. Further it indicates if a server or a client implementation are required (mandatory), recommended or if it is optional to implement the method.

Note on Table 7: GET_PARAMETER is optional. For example, a fully functional server can be built to deliver media without any parameters. However, SET_PARAMETER is required, i.e., mandatory to implement for the server, this is due to its usage for keep-alive. PAUSE is required because it is the only way of leaving the Play state without terminating the whole session.

If an RTSP agent does not support a particular method, it MUST return 501 (Not Implemented) and the requesting RTSP agent, in turn, SHOULD NOT try this method again for the given agent / resource combination. An RTSP proxy whose main function is to log or audit and not modify

transport or media handling in any way MAY forward RTSP messages with unknown methods. Note that the proxy still needs to perform the minimal required processing, like adding the Via header.

13.1. OPTIONS

The semantics of the RTSP OPTIONS method is similar to that of the HTTP OPTIONS method described in [H9.2]. In RTSP however, OPTIONS is bi-directional, in that a client can send the request to a server and vice versa. A client MUST implement the capability to send an OPTIONS request and a server or a proxy MUST implement the capability to respond to an OPTIONS request. In addition to this "MUST implement" functionality, clients, servers and proxies MAY provide support both for sending OPTIONS requests and generating responses to the requests.

An OPTIONS request may be issued at any time. Such a request does not modify the session state. However, it may prolong the session lifespan (see below). The URI in an OPTIONS request determines the scope of the request and the corresponding response. If the Request-URI refers to a specific media resource on a given host, the scope is limited to the set of methods supported for that media resource by the indicated RTSP agent. A Request-URI with only the host address limits the scope to the specified RTSP agent's general capabilities without regard to any specific media. If the Request-URI is an asterisk ("*"), the scope is limited to the general capabilities of the next hop (i.e., the RTSP agent in direct communication with the request sender).

Regardless of the scope of the request, the Public header MUST always be included in the OPTIONS response listing the methods that are supported by the responding RTSP agent. In addition, if the scope of the request is limited to a media resource, the Allow header MUST be included in the response to enumerate the set of methods that are allowed for that resource unless the set of methods completely matches the set in the Public header. If the given resource is not available, the RTSP agent SHOULD return an appropriate response code such as 3rr or 4xx. The Supported header MAY be included in the request to query the set of features that are supported by the responding RTSP agent.

The OPTIONS method can be used to keep an RTSP session alive. However, this is not the preferred way of session keep-alive signaling, see Section 18.49. An OPTIONS request intended for keeping alive an RTSP session MUST include the Session header with the associated session identifier. Such a request SHOULD also use the media or the aggregated control URI as the Request-URI.

Example:

```
C->S:  OPTIONS rtsp://server.example.com RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2
      Proxy-Require: gzipped-messages
      Supported: play.basic

S->C:  RTSP/2.0 200 OK
      CSeq: 1
      Public: DESCRIBE, SETUP, TEARDOWN, PLAY, PAUSE, OPTIONS
      Supported: play.basic, setup.rtp.rtcp.mux, play.scale
      Server: PhonyServer/1.1
```

Note that some of the feature-tags in Supported and Proxy-Require are fictitious features.

13.2. DESCRIBE

The DESCRIBE method is used to retrieve the description of a presentation or media object from a server. The Request-URI of the DESCRIBE request identifies the media resource of interest. The client MAY include the Accept header in the request to list the description formats that it understands. The server MUST respond with a description of the requested resource and return the description in the message body of the response, if the DESCRIBE method request can be successfully fulfilled. The DESCRIBE reply-response pair constitutes the media initialization phase of RTSP.

The DESCRIBE response SHOULD contain all media initialization information for the resource(s) that it describes. Servers SHOULD NOT use the DESCRIBE response as a means of media indirection by having the description point at another server; instead, using the 3rr responses is RECOMMENDED.

By forcing a DESCRIBE response to contain all media initialization information for the set of streams that it describes, and discouraging the use of DESCRIBE for media indirection, any looping problems can be avoided that might have resulted from other approaches.

Example:

```
C->S: DESCRIBE rtsp://server.example.com/fizzle/foo RTSP/2.0
      CSeq: 312
      User-Agent: PhonyClient/1.2
      Accept: application/sdp, application/example

S->C: RTSP/2.0 200 OK
      CSeq: 312
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Server: PhonyServer/1.1
      Content-Base: rtsp://server.example.com/fizzle/foo/
      Content-Type: application/sdp
      Content-Length: 358

      v=0
      o=MNobody 2890844526 2890842807 IN IP4 192.0.2.46
      s=SDP Seminar
      i=A Seminar on the session description protocol
      u=http://www.example.com/lectures/sdp.ps
      e=seminar@example.com (Seminar Management)
      c=IN IP4 0.0.0.0
      a=control:*
      t=2873397496 2873404696
      m=audio 3456 RTP/AVP 0
      a=control:audio
      m=video 2232 RTP/AVP 31
      a=control:video
```

Media initialization is a requirement for any RTSP-based system, but the RTSP specification does not dictate that this is required to be done via the DESCRIBE method. There are three ways that an RTSP client may receive initialization information:

- o via an RTSP DESCRIBE request
- o via some other protocol (HTTP, email attachment, etc.)
- o via some form of user interface

If a client obtains a valid description from an alternate source, the client MAY use this description for initialization purposes without issuing a DESCRIBE request for the same media. The client should use any MTag to either validate the presentation description or make the session establishment conditional on being valid.

It is RECOMMENDED that minimal servers support the DESCRIBE method, and highly recommended that minimal clients support the ability to act as "helper applications" that accept a media initialization file

from a user interface, and/or other means that are appropriate to the operating environment of the clients.

13.3. SETUP

The description below uses the following states in a protocol state machine that is related to a specific session when that session has been created. The state transitions are driven by protocol interactions. For additional information about the state machine see Appendix B.

Init: Initial state: no session exists.

Ready: Session is ready to start playing.

Play: Session is playing, i.e., sending media stream data in the direction S->C.

The SETUP request for a URI specifies the transport mechanism to be used for the streamed media. The SETUP method may be used in two different cases; Create an RTSP session and change the transport parameters of already set up media stream(s). SETUP can be used in all three states; Init, and Ready, for both purposes and in PLAY to change the transport parameters. The usage of SETUP method in the Play state to add a media resource to the session is unspecified (Section 3.1).

The Transport header, see Section 18.54, specifies the media transport parameters acceptable to the client for data transmission; the response will contain the transport parameters selected by the server. This allows the client to enumerate in descending order of preference the transport mechanisms and parameters acceptable to it, while the server can select the most appropriate. It is expected that the session description format used will enable the client to select a limited number of possible configurations that are offered to the server to choose from. All transport related parameters SHALL be included in the Transport header; the use of other headers for this purpose is NOT RECOMMENDED due to middleboxes, such as firewalls or NATs.

For the benefit of any intervening firewalls, a client MUST indicate the known transport parameters, even if it has no influence over these parameters, for example, where the server advertises a fixed multicast address as destination.

Since SETUP includes all transport initialization information, firewalls and other intermediate network devices (which need this information) are spared the more arduous task of parsing the

DESCRIBE response, which has been reserved for media initialization.

The client MUST include the Accept-Ranges header in the request indicating all supported unit formats in the Range header. This allows the server to know which formats it may use in future session related responses, such as a PLAY response without any range in the request. If the client does not support a time format necessary for the presentation, the server MUST respond using 456 (Header Field Not Valid for Resource) and include the Accept-Ranges header with the range unit formats supported for the resource.

In a SETUP response the server MUST include the Accept-Ranges header (see Section 18.5) to indicate which time formats are acceptable to use for this media resource.

The SETUP response 200 OK MUST include the Media-Properties header (see Section 18.29). The combination of the parameters of the Media-Properties header indicates the nature of the content present in the session (see also Section 4.7). For example, a live stream with time shifting is indicated by

- o Random Access set to Random-Access,
- o Content Modifications set to Time Progressing,
- o Retention set to Time-Duration (with specific recording window time value).

The SETUP response 200 OK MUST include the Media-Range header (see Section 18.30) if the media is Time-Progressing.

A basic example for SETUP:

```

C->S: SETUP rtsp://example.com/foo/bar/baz.rm RTSP/2.0
      CSeq: 302
      Transport: RTP/AVP;unicast;dest_addr=":4588"/":4589",
                RTP/AVP/TCP;unicast;interleaved=0-1
      Accept-Ranges: npt, clock
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 302
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Server: PhonyServer/1.1
      Session: 47112344;timeout=60
      Transport: RTP/AVP;unicast;dest_addr="192.0.2.53:4588"/
                "192.0.2.53:4589"; src_addr="198.51.100.241:6256"/
                "198.51.100.241:6257"; ssrc=2A3F93ED
      Accept-Ranges: npt
      Media-Properties: Random-Access=3.2, Time-Progressing,
                      Time-Duration=3600.0
      Media-Range: npt=0-2893.23

```

In the above example the client wants to create an RTSP session containing the media resource "rtsp://example.com/foo/bar/baz.rm". The transport parameters acceptable to the client are either RTP/AVP/UDP (UDP per default) to be received on client port 4588 and 4589 at the address the RTSP setup connection comes from or RTP/AVP interleaved on the RTSP control channel. The server selects the RTP/AVP/UDP transport and adds the address and ports it will send and receive RTP and RTCP from, and the RTP SSRC that will be used by the server.

The server MUST generate a session identifier in response to a successful SETUP request, unless a SETUP request to a server includes a session identifier or a Pipelined-Requests header referencing an existing session context, in which case the server MUST bundle this SETUP request into the existing session (aggregated session) or return error 459 (Aggregate Operation Not Allowed) (see Section 17.4.24). An Aggregate control URI MUST be used to control an aggregated session. This URI MUST be different from the stream control URIs of the individual media streams included in the aggregate (see Section 13.4.2 for aggregated sessions and for the particular URIs see Appendix D.1.1). The Aggregate control URI is to be specified by the session description if the server supports aggregated control and aggregated control is desired for the session. However, even if aggregated control is offered the client MAY chose to not set up the session in aggregated control. If an Aggregate control URI is not specified in the session description, it is normally an indication that non-aggregated control should be used.

The SETUP of media streams in an aggregate which has not been given an aggregated control URI is unspecified.

While the session ID sometimes carries enough information for aggregate control of a session, the Aggregate control URI is still important for some methods such as SET_PARAMETER where the control URI enables the resource in question to be easily identified. The Aggregate control URI is also useful for proxies, enabling them to route the request to the appropriate server, and for logging, where it is useful to note the actual resource that a request was operating on.

A session will exist until it is either removed by a TEARDOWN request or is timed-out by the server. The server MAY remove a session that has not demonstrated liveness signs from the client(s) within a certain timeout period. The default timeout value is 60 seconds; the server MAY set this to a different value and indicate so in the timeout field of the Session header in the SETUP response. For further discussion see Section 18.49. Signs of liveness for an RTSP session are:

- o Any RTSP request from a client which includes a Session header with that session's ID.
- o If RTP is used as a transport for the underlying media streams, an RTCP sender or receiver report from the client(s) for any of the media streams in that RTSP session. RTCP Sender Reports may for example be received in sessions where the server is invited into a conference session and is valid for keep-alive.

If a SETUP request on a session fails for any reason, the session state, as well as transport and other parameters for associated streams MUST remain unchanged from their values as if the SETUP request had never been received by the server.

13.3.1. Changing Transport Parameters

A client MAY issue a SETUP request for a stream that is already set up or playing in the session to change transport parameters, which a server MAY allow. If it does not allow changing of parameters, it MUST respond with error 455 (Method Not Valid In This State). The reasons to support changing transport parameters include allowing application layer mobility and flexibility to utilize the best available transport as it becomes available. If a client receives a 455 when trying to change transport parameters while the server is in Play state, it MAY try to put the server in Ready state using PAUSE, before issuing the SETUP request again. If that also fails the changing of transport parameters will require that the client

performs a TEARDOWN of the affected media and then to set it up again. For an aggregated session avoiding tearing down all the media at the same time will avoid the creation of a new session.

All transport parameters MAY be changed. However, the primary usage expected is to either change the transport protocol completely, like switching from Interleaved TCP mode to UDP or vice versa, or to change the delivery address.

In a SETUP response for a request to change the transport parameters while in Play state, the server MUST include the Range to indicate at what point the new transport parameters will be used. Further, if RTP is used for delivery, the server MUST also include the RTP-Info header to indicate at what timestamp and RTP sequence number the change will take place. If both RTP-Info and Range are included in the response the "rtp_time" parameter and start point in the Range header MUST be for the corresponding time, i.e., be used in the same way as for PLAY to ensure the correct synchronization information is available.

If the transport parameters change while in Play state results in a change of synchronization related information, for example changing RTP SSRC, the server MUST provide in the SETUP response the necessary synchronization information. However, the server is RECOMMENDED to avoid changing the synchronization information if possible.

13.4. PLAY

This section describes the usage of the PLAY method in general, for aggregated sessions, and in different usage scenarios.

13.4.1. General Usage

The PLAY method tells the server to start sending data via the mechanism specified in SETUP and which part of the media should be played out. PLAY requests are valid when the session is in Ready or Play states. A PLAY request MUST include a Session header to indicate which session the request applies to.

Upon receipt of the PLAY request, the server MUST position the normal play time to the beginning of the range specified in the received Range header, within the limits of the media resource and in accordance with the Seek-Style header (Section 18.47) and deliver stream data until the end of the range if given, until a new PLAY request is received, or until the end of the media is reached. If no Range header is present in the PLAY request the server SHALL play from current pause point until the end of media. The pause point defaults at session start to the beginning of the media. For media

that is time-progressing and has no retention, the pause point will always be set equal to NPT "now", i.e., the current delivery point. The pause point may also be set to a particular point in the media by the PAUSE method, see Section 13.6. The pause point for media that is currently playing is equal to the current media position. For time-progressing media with time-limited retention, if the pause point represents a position that is older than what is retained by the server, the pause point will be moved to the oldest retained.

What range values are valid depends on the type of content. For content that isn't time progressing the range value is valid if the given range is part of any media within the aggregate. In other words the valid media range for the aggregate is the union of all of the media components in the aggregate. If a given range value points outside of the media, the response MUST be the 457 (Invalid Range) error code and include the Media-Range header (Section 18.30) with the valid range for the media. Except for time progressing content where the client requests a start point prior to what is retained, the start point is adjusted to the oldest retained content. For a start point that is beyond the media front edge, i.e., beyond the current value for "now", the server SHALL adjust the start value to the current front edge. The Range header's stop point value may point beyond the current media edge. In that case, the server SHALL deliver media from the requested (and possibly adjusted) start point until the provided stop point, or the end of the media is reached prior to the specified stop point. Please note that if one simply wants to play from a particular start point until the end of media using a Range header with an implicit stop point is RECOMMENDED.

If a client requests to start playing at the end of media, either explicitly with a Range header or implicitly with a pause point that is at the end of media, a 457 (Invalid Range) error MUST be sent and include the Media-Range header (Section 18.30). It is specified below that the Range header also must be included in the response and that it will carry the pause point in the media, in the case of the session being in Ready State. Note that this also applies if the pause point or requested start point is at the beginning of the media and a Scale header (Section 18.46) is included with a negative value (playing backwards).

For media with random access properties a client may express its preference on which policy for start point selection the server shall use. This is done by including the Seek-Style header (Section 18.47) in the PLAY request. The Seek-Style applied will affect the content of the Range header as it will be adjusted to indicate from what point the media actually is delivered.

A client desiring to play the media from the beginning MUST send a PLAY request with a Range header pointing at the beginning, e.g., "npt=0-". If a PLAY request is received without a Range header and media delivery has stopped at the end, the server SHOULD respond with a 457 "Invalid Range" error response. In that response, the current pause point MUST be included in a Range header.

All range specifiers in this specification allow for ranges with an implicit start point (e.g., "npt=-30"). When used in a PLAY request, the server treats this as a request to start or resume delivery from the current pause point, ending at the end time specified in the Range header. If the pause point is located later than the given end value, a 457 (Invalid Range) response MUST be given.

The example below will play seconds 10 through 25. It also requests the server to deliver media from the first Random Access Point prior to the indicated start point.

```
C->S: PLAY rtsp://audio.example.com/audio RTSP/2.0
      CSeq: 835
      Session: 12345678
      Range: npt=10-25
      Seek-Style: RAP
      User-Agent: PhonyClient/1.2
```

Servers MUST include a "Range" header in any PLAY response, even if no Range header was present in the request. The response MUST use the same format as the request's range header contained. If no Range header was in the request, the format used in any previous PLAY request within the session SHOULD be used. If no format has been indicated in a previous request the server MAY use any time format supported by the media and indicated in the Accept-Ranges header in the SETUP request. It is RECOMMENDED that NPT is used if supported by the media.

For any error response to a PLAY request, the server's response depends on the current session state. If the session is in Ready state, the current pause-point is returned using Range header with the pause point as the explicit start-point and an implicit stop-point. For time-progressing content where the pause-point moves with real-time due to limited retention, the current pause point is returned. For sessions in Play state, the current playout point and the remaining parts of the range request is returned. For any media with retention longer than 0 seconds the currently valid Media-Range header SHALL also be included in the response.

A PLAY response MAY include a header carrying synchronization information. As the information necessary is dependent on the media

transport format, further rules specifying the header and its usage are needed. For RTP the RTP-Info header is specified, see Section 18.45, and used in the following example.

Here is a simple example for a single audio stream where the client requests the media starting from 3.52 seconds and to the end. The server sends a 200 OK response with the actual play time which is 10 ms prior (3.51) and the RTP-Info header that contains the necessary parameters for the RTP stack.

```
C->S: PLAY rtsp://example.com/audio RTSP/2.0
      CSeq: 836
      Session: 12345678
      Range: npt=3.52-
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 836
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Server: PhonyServer/1.0
      Range: npt=3.51-324.39
      Seek-Style: First-Prior
      RTP-Info:url="rtsp://example.com/audio"
               ssrc=0D12F123:seq=14783;rtptime=2345962545

S->C: RTP Packet TS=2345962545 => NPT=3.51
      Media duration=0.16 seconds
```

The server replies with the actual start point that will be delivered. This may differ from the requested range if alignment of the requested range to valid frame boundaries is required for the media source. Note that some media streams in an aggregate may need to be delivered from even earlier points. Also, some media formats have a very long duration per individual data unit, therefore it might be necessary for the client to parse the data unit, and select where to start. The server SHALL also indicate which policy it uses for selecting the actual start point by including a Seek-Style header.

In the following example the client receives the first media packet that stretches all the way up and past the requested playtime. Thus, it is the client's decision whether to render to the user the time between 3.52 and 7.05, or to skip it. In most cases it is probably most suitable not to render that time period.


```
C->S: PLAY rtsp://example.com/audio RTSP/2.0
      CSeq: 836
      Session: 12345678
      Range: npt=7.05-
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 836
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Server: PhonyServer/1.0
      Range: npt=3.52-
      Seek-Style: First-Prior
      RTP-Info:url="rtsp://example.com/audio"
               ssrc=0D12F123:seq=14783;rtptime=2345962545

S->C: RTP Packet TS=2345962545 => NPT=3.52
      Duration=4.15 seconds
```

After playing the desired range, the presentation does NOT change to the Ready state, media delivery simply stops. If it is necessary to put the stream into the Ready state, a PAUSE request MUST be issued to do that. A PLAY request while the stream is still in the Play state is legal, and can be issued without an intervening PAUSE request. Such a request MUST replace the current PLAY action with the new one requested, i.e., being handled in the same way as if as the request was received in Ready state. In the case that the range in Range header has an implicit start time ("-endtime"), the server MUST continue to play from where it currently was until the specified end point. This is useful to change the end to at another point than in the previous request.

The following example plays the whole presentation starting at SMPTE time code 0:10:20 until the end of the clip. Note: The RTP-Info headers has been broken into several lines, where following lines start with whitespace as allowed by the syntax.

```
C->S: PLAY rtsp://audio.example.com/twister.en RTSP/2.0
      CSeq: 833
      Session: 12345678
      Range: smpte=0:10:20-
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 833
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: smpte=0:10:22-0:15:45
      Seek-Style: Next
      RTP-Info:url="rtsp://example.com/twister.en"
               ssrc=0D12F123:seq=14783;rtptime=2345962545
```

For playing back a recording of a live presentation, it may be desirable to use clock units:

```
C->S: PLAY rtsp://audio.example.com/meeting.en RTSP/2.0
      CSeq: 835
      Session: 12345678
      Range: clock=19961108T142300Z-19961108T143520Z
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 835
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: clock=19961108T142300Z-19961108T143520Z
      Seek-Style: Next
      RTP-Info:url="rtsp://example.com/meeting.en"
               ssrc=0D12F123:seq=53745;rtptime=484589019
```

13.4.2. Aggregated Sessions

PLAY requests can operate on sessions controlling a single media and on aggregated sessions controlling multiple media.

In an aggregated session the PLAY request MUST contain an aggregated control URI. A server MUST respond with error 460 (Only Aggregate Operation Allowed) if the client PLAY Request-URI is for a single media. The media in an aggregate MUST be played in sync. If a client wants individual control of the media, it needs to use separate RTSP sessions for each media.

For aggregated sessions where the initial SETUP request (creating a session) is followed by one or more additional SETUP requests, a PLAY request MAY be pipelined after those additional SETUP requests without awaiting their responses. This procedure can reduce the delay from start of session establishment until media play-out has started with one round trip time. However, a client needs to be aware that using this procedure will result in the playout of the server state established at the time of processing the PLAY, i.e., after the processing of all the requests prior to the PLAY request in the pipeline. This state may not be the intended one due to failure of any of the prior requests. A client can easily determine this based on the responses from those requests. In case of failure, the client can halt the media playout using PAUSE and try to establish the intended state again before issuing another PLAY request.

13.4.3. Updating current PLAY Requests

Clients can issue PLAY requests while the stream is in Play state and thus updating their request.

The important difference compared to a PLAY request in Ready state is the handling of the current play point and how the Range header in the request is constructed. The session is actively playing media and the play point will be moving, making the exact time a request will take effect hard to predict. Depending on how the PLAY header appears two different cases exist: total replacement or continuation. A total replacement is signaled by having the first range specification have an explicit start value, e.g., "npt=45-" or "npt=45-60", in which case the server stops playout at the current playout point and then starts delivering media according to the Range header. This is equivalent to having the client first send a PAUSE and then a new PLAY request that isn't based on the pause point. In the case of continuation the first range specifier has an implicit start point and an explicit stop value (Z), e.g., "npt=-60", which indicate that it MUST convert the range specifier being played prior to this PLAY request (X to Y) into (X to Z) and continue as this was the request originally played. If the current delivery point is beyond the stop point, the server SHALL immediately pause delivery. As the request has been completed successfully it shall be responded with 200 OK. A PLAY_NOTIFY with end-of-stream is also sent to indicate the actual stop point. The pause point is set to the requested stop point.

Following is an example of this behavior: The server has received requests to play ranges 10 to 15. If the new PLAY request arrives at the server 4 seconds after the previous one, it will take effect while the server still plays the first range (10-15). The server

changes the current play to continue to 25 seconds, i.e., the equivalent single request would be PLAY with "range: npt=10-25".

```
C->S: PLAY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 834
      Session: 12345678
      Range: npt=10-15
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 834
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: npt=10-15
      Seek-Style: Next
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
                  ssrc=0D12F123:seq=5712;rtptime=934207921,
                  url="rtsp://example.com/fizzle/videotrack"
                  ssrc=789DAF12:seq=57654;rtptime=2792482193
      Session: 12345678

C->S: PLAY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 835
      Session: 12345678
      Range: npt=-25
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 835
      Date: Thu, 23 Jan 1997 15:35:09 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: npt=14-25
      Seek-Style: Next
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
                  ssrc=0D12F123:seq=5712;rtptime=934239921,
                  url="rtsp://example.com/fizzle/videotrack"
                  ssrc=789DAF12:seq=57654;rtptime=2792842193
```

A common use of a PLAY request while in Play state is changing the scale of the media, i.e., entering or leaving fast forward or fast rewind. The client can issue an updating PLAY request that is either a continuation or a complete replacement, as discussed above this section. Below is an example of a client that is requesting a fast forward (scale=2) without giving a stop point and then change from fast forward to regular playout (scale = 1). In the second PLAY request the time is set explicitly to be where ever the server

currently plays out (npt=now-) and the server responds with the actual playback point where the new scale actually takes effect (npt=02:17:27.144-).

```
C->S: PLAY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 2034
      Session: 12345678
      Range: npt=now-
      Scale: 2.0
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 2034
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: npt=02:17:21.394-
      Seek-Style: Next
      Scale: 2.0
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
                ssrc=0D12F123:seq=5712;rtptime=934207921,
                url="rtsp://example.com/fizzle/videotrack"
                ssrc=789DAF12:seq=57654;rtptime=2792482193
```

[playing in fast forward and now returning to scale = 1]

```
C->S: PLAY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 2035
      Session: 12345678
      Range: npt=now-
      Scale: 1.0
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 2035
      Date: Thu, 23 Jan 1997 15:35:09 GMT
      Session: 12345678
      Server: PhonyServer/1.0
      Range: npt=02:17:27.144-
      Seek-Style: Next
      Scale: 1.0
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
                ssrc=0D12F123:seq=5712;rtptime=934239921,
                url="rtsp://example.com/fizzle/videotrack"
                ssrc=789DAF12:seq=57654;rtptime=2792842193
```

13.4.4. Playing On-Demand Media

On-demand media is indicated by the content of the Media-Properties header in the SETUP response by (see also Section 18.29):

- o Random Access property is set to Random-Access;
- o Content Modifications set to Immutable;
- o Retention set to Unlimited or Time-Limited.

Playing on-demand media follows the general usage as described in Section 13.4.1.

13.4.5. Playing Dynamic On-Demand Media

Dynamic on-demand media is indicated by the content of the Media-Properties header in the SETUP response by (see also Section 18.29):

- o Random Access set to Random-Access;
- o Content Modifications set to Dynamic;
- o Retention set to Unlimited or Time-Limited.

Playing on-demand media follows the general usage as described in Section 13.4.1 as long as the media has not been changed.

There are two ways for the client to be informed about changes of media resources in Play state. The client will receive a PLAY_NOTIFY request with Notify-Reason header set to media-properties-update (see Section 13.5.2. The client can use the value of the Media-Range to decide further actions, if the Media-Range header is present in the PLAY_NOTIFY request. The second way is that the client issues a GET_PARAMETER request without a body but including a Media-Range header. The 200 OK response MUST include the current Media-Range header (see Section 18.30).

13.4.6. Playing Live Media

Live media is indicated by the content of the Media-Properties header in the SETUP response by (see also Section 18.29):

- o Random-Access set to No-Seeking;
- o Content Modifications set to Time-Progressing;
- o Retention with Time-Duration set to 0.0.

For live media, the SETUP response 200 OK MUST include the Media-Range header (see Section 18.30).

A client MAY send PLAY requests without the Range header. If the request includes the Range header it MUST use a symbolic value representing "now". For NPT that range specification is "npt=now-". The server MUST include the Range header in the response and it MUST indicate an explicit time value and not a symbolic value. In other words, "npt=now-" is not valid to be used in the response. Instead the time since session start is recommended expressed as an open interval, e.g., "npt=96.23-". An absolute time value (clock) for the corresponding time MAY be given, i.e., "clock=20030213T143205Z-". The Absolute Time format can only be used if client has shown support for it using the Accept-Ranges header.

13.4.7. Playing Live with Recording

Certain media servers may offer recording services of live sessions to their clients. This recording would normally be from the beginning of the media session. Clients can randomly access the media between now and the beginning of the media session. This live media with recording is indicated by the content of the Media-Properties header in the SETUP response by (see also Section 18.29):

- o Random Access set to Random-Access;
- o Content Modifications set to Time-Progressing;
- o Retention set to Time-Limited or Unlimited

The SETUP response 200 OK MUST include the Media-Range header (see Section 18.30) for this type of media. For live media with recording, the Range header indicates the current delivery point in the media and the Media-Range header indicates the currently available media window around the current time. This window can cover recorded content in the past (seen from current time in the media) or recorded content in the future (seen from current time in the media). The server adjusts the delivery point to the requested border of the window. If the client requests a delivery point that is located outside the recording window, e.g., if the requested point is too far in the past, the server selects the oldest point in the recording. The considerations in Section 13.5.3 apply if a client requests delivery with Scale (Section 18.46) values other than 1.0 (Normal playback rate) while delivering live media with recording.

13.4.8. Playing Live with Time-Shift

Certain media servers may offer time-shift services to their clients. This time shift records a fixed interval in the past, i.e., a sliding window recording mechanism, but not past this interval. Clients can randomly access the media between now and the interval. This live media with recording is indicated by the content of the Media-Properties header in the SETUP response by (see also Section 18.29):

- o Random Access set to Random-Access;
- o Content Modifications set to Time-Progressing;
- o Retention set to Time-Duration and a value indicating the recording interval (>0).

The SETUP response 200 OK MUST include the Media-Range header (see Section 18.30) for this type of media. For live media with recording the Range header indicates the current time in the media and the Media Range indicates a window around the current time. This window can cover recorded content in the past (seen from current time in the media) or recorded content in the future (seen from current time in the media). The server adjusts the play point to the requested border of the window, if the client requests a play point that is located outside the recording windows, e.g., if requested too far in the past, the server selects the oldest range in the recording. The considerations in Section 13.5.3 apply, if a client requests delivery using a Scale (Section 18.46) value other than 1.0 (Normal playback rate) while delivering live media with time-shift.

13.5. PLAY_NOTIFY

The PLAY_NOTIFY method is issued by a server to inform a client about an asynchronous event for a session in Play state. The Session header MUST be presented in a PLAY_NOTIFY request and indicates the scope of the request. Sending of PLAY_NOTIFY requests requires a persistent connection between server and client, otherwise there is no way for the server to send this request method to the client.

PLAY_NOTIFY requests have an end-to-end (i.e., server to client) scope, as they carry the Session header, and apply only to the given session. The client SHOULD immediately return a response to the server.

PLAY_NOTIFY requests MAY use both aggregate control URI and individual media resource URIs depending on the scope of the notification. This scope may have important distinctions for aggregated sessions, and each reason for a PLAY_NOTIFY request needs

to specify the interpretation and if aggregated control URIs or individual URIs may be used in requests.

PLAY_NOTIFY requests can be used with a message body, depending on the value of the Notify-Reason header. It is described in the particular section for each Notify-Reason if a message body is used. However, currently there is no Notify-Reason that allows using a message body. In this case, there is a need to obey some limitations when adding new Notify-Reasons that intend to use a message body: the server can send any type of message body, but it is not ensured that the client can understand the received message body. This is related to DESCRIBE (see Section 13.2), but in this particular case the client can state its acceptable message bodies by using the Accept header. In the case of PLAY_NOTIFY, the server does not know which message bodies are understood by the client.

The Notify-Reason header (see Section 18.32) specifies the reason why the server sends the PLAY_NOTIFY request. This is extensible and new reasons can be added in the future (see Section 22.8). In case the client does not understand the reason for the notification it MUST respond with a 465 (Notification Reason Unknown) (Section 17.4.30) error code. Servers can send PLAY_NOTIFY with these types:

- o end-of-stream (see Section 13.5.1);
- o media-properties-update (see Section 13.5.2);
- o scale-change (see Section 13.5.3).

13.5.1. End-of-Stream

A PLAY_NOTIFY request with Notify-Reason header set to end-of-stream indicates the completion or near completion of the PLAY request and the ending delivery of the media stream(s). The request MUST NOT be issued unless the server is in the Play state. The end of the media stream delivery notification may be used to indicate either a successful completion of the PLAY request currently being served, or to indicate some error resulting in failure to complete the request. The Request-Status header (Section 18.42) MUST be included to indicate which request the notification is for and its completion status. The message response status codes (Section 8.1.1) are used to indicate how the PLAY request concluded. The sender of a PLAY_NOTIFY can issue an updated PLAY_NOTIFY, in the case of a PLAY_NOTIFY sent with wrong information. For instance, a PLAY_NOTIFY was issued before reaching the end-of-stream, but some error occurred resulting in that the previously sent PLAY_NOTIFY contained a wrong time when the stream will end. In this case a new PLAY_NOTIFY MUST

be sent including the correct status for the completion and all additional information.

PLAY_NOTIFY requests with Notify-Reason header set to end-of-stream MUST include a Range header and the Scale header if the scale value is not 1. The Range header indicates the point in the stream or streams where delivery is ending with the timescale that was used by the server in the PLAY response for the request being fulfilled. The server MUST NOT use the "now" constant in the Range header; it MUST use the actual numeric end position in the proper timescale. When end-of-stream notifications are issued prior to having sent the last media packets, this is evident as the end time in the Range header is beyond the current time in the media being received by the client, e.g., "npt=-15", if npt is currently at 14.2 seconds. The Scale header is to be included so that it is evident if the media time scale is moving backwards and/or have a non-default pace. The end-of-stream notification does not prevent the client from sending a new PLAY request.

If RTP is used as media transport, a RTP-Info header MUST be included, and the RTP-Info header MUST indicate the last sequence number in the seq parameter.

For an RTSP Session where media resources are under aggregated control the media resources will normally end at approximately the same time, although some small differences may exist, on the scale of a few hundred of milliseconds. In those cases a RTSP session under aggregated control SHOULD send only a single PLAY_NOTIFY request. By using the aggregate control URI in the PLAY_NOTIFY request the RTSP server indicates that this applies to all media resources within the session. In cases RTP is used for media delivery corresponding RTP-Info needs to be included for all media resources. In cases where one or more media resource has significantly shorter duration than some other resources in the aggregated session the server MAY send end-of-stream notifications using individual media resource URIs to indicate to agents that there will be no more media for this particular media resource related to the current active PLAY request. In such cases when the remaining media resources comes to end-of-stream they MUST send a PLAY_NOTIFY request using the aggregate control URI to indicate that no more resources remain.

A PLAY_NOTIFY request with Notify-Reason header set to end-of-stream MUST NOT carry a message body.

This example request notifies the client about a future end-of-stream event:

```
S->C: PLAY_NOTIFY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 854
      Notify-Reason: end-of-stream
      Request-Status: cseq=853 status=200 reason="OK"
      Range: npt=-145
      RTP-Info:url="rtsp://example.com/fizzle/foo/audio"
                ssrc=0D12F123:seq=14783;rtptime=2345962545,
                url="rtsp://example.com/fizzle/video"
                ssrc=789DAF12:seq=57654;rtptime=2792482193

      Session: uZ3ci0K+Ld-M
      Date: Mon, 08 Mar 2010 13:37:16 GMT

C->S: RTSP/2.0 200 OK
      CSeq: 854
      User-Agent: PhonyClient/1.2
      Session: uZ3ci0K+Ld-M
```

13.5.2. Media-Properties-Update

A PLAY_NOTIFY request with Notify-Reason header set to media-properties-update indicates an update of the media properties for the given session (see Section 18.29) and/or the available media range that can be played as indicated by Media-Range (Section 18.30). PLAY_NOTIFY requests with Notify-Reason header set to media-properties-update MUST include a Media-Properties and Date header and SHOULD include a Media-Range header. The Media-Properties header has session scope, thus for aggregated sessions the PLAY_NOTIFY request MUST be using the aggregated control URI.

This notification MUST be sent for media that are Time-Progressing every time an event happens that changes the basis for making estimates on how the available for play-back media range will progress with wall clock time. In addition it is RECOMMENDED that the server sends these notifications approximately every 5 minutes for time-progressing content to ensure the long-term stability of the client estimation and allowing for clock skew detection by the client. The time between notifications should be greater than 1 minute and less than 2 hours. For the reasons just explained, requests MUST include a Media-Range header to provide current Media duration and a Range header to indicate the current playing point and any remaining parts of the requested range.

The recommendation for sending updates every 5 minutes is due to any clock skew issues. In 5 minutes the clock skew should not become too significant as this is not used for media playback and synchronization, only for determining which content is available to the user.

A PLAY_NOTIFY request with Notify-Reason header set to media-properties-update MUST NOT carry a message body.

```
S->C: PLAY_NOTIFY rtsp://example.com/fizzle/foo RTSP/2.0
      Date: Tue, 14 Apr 2008 15:48:06 GMT
      CSeq: 854
      Notify-Reason: media-properties-update
      Session: uZ3ci0K+Ld-M
      Media-Properties: Time-Progressing,
                       Time-Limited=20080415T153919.36Z, Random-Access=5.0
      Media-Range: npt=00:00:00-01:37:21.394
      Range: npt=01:15:49.873-

C->S: RTSP/2.0 200 OK
      CSeq: 854
      User-Agent: PhonyClient/1.2
      Session: uZ3ci0K+Ld-M
```

13.5.3. Scale-Change

The server may be forced to change the rate of media time per playback time when a client requests delivery using a Scale (Section 18.46) value other than 1.0 (normal playback rate). For time progressing media with some retention, i.e., the server stores already sent content, a client requesting to play with Scale values larger than 1 may catch up with the front end of the media. The server will then be unable to continue to provide content at Scale larger than 1 as content is only made available by the server at Scale=1. Another case is when Scale < 1 and the media retention is time-duration limited. In this case the delivery point can reach the oldest media unit available, and further playback at this scale becomes impossible as there will be no media available. To avoid having the client lose any media, the scale will need to be adjusted to the same rate at which the media is removed from the storage buffer, commonly Scale = 1.0.

Another case is when the content itself consists of spliced pieces or is dynamically updated. In these cases the server may be required to change from one supported scale value (different than Scale=1.0) to another. In this case the server will pick the closest value and inform the client of what it has picked. In these cases the media properties will also be sent updating the supported Scale values. This enables a client to adjust the Scale value used.

To minimize impact on playback in any of the above cases the server MUST modify the playback properties and set Scale to a supportable value and continue delivery of the media. When doing this modification it MUST send a PLAY_NOTIFY message with the Notify-

Reason header set to "scale-change". The request MUST contain a Range header with the media time when the change took effect, a Scale header with the new value in use, Session header with the identifier for the session it applies to and a Date header with the server wallclock time of the change. For time progressing content also the Media-Range and the Media-Properties at this point in time MUST be included. The Media-Properties header MUST be included if the scale change was due to the content changing what scale values that is supported.

For media streams being delivered using RTP also a RTP-Info header MUST be included. It MUST contain the rtptime parameter with a value corresponding to the point of change in that media and optionally also the sequence number.

PLAY_NOTIFY requests for aggregated sessions MUST use the aggregated control URI in the request. The scale change for any aggregated session applies to all media streams part of the aggregate.

A PLAY_NOTIFY request with Notify-Reason header set to "Scale-Change" MUST NOT carry a message body.

```
S->C: PLAY_NOTIFY rtsp://example.com/fizzle/foo RTSP/2.0
      Date: Tue, 14 Apr 2008 15:48:06 GMT
      CSeq: 854
      Notify-Reason: scale-change
      Session: uZ3ci0K+Ld-M
      Media-Properties: Time-Progressing,
                      Time-Limited=20080415T153919.36Z, Random-Access=5.0
      Media-Range: npt=00:00:00-01:37:21.394
      Range: npt=01:37:21.394-
      Scale: 1
      RTP-Info: url="rtsp://example.com/fizzle/foo/audio"
                ssrc=0D12F123:rtptime=2345962545,
                url="rtsp://example.com/fizzle/videotrack"
                ssrc=789DAF12:seq=57654;rtptime=2792482193
```

```
C->S: RTSP/2.0 200 OK
      CSeq: 854
      User-Agent: PhonyClient/1.2
      Session: uZ3ci0K+Ld-M
```

13.6. PAUSE

The PAUSE request causes the stream delivery to immediately be interrupted (halted). A PAUSE request MUST be done either with the aggregated control URI for aggregated sessions, resulting in all media being halted, or the media URI for non-aggregated sessions.

Any attempt to do muting of a single media with a PAUSE request in an aggregated session MUST be responded to with error 460 (Only Aggregate Operation Allowed). After resuming playback, synchronization of the tracks MUST be maintained. Any server resources are kept, though servers MAY close the session and free resources after being paused for the duration specified with the timeout parameter of the Session header in the SETUP message.

Example:

```
C->S: PAUSE rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 834
      Session: 12345678
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 834
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Range: npt=45.76-75.00
```

The PAUSE request causes stream delivery to be interrupted immediately on receipt of the message and the pause point is set to the current point in the presentation. That pause point in the media stream needs to be maintained. A subsequent PLAY request without Range header resumes from the pause point and plays until media end.

The pause point after any PAUSE request MUST be returned to the client by adding a Range header with what remains unplayed of the PLAY request's range. For media with random access properties, if one desires to resume playing a ranged request, one simply includes the Range header from the PAUSE response and includes the Seek-Style header with the Next policy in the PLAY request. For media that is time-progressing and has retention duration=0 the follow-up PLAY request to start media delivery again, MUST use "npt=now-" and not the answer given in the response to PAUSE.

```
C->S: PLAY rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 834
      Session: 12345678
      Range: npt=10-30
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 834
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Server: PhonyServer/1.0
      Range: npt=10-30
      Seek-Style: First-Prior
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
                ssrc=0D12F123:seq=5712;rtptime=934207921,
                url="rtsp://example.com/fizzle/videotrack"
                ssrc=4FAD8726:seq=57654;rtptime=2792482193
      Session: 12345678
```

After 11 seconds, i.e., at 21 seconds into the presentation:

```
C->S: PAUSE rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 835
      Session: 12345678
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 835
      Date: 23 Jan 1997 15:35:17 GMT
      Server: PhonyServer/1.0
      Range: npt=21-30
      Session: 12345678
```

If a client issues a PAUSE request and the server acknowledges and enters the Ready state, the proper server response, if the player issues another PAUSE, is still 200 OK. The 200 OK response MUST include the Range header with the current pause point. See examples below:

```
C->S: PAUSE rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 834
      Session: 12345678
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 834
      Session: 12345678
      Date: Thu, 23 Jan 1997 15:35:06 GMT
      Range: npt=45.76-98.36
```

```
C->S: PAUSE rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 835
      Session: 12345678
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 835
      Session: 12345678
      Date: 23 Jan 1997 15:35:07 GMT
      Range: npt=45.76-98.36
```

13.7. TEARDOWN

13.7.1. Client to Server

The TEARDOWN client to server request stops the stream delivery for the given URI, freeing the resources associated with it. A TEARDOWN request can be performed on either an aggregated or a media control URI. However, some restrictions apply depending on the current state. The TEARDOWN request **MUST** contain a Session header indicating what session the request applies to. The TEARDOWN request **MUST NOT** include a Terminate-Reason header.

A TEARDOWN using the aggregated control URI or the media URI in a session under non-aggregated control (single media session) **MAY** be done in any state (Ready and Play). A successful request **MUST** result in that media delivery being immediately halted and the session state being destroyed. This **MUST** be indicated through the lack of a Session header in the response.

A TEARDOWN using a media URI in an aggregated session can only be done in Ready state. Such a request only removes the indicated media stream and associated resources from the session. This may result in a session returning to non-aggregated control, because it only contains a single media after the request's completion. A session that will exist after the processing of the TEARDOWN request **MUST** in the response to that TEARDOWN request contain a Session header. Thus

the presence of the Session header indicates to the receiver of the response if the session is still extant or has been removed.

Example:

```
C->S: TEARDOWN rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 892
      Session: 12345678
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 892
      Server: PhonyServer/1.0
```

13.7.2. Server to Client

The server can send TEARDOWN requests in the server to client direction to indicate that the server has been forced to terminate the ongoing session. This may happen for several reasons, such as server maintenance without available backup, or that the session has been inactive for extended periods of time. The reason is provided in the Terminate-Reason header (Section 18.52).

When a RTSP client has maintained a RTSP session that otherwise is inactive for an extended period of time the server may reclaim the resources. That is done by issuing a TEARDOWN request with the Terminate-Reason set to "Session-Timeout". This MAY be done when the client has been inactive in the RTSP session for more than one Session Timeout period (Section 18.49). However, the server is RECOMMENDED to not perform this operation until an extended period of inactivity of 10 times the Session Timeout period has passed. It is up to the operator of the RTSP server to actually configure how long this extended period of inactivity is. An operator should take into account when doing this configuration what the served content is and what this means for the extended period of inactivity.

In case the server needs to stop providing service to the established sessions and there is no server to point at in a REDIRECT request, then TEARDOWN SHALL be used to terminate the session. This method can also be used when non-recoverable internal errors have happened and the server has no other option then to terminate the sessions.

The TEARDOWN request MUST be done only on the session aggregate control URI (i.e., it is not allowed to terminate individual media streams, if it is a session aggregate) and MUST include the following headers; Session and Terminate-Reason headers. The request only applies to the session identified in the Session header. The server

may include a message to the client's user with the "user-msg" parameter.

The TEARDOWN request may alternatively be done on the wild card URI * and without any session header. The scope of such a request is limited to the next-hop (i.e., the RTSP agent in direct communication with the server) and applies, as well, to the RTSP connection between the next-hop RTSP agent and the server. This request indicates that all sessions and pending requests being managed via the connection are terminated. Any intervening proxies SHOULD do all of the following in the order listed:

1. respond to the TEARDOWN request
2. disconnect the control channel from the requesting server
3. pass the TEARDOWN request to each applicable client (typically those clients with an active session or an unanswered request)

Note: The proxy is responsible for accepting TEARDOWN responses from its clients; these responses MUST NOT be passed on to either the original server or the target server in the redirect.

13.8. GET_PARAMETER

The GET_PARAMETER request retrieves the value of any specified parameter or parameters for a presentation or stream specified in the URI. If the Session header is present in a request, the value of a parameter MUST be retrieved in the specified session context. There are two ways of specifying the parameters to be retrieved.

The first is by including headers which have been defined such that you can use them for this purpose. Headers for this purpose should allow empty, or stripped value parts to avoid having to specify bogus data when indicating the desire to retrieve a value. The successful completion of the request should also be evident from any filled out values in the response. The headers in this specification that MAY be used for retrieving their current value using GET_PARAMETER are listed below; additional headers MAY be specified in the future:

- o Accept-Ranges
- o Media-Range
- o Media-Properties
- o Range

o RTP-Info

The other way is to specify a message body that lists the parameter(s) that are desired to be retrieved. The Content-Type header (Section 18.19) is used to specify which format the message body has. If the receiver of the request does not support the media type used for the message body, it SHALL respond using the error code 415 (Unsupported Media Type). The responder to a GET_PARAMETER request MUST use the media type of the request for the response. For additional considerations regarding message body negotiation see Section 9.3.

RTSP Agents implementing support for responding to GET_PARAMETER requests SHALL implement the text/parameters format (Appendix F). This to ensure that at least one known format for parameters is implemented and thus prevent parameter format negotiation failure.

Parameters specified within the body of the message must all be understood by the request receiving agent. If one or more parameters are not understood a 451 (Parameter Not Understood) MUST be sent including a body listing these parameters that weren't understood. If all parameters are understood their values are filled in and returned in the response message body.

The method can also be used without a message body or any header that requests parameters for keep-alive purpose. The keep-alive timer has been updated for any request that is successful, i.e., a 200 OK response is received. Any non-required header present in such a request may or may not have been processed. Normally the presence of filled out values in the header will be indication that the header has been processed. However, for cases when this is difficult to determine, it is recommended to use a feature-tag and the Require header. For this reason it is usually easier if any parameters to be retrieved are sent in the body, rather than using any header.

Example:

```
S->C: GET_PARAMETER rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 431
      User-Agent: PhonyClient/1.2
      Session: 12345678
      Content-Length: 26
      Content-Type: text/parameters
```

```
      packets_received
      jitter
```

```
C->S: RTSP/2.0 200 OK
      CSeq: 431
      Session: 12345678
      Server: PhonyServer/1.1
      Date: Mon, 08 Mar 2010 13:43:23 GMT
      Content-Length: 38
      Content-Type: text/parameters
```

```
      packets_received: 10
      jitter: 0.3838
```

13.9. SET_PARAMETER

This method requests to set the value of a parameter or a set of parameters for a presentation or stream specified by the URI. The method MAY also be used without a message body. It is the RECOMMENDED method to be used in a request sent for the sole purpose of updating the keep-alive timer. If this request is successful, i.e., a 200 OK response is received, then the keep-alive timer has been updated. Any non-required header present in such a request may or may not have been processed. To allow a client to determine if any such header has been processed, it is necessary to use a feature tag and the Require header. Due to this reason it is RECOMMENDED that any parameters are sent in the body, rather than using any header.

When using a message body to list the parameter(s) that are desired to be set the Content-Type header (Section 18.19) is used to specify which format the message body has. If the receiver of the request is not supporting the media type used for the message body, it SHALL respond using the error code 415 (Unsupported Media Type). For additional considerations regarding message body negotiation see Section 9.3.

RTSP Agents implementing support for responding to SET_PARAMETER requests SHALL implement the text/parameters format (Appendix F). This to ensure that at least one known format for parameters is implemented and thus prevent parameter format negotiation failure.

A request is RECOMMENDED to only contain a single parameter to allow the client to determine why a particular request failed. If the request contains several parameters, the server MUST only act on the request if all of the parameters can be set successfully. A server MUST allow a parameter to be set repeatedly to the same value, but it MAY disallow changing parameter values. If the receiver of the request does not understand or cannot locate a parameter, error 451 (Parameter Not Understood) MUST be used. When a parameter is not allowed to change, the error code is 458 (Parameter Is Read-Only). The response body MUST contain only the parameters that have errors. Otherwise, a body MUST NOT be returned. The response body MUST use the media type of the request for the response.

Note: transport parameters for the media stream MUST only be set with the SETUP command.

Restricting setting transport parameters to SETUP is for the benefit of firewalls.

The parameters are split in a fine-grained fashion so that there can be more meaningful error indications. However, it may make sense to allow the setting of several parameters if an atomic setting is desirable. Imagine device control where the client does not want the camera to pan unless it can also tilt to the right angle at the same time.

Example:

```
C->S: SET_PARAMETER rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 421
      User-Agent: PhonyClient/1.2
      Session: iixT43KLc
      Date: Mon, 08 Mar 2010 14:45:04 GMT
      Content-length: 20
      Content-type: text/parameters
```

```
      barparam: barstuff
```

```
S->C: RTSP/2.0 451 Parameter Not Understood
      CSeq: 421
      Session: iixT43KLc
      Server: PhonyServer/1.0
      Date: Mon, 08 Mar 2010 14:44:56 GMT
      Content-length: 20
      Content-type: text/parameters
```

```
      barparam: barstuff
```

13.10. REDIRECT

The REDIRECT method is issued by a server to inform a client that the service provided will be terminated and where a corresponding service can be provided instead. This may happen for different reasons. One is that the server is being administered such that it must stop providing service. Thus the client is required to connect to another server location to access the resource indicated by the Request-URI.

The REDIRECT request SHALL contain a Terminate-Reason header (Section 18.52) to inform the client of the reason for the request. Additional parameters related to the reason may also be included. The intention here is to allow a server administrator to do a controlled shutdown of the RTSP server. That requires sufficient time to inform all entities having associated state with the server and for them to perform a controlled migration from this server to a fall back server.

A REDIRECT request with a Session header has end-to-end (i.e., server to client) scope and applies only to the given session. Any intervening proxies SHOULD NOT disconnect the control channel while there are other remaining end-to-end sessions. The REQUIRED Location header MUST contain a complete absolute URI pointing to the resource to which the client SHOULD reconnect. Specifically, the Location MUST NOT contain just the host and port. A client may receive a REDIRECT request with a Session header, if and only if, an end-to-end session has been established.

A client may receive a REDIRECT request without a Session header at any time when it has communication or a connection established with a server. The scope of such a request is limited to the next-hop (i.e., the RTSP agent in direct communication with the server) and applies to all sessions controlled, as well as the connection between the next-hop RTSP agent and the server. A REDIRECT request without a Session header indicates that all sessions and pending requests being managed via the connection MUST be redirected. The Location header, if included in such a request, SHOULD contain an absolute URI with only the host address and the OPTIONAL port number of the server to which the RTSP agent SHOULD reconnect. Any intervening proxies SHOULD do all of the following in the order listed:

1. respond to the REDIRECT request
2. disconnect the control channel from the requesting server
3. connect to the server at the given host address

4. pass the REDIRECT request to each applicable client (typically those clients with an active session or an unanswered request)

Note: The proxy is responsible for accepting REDIRECT responses from its clients; these responses MUST NOT be passed on to either the original server or the redirected server.

When the server lacks any alternative server and needs to terminate a session or all sessions the TEARDOWN request SHALL be used instead.

When no Terminate-Reason "time" parameter is included in a REDIRECT request, the client SHALL perform the redirection immediately and return a response to the server. The server shall consider the session as terminated and can free any associated state after it receives the successful (2xx) response. The server MAY close the signaling connection upon receiving the response and the client SHOULD close the signaling connection after sending the 2xx response. The exception to this is when the client has several sessions on the server being managed by the given signaling connection. In this case, the client SHOULD close the connection when it has received and responded to REDIRECT requests for all the sessions managed by the signaling connection.

The Terminate-Reason header "time" parameter MAY be used to indicate the wallclock time by when the redirection MUST have taken place. To allow a client to determine that redirect time without being time synchronized with the server, the server MUST include a Date header in the request. The client should have terminated the session and closed the connection before the redirection time-line terminated. The server MAY simply cease to provide service when the deadline time has been reached, or it may issue TEARDOWN requests to the remaining sessions.

If the REDIRECT request times out following the rules in Section 10.4 the server MAY terminate the session or transport connection that would be redirected by the request. This is a safeguard against misbehaving clients that refuse to respond to a REDIRECT request. Thus, removing any incentive to not acknowledge the reception of a REDIRECT request.

After a REDIRECT request has been processed, a client that wants to continue to receive media for the resource identified by the Request-URI will have to establish a new session with the designated host. If the URI given in the Location header is a valid resource URI, a client SHOULD issue a DESCRIBE request for the URI.

Note: The media resource indicated by the Location header can be identical, slightly different or totally different. This is the reason why a new DESCRIBE request SHOULD be issued.

If the Location header contains only a host address, the client may assume that the media on the new server is identical to the media on the old server, i.e., all media configuration information from the old session is still valid except for the host address. However, the usage of conditional SETUP using MTag identifiers is RECOMMENDED as a means to verify the assumption.

This example request redirects traffic for this session to the new server at the given absolute time:

```
S->C: REDIRECT rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 732
      Location: rtsp://s2.example.com:8001
      Terminate-Reason: Server-Admin ;time=19960213T143205Z
      Session: uZ3ci0K+Ld-M
      Date: Thu, 13 Feb 1996 14:30:43 GMT

C->S: RTSP/2.0 200 OK
      CSeq: 732
      User-Agent: PhonyClient/1.2
      Session: uZ3ci0K+Ld-M
```

14. Embedded (Interleaved) Binary Data

In order to fulfill certain requirements on the network side, e.g., in conjunction with network address translators that block RTP traffic over UDP, it may be necessary to interleave RTSP messages and media stream data. This interleaving should generally be avoided unless necessary since it complicates client and server operation and imposes additional overhead. Also, head-of-line blocking may cause problems. Interleaved binary data SHOULD only be used if RTSP is carried over TCP. Interleaved data is not allowed inside RTSP messages.

Stream data such as RTP packets is encapsulated by an ASCII dollar sign (36 decimal), followed by a one-octet channel identifier, followed by the length of the encapsulated binary data as a binary, two-octet unsigned integer in network octet order (Appendix B of [RFC0791]). The stream data follows immediately afterwards, without a CRLF, but including the upper-layer protocol headers. Each \$ block MUST contain exactly one upper-layer protocol data unit, e.g., one RTP packet.

Note that this mechanism does not support PDUs larger than 65535 octets, which matches the maximum payload size of regular, non-jumbo IPv4 and IPv6 packets. If the media delivery protocol intended to be used has larger PDUs than that, definition of a PDU fragmentation mechanism will be required to support embedded binary data.

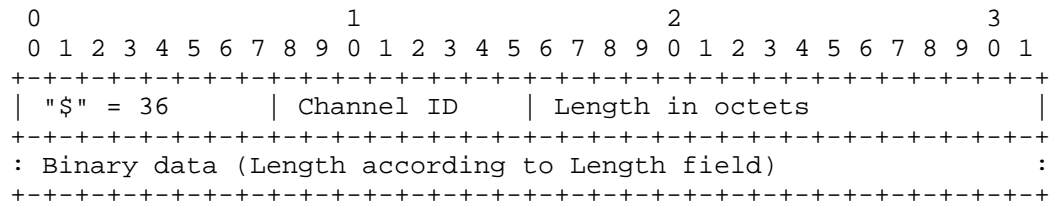


Figure 1: Embedded Interleaved Binary Data Format

The channel identifier is defined in the Transport header with the interleaved parameter (Section 18.54).

When the transport choice is RTP, RTCP messages are also interleaved by the server over the TCP connection. The usage of RTCP messages is indicated by including an interval containing a second channel in the interleaved parameter of the Transport header, see Section 18.54. If RTCP is used, packets MUST be sent on the first available channel higher than the RTP channel. The channels are bi-directional, using the same Channel ID in both directions, and therefore RTCP traffic is sent on the second channel in both directions.

RTCP is sometimes needed for synchronization when two or more streams are interleaved in such a fashion. Also, this provides a convenient way to tunnel RTP/RTCP packets through the RTSP connection (TCP or TCP/TLS) when required by the network configuration and transfer them onto UDP when possible.

```
C->S: SETUP rtsp://example.com/bar.file RTSP/2.0
      CSeq: 2
      Transport: RTP/AVP/TCP;unicast;interleaved=0-1
      Accept-Ranges: npt, smpte, clock
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 2
      Date: Thu, 05 Jun 1997 18:57:18 GMT
      Transport: RTP/AVP/TCP;unicast;interleaved=5-6
      Session: 12345678
      Accept-Ranges: npt
      Media-Properties: Random-Access=0.2, Immutable, Unlimited

C->S: PLAY rtsp://example.com/bar.file RTSP/2.0
      CSeq: 3
      Session: 12345678
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 3
      Session: 12345678
      Date: Thu, 05 Jun 1997 18:57:19 GMT
      RTP-Info: url="rtsp://example.com/bar.file"
                ssrc=0D12F123:seq=232433;rtptime=972948234
      Range: npt=0-56.8
      Seek-Style: RAP

S->C: $005{2 octet length}{"length" octets data, w/RTP header}
S->C: $005{2 octet length}{"length" octets data, w/RTP header}
S->C: $006{2 octet length}{"length" octets RTCP packet}
```

15. Proxies

RTSP Proxies are RTSP agents that are located in between a client and a server. A proxy can take on both the role as a client and as server depending on what it tries to accomplish. RTSP proxies use two transport layer connections, one from the RTSP client to the RTSP proxy and a second from the RTSP proxy to the RTSP server. Proxies are introduced for several different reasons and those listed below are often combined.

Caching Proxy: This type of proxy is used to reduce the workload on servers and connections. By caching the description and media streams, i.e., the presentation, the proxy can serve a client with content, but without requesting it from the server once it has been cached and has not become stale. See the caching Section 16. This type of proxy is also expected to understand

RTSP end-point functionality, i.e., functionality identified in the Require header in addition to what Proxy-Require demands.

Translator Proxy: This type of proxy is used to ensure that an RTSP client gets access to servers and content on an external network or using content encodings not supported by the client. The proxy performs the necessary translation of addresses, protocols or encodings. This type of proxy is expected to also understand RTSP end-point functionality, i.e., functionality identified in the Require header in addition to what Proxy-Require demands.

Access Proxy: This type of proxy is used to ensure that an RTSP client gets access to servers on an external network. Thus this proxy is placed on the border between two domains, e.g., a private address space and the public Internet. The proxy performs the necessary translation, usually addresses. This type of proxy is required to redirect the media to itself or a controlled gateway that performs the translation before the media can reach the client.

Security Proxy: This type of proxy is used to help facilitate security functions around RTSP. For example when having a firewalled network, the security proxy requests that the necessary pinholes in the firewall are opened when a client in the protected network wants to access media streams on the external side. This proxy can perform its function without redirecting the media between the server and client. However, in deployments with private address spaces this proxy is likely to be combined with the access proxy. Anyway, the functionality of this proxy is usually closely tied into understanding all aspects of the media transport.

Auditing Proxy: RTSP proxies can also provide network owners with a logging and audit point for RTSP sessions, e.g., for corporations that track their employees usage of the network. This type of proxy can perform its function without inserting itself or any other node in the media transport. This proxy type can also accept unknown methods as it doesn't interfere with the clients' requests.

All types of proxies can also be used when using secured communication with TLS as RTSP 2.0 allows the client to approve certificate chains used for connection establishment from a proxy, see Section 19.3.2. However, that trust model may not be suitable for all types of deployment. In those cases, the secured sessions do by-pass the proxies.

Access proxies SHOULD NOT be used in equipment like NATs and firewalls that aren't expected to be regularly maintained, like home or small office equipment. In these cases it is better to use the NAT traversal procedures defined for RTSP 2.0 [I-D.ietf-mmusic-rtsp-nat]. The reason for these recommendations is that any extensions of RTSP resulting in new media transport protocols or profiles, new parameters, etc. may fail in a proxy that isn't maintained. This would impede RTSP's future development and usage.

15.1. Proxies and Protocol Extensions

The existence of proxies must always be considered when developing new RTSP extensions. Most types of proxies will need to implement any new method to operate correctly in the presence of that extension. New headers can be introduced and will not be blocked by older proxies. However, it is important to consider if this header and its function is required to be understood by the proxy or can be simply forwarded. If the header needs to be understood, a feature-tag representing the functionality MUST be included in the Proxy-Require header. Below are guidelines for analysis whether the header needs to be understood. The transport header and its parameters are extensible which on the other hand requires handling rules for a proxy in order to ensure a correct interpretation.

Whether a proxy needs to understand a header is not easy to determine, as they serve a broad variety of functions. When evaluating if a header needs to be understood, one can divide the functionality into three main categories:

Media modifying: The caching and translator proxies are modifying the actual media and therefore need to understand also the request directed to the server that affects how the media is rendered. Thus, this type of proxy needs to also understand the server side functionality.

Transport modifying: The access and the security proxy both need to understand how the media transport is performed, either for opening pinholes or to translate the outer headers, e.g., IP and UDP or TCP.

Non-modifying: The audit proxy is special in that it does not modify the messages in other ways than to insert the Via header. That makes it possible for this type to forward RTSP messages that contain different types of unknown methods, headers or header parameters.

Based on the above classification, one should evaluate if the new functionality requires the Transport modifying type of proxies to understand it or not.

15.2. Multiplexing and Demultiplexing of Messages

RTSP proxies may have to multiplex multiple RTSP sessions from their clients towards RTSP servers. This requires that RTSP requests from multiple clients are multiplexed onto a common connection for requests outgoing to an RTSP server and on the way back the responses are demultiplexed from the server to per client responses. On the protocol level this requires that request and response messages are handled in both ways, requiring that there is a mechanism to correlate what request/response pair exchanged between proxy and server is mapped to what client (or client request).

This multiplexing of requests and demultiplexing of responses is done by using the CSeq header field. The proxy has to rewrite the CSeq in requests to the server and responses from the server and remember what CSeq is mapped to what client. The proxy also needs to ensure that the order of the message related to each client is maintained. Section 18.20 is defining the handling of how requests and responses are rewritten.

16. Caching

In HTTP, request-response pairs are cached. RTSP differs significantly in that respect. Responses are not cacheable, with the exception of the presentation description returned by DESCRIBE. (Since the responses for anything but DESCRIBE and GET_PARAMETER do not return any data, caching is not really an issue for these requests.) However, it is desirable for the continuous media data, typically delivered out-of-band with respect to RTSP, to be cached, as well as the session description.

On receiving a SETUP or PLAY request, a proxy ascertains whether it has an up-to-date copy of the continuous media content and its description. It can determine whether the copy is up-to-date by issuing a SETUP or DESCRIBE request, respectively, and comparing the Last-Modified header with that of the cached copy. If the copy is not up-to-date, it modifies the SETUP transport parameters as appropriate and forwards the request to the origin server. Subsequent control commands such as PLAY or PAUSE then pass the proxy unmodified. The proxy delivers the continuous media data to the client, while possibly making a local copy for later reuse. The exact allowed behavior of the cache is given by the cache-response directives described in Section 18.11. A cache MUST answer any DESCRIBE requests if it is currently serving the stream to the

requester, as it is possible that low-level details of the stream description may have changed on the origin-server.

Note that an RTSP cache, is of the "cut-through" variety. Rather than retrieving the whole resource from the origin server, the cache simply copies the streaming data as it passes by on its way to the client. Thus, it does not introduce additional latency.

To the client, an RTSP proxy cache appears like a regular media server. To the media origin server an RTSP proxy cache appears like a client. Just as an HTTP cache has to store the content type, content language, and so on for the objects it caches, a media cache has to store the presentation description. Typically, a cache eliminates all transport-references (e.g., multicast information) from the presentation description, since these are independent of the data delivery from the cache to the client. Information on the encodings remains the same. If the cache is able to translate the cached media data, it would create a new presentation description with all the encoding possibilities it can offer.

16.1. Validation Model

When a cache has a stale entry that it would like to use as a response to a client's request, it first has to check with the origin server (or possibly an intermediate cache with a fresh response) to see if its cached entry is still usable. This is called "validating" the cache entry. To avoid having to pay the overhead of retransmitting the full response if the cached entry is good, and at the same time avoiding to pay the overhead of an extra round trip if the cached entry is invalid, the RTSP protocol supports the use of conditional methods.

The key protocol features for supporting conditional methods are those concerned with "cache validators." When an origin server generates a full response, it attaches some sort of validator to it, which is kept with the cache entry. When a client (user agent or proxy cache) makes a conditional request for a resource for which it has a cache entry, it includes the associated validator in the request.

The server then checks that validator against the current validator for the requested resource, and, if they match (see Section 16.1.3), it responds with a special status code (usually, 304 (Not Modified)) and no message body. Otherwise, it returns a full response (including message body). Thus, avoiding transmitting the full response if the validator matches, and avoiding an extra round trip if it does not match.

In RTSP, a conditional request looks exactly the same as a normal request for the same resource, except that it carries a special header (which includes the validator) that implicitly turns the method (usually DESCRIBE or SETUP) into a conditional.

The protocol includes both positive and negative senses of cache-validating conditions. That is, it is possible to request either that a method be performed if and only if a validator matches or if and only if no validators match.

Note: a response that lacks a validator may still be cached, and served from cache until it expires, unless this is explicitly prohibited by a cache-control directive (see Section 18.11). However, a cache cannot do a conditional retrieval if it does not have a validator for the resource, which means it will not be refreshable after it expires.

Media streams that are being adapted based on the transport capacity between the server and the cache makes caching more difficult. A server needs to consider how it views caching of media streams that it adapts and potentially instruct any caches to not cache such streams.

16.1.1.1. Last-Modified Dates

The Last-Modified header (Section 18.27) value is often used as a cache validator. In simple terms, a cache entry is considered to be valid if the cache entry was created after the Last-Modified time.

16.1.1.2. Message Body Tag Cache Validators

The MTag response-header field value, a message body tag, provides for an "opaque" cache validator. This might allow more reliable validation in situations where it is inconvenient to store modification dates, where the one-second resolution of RTSP-date values is not sufficient, or where the origin server wishes to avoid certain paradoxes that might arise from the use of modification dates.

Message body tags are described in Section 4.6

16.1.1.3. Weak and Strong Validators

Since both origin servers and caches will compare two validators to decide if they represent the same or different entities, one normally would expect that if the message body (i.e., the presentation description) or any associated message body headers changes in any way, then the associated validator would change as well. If this is

true, then this validator is a "strong validator." The Message body (i.e., the presentation description) or any associated message body headers is named an entity for a better understanding.

However, there might be cases when a server prefers to change the validator only on semantically significant changes, and not when insignificant aspects of the entity change. A validator that does not always change when the resource changes is a "weak validator."

Message body tags are normally "strong validators," but the protocol provides a mechanism to tag a message body tag as "weak." One can think of a strong validator as one that changes whenever the bits of an entity changes, while a weak value changes whenever the meaning of an entity changes. Alternatively, one can think of a strong validator as part of an identifier for a specific entity, while a weak validator is part of an identifier for a set of semantically equivalent entities.

Note: One example of a strong validator is an integer that is incremented in stable storage every time an entity is changed.

An entity's modification time, if represented with one-second resolution, could be a weak validator, since it is possible that the resource might be modified twice during a single second.

Support for weak validators is optional. However, weak validators allow for more efficient caching of equivalent objects.

A "use" of a validator is either when a client generates a request and includes the validator in a validating header field, or when a server compares two validators.

Strong validators are usable in any context. Weak validators are only usable in contexts that do not depend on exact equality of an entity. For example, either kind is usable for a conditional DESCRIBE of a full entity. However, only a strong validator is usable for a sub-range retrieval, since otherwise the client might end up with an internally inconsistent entity.

Clients MAY issue DESCRIBE requests with either weak validators or strong validators. Clients MUST NOT use weak validators in other forms of requests.

The only function that the RTSP protocol defines on validators is comparison. There are two validator comparison functions, depending on whether the comparison context allows the use of weak validators or not:

- o The strong comparison function: in order to be considered equal, both validators MUST be identical in every way, and both MUST NOT be weak.
- o The weak comparison function: in order to be considered equal, both validators MUST be identical in every way, but either or both of them MAY be tagged as "weak" without affecting the result.

A message body tag is strong unless it is explicitly tagged as weak.

A Last-Modified time, when used as a validator in a request, is implicitly weak unless it is possible to deduce that it is strong, using the following rules:

- o The validator is being compared by an origin server to the actual current validator for the entity and,
- o That origin server reliably knows that the associated entity did not change more than once during the second covered by the presented validator.

OR

- o The validator is about to be used by a client in an If-Modified-Since, because the client has a cache entry for the associated entity, and
- o That cache entry includes a Date value, which gives the time when the origin server sent the original response, and
- o The presented Last-Modified time is at least 60 seconds before the Date value.

OR

- o The validator is being compared by an intermediate cache to the validator stored in its cache entry for the entity, and
- o That cache entry includes a Date value, which gives the time when the origin server sent the original response, and
- o The presented Last-Modified time is at least 60 seconds before the Date value.

This method relies on the fact that if two different responses were sent by the origin server during the same second, but both had the same Last-Modified time, then at least one of those responses would have a Date value equal to its Last-Modified time. The arbitrary 60-

second limit guards against the possibility that the Date and Last-Modified values are generated from different clocks, or at somewhat different times during the preparation of the response. An implementation MAY use a value larger than 60 seconds, if it is believed that 60 seconds is too short.

If a client wishes to perform a sub-range retrieval on a value for which it has only a Last-Modified time and no opaque validator, it MAY do this only if the Last-Modified time is strong in the sense described here.

16.1.4. Rules for When to Use Message Body Tags and Last-Modified Dates

This document adopt a set of rules and recommendations for origin servers, clients, and caches regarding when various validator types ought to be used, and for what purposes.

RTSP origin servers:

- o SHOULD send a message body tag validator unless it is not feasible to generate one.
- o MAY send a weak message body tag instead of a strong message body tag, if performance considerations support the use of weak message body tags, or if it is unfeasible to send a strong message body tag.
- o SHOULD send a Last-Modified value if it is feasible to send one, unless the risk of a breakdown in semantic transparency that could result from using this date in an If-Modified-Since header would lead to serious problems.

In other words, the preferred behavior for an RTSP origin server is to send both a strong message body tag and a Last-Modified value.

In order to be legal, a strong message body tag MUST change whenever the associated entity value changes in any way. A weak message body tag SHOULD change whenever the associated entity changes in a semantically significant way.

Note: in order to provide semantically transparent caching, an origin server MUST avoid reusing a specific strong message body tag value for two different entities, or reusing a specific weak message body tag value for two semantically different entities. Cache entries might persist for arbitrarily long periods, regardless of expiration times, so it might be inappropriate to expect that a cache will never again attempt to validate an entry using a validator that it obtained at some point in the past.

RTSP clients:

- o If a message body tag has been provided by the origin server, MUST use that message body tag in any cache-conditional request (using If-Match or If-None-Match).
- o If only a Last-Modified value has been provided by the origin server, SHOULD use that value in non-subrange cache-conditional requests (using If-Modified-Since).
- o If both a message body tag and a Last-Modified value have been provided by the origin server, SHOULD use both validators in cache-conditional requests.

An RTSP origin server, upon receiving a conditional request that includes both a Last-Modified date (e.g., in an If-Modified-Since header) and one or more message body tags (e.g., in an If-Match, If-None-Match, or If-Range header field) as cache validators, MUST NOT return a response status of 304 (Not Modified) unless doing so is consistent with all of the conditional header fields in the request.

Note: The general principle behind these rules is that RTSP servers and clients should transmit as much non-redundant information as is available in their responses and requests. RTSP systems receiving this information will make the most conservative assumptions about the validators they receive.

16.1.5. Non-validating Conditionals

The principle behind message body tags is that only the service author knows the semantics of a resource well enough to select an appropriate cache validation mechanism, and the specification of any validator comparison function more complex than octet-equality would open up a can of worms. Thus, comparisons of any other headers are never used for purposes of validating a cache entry.

16.2. Invalidation After Updates or Deletions

The effect of certain methods performed on a resource at the origin server might cause one or more existing cache entries to become non-transparently invalid. That is, although they might continue to be "fresh," they do not accurately reflect what the origin server would return for a new request on that resource.

There is no way for the RTSP protocol to guarantee that all such cache entries are marked invalid. For example, the request that caused the change at the origin server might not have gone through

the proxy where a cache entry is stored. However, several rules help reduce the likelihood of erroneous behavior.

In this section, the phrase "invalidate an entity" means that the cache will either remove all instances of that entity from its storage, or will mark these as "invalid" and in need of a mandatory revalidation before they can be returned in response to a subsequent request.

Some RTSP methods MUST cause a cache to invalidate an entity. This is either the entity referred to by the Request-URI, or by the Location or Content-Location headers (if present). These methods are:

- o DESCRIBE
- o SETUP

In order to prevent denial of service attacks, an invalidation based on the URI in a Location or Content-Location header MUST only be performed if the host part is the same as in the Request-URI.

A cache that passes through requests for methods it does not understand SHOULD invalidate any entities referred to by the Request-URI.

17. Status Code Definitions

Where applicable, HTTP status [H10] codes are reused. See Table 4 in Section 8.1 for a listing of which status codes may be returned by which requests. All error messages, 4xx and 5xx MAY return a body containing further information about the error.

17.1. Informational 1xx

17.1.1. 100 Continue

The client SHOULD continue with its request. This interim response is used to inform the client that the initial part of the request has been received and has not yet been rejected by the server. The client SHOULD continue by sending the remainder of the request or, if the request has already been completed, ignore this response. The server MUST send a final response after the request has been completed.

17.2. Success 2xx

This class of status code indicates that the client's request was successfully received, understood, and accepted.

17.2.1. 200 OK

The request has succeeded. The information returned with the response is dependent on the method used in the request.

17.3. Redirection 3xx

The notation "3xx" indicates response codes from 300 to 399 inclusive which are meant for redirection. The response code 304 is excluded, as it is not used for redirection and instead the "3rr" notation is used. The 304 response code appears here, rather than a 2xx response code which would have been appropriate, this as 304 has been used also in RTSP 1.0 [RFC2326].

Within RTSP, redirection may be used for load balancing or redirecting stream requests to a server topologically closer to the client. Mechanisms to determine topological proximity are beyond the scope of this specification.

A 3rr code MAY be used to respond to any request. The Location header MUST be included in any 3rr response. It is RECOMMENDED that they are used if necessary before a session is established, i.e., in response to DESCRIBE or SETUP. However, in cases where a server is not able to send a REDIRECT request to the client, the server MAY need to resort to using 3rr responses to inform a client with an established session about the need for redirecting the session. If a 3rr response is received for a request in relation to an established session, the client SHOULD send a TEARDOWN request for the session, and MAY reestablish the session using the resource indicated by the Location.

If the Location header is used in a response it MUST contain an absolute URI pointing out the media resource the client is redirected to, the URI MUST NOT only contain the host name.

In the event that an unknown 3rr status code is received, the agent SHOULD behave as if a 302 response code had been received (Section 17.3.3).

17.3.1. 300

This response code is not used in RTSP 2.0.

17.3.2. 301 Moved Permanently

The requested resource is moved permanently and resides now at the URI given by the Location header. The user client SHOULD redirect automatically to the given URI. This response MUST NOT contain a message-body. The Location header MUST be included in the response.

17.3.3. 302 Found

The requested resource resides temporarily at the URI given by the Location header. This response is intended to be used for many types of temporary redirects; e.g., load balancing. It is RECOMMENDED that the server set the reason phrase to something more meaningful than "Found" in these cases. The user client SHOULD redirect automatically to the given URI. This response MUST NOT contain a message-body.

This example shows a client being redirected to a different server:

```
C->S: SETUP rtsp://example.com/fizzle/foo RTSP/2.0
      CSeq: 2
      Transport: RTP/AVP/TCP;unicast;interleaved=0-1
      Accept-Ranges: npt, smpte, clock
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 302 Try Other Server
      CSeq: 2
      Location: rtsp://s2.example.com:8001/fizzle/foo
```

17.3.4. 303 See Other

This status code MUST NOT be used in RTSP 2.0. However, it was allowed in RTSP 1.0 [RFC2326].

17.3.5. 304 Not Modified

If the client has performed a conditional DESCRIBE or SETUP (see Section 18.25) and the requested resource has not been modified, the server SHOULD send a 304 response. This response MUST NOT contain a message-body.

The response MUST include the following header fields:

- o Date

- o MTag and/or Content-Location, if the header(s) would have been sent in a 200 response to the same request.
- o Expires and Cache-Control if the field-value might differ from that sent in any previous response for the same variant.

This response is independent for the DESCRIBE and SETUP requests. That is, a 304 response to DESCRIBE does NOT imply that the resource content is unchanged (only the session description) and a 304 response to SETUP does NOT imply that the resource description is unchanged. The MTag and If-Match headers may be used to link the DESCRIBE and SETUP in this manner.

17.3.6. 305 Use Proxy

The requested resource MUST be accessed through the proxy given by the Location field. The Location field gives the URI of the proxy. The recipient is expected to repeat this single request via the proxy. 305 responses MUST only be generated by origin servers.

17.4. Client Error 4xx

17.4.1. 400 Bad Request

The request could not be understood by the server due to malformed syntax. The client SHOULD NOT repeat the request without modifications. If the request does not have a CSeq header, the server MUST NOT include a CSeq in the response.

17.4.2. 401 Unauthorized

The request requires user authentication. The response MUST include a WWW-Authenticate header (Section 18.58) field containing a challenge applicable to the requested resource. The client MAY repeat the request with a suitable Authorization header field. If the request already included Authorization credentials, then the 401 response indicates that authorization has been refused for those credentials. If the 401 response contains the same challenge as the prior response, and the user agent has already attempted authentication at least once, then the user SHOULD be presented the message body that was given in the response, since that message body might include relevant diagnostic information. HTTP access authentication is explained in [RFC2617].

17.4.3. 402 Payment Required

This code is reserved for future use.

17.4.4. 403 Forbidden

The server understood the request, but is refusing to fulfill it. Authorization will not help and the request SHOULD NOT be repeated. If the server wishes to make public why the request has not been fulfilled, it SHOULD describe the reason for the refusal in the message body. If the server does not wish to make this information available to the client, the status code 404 (Not Found) can be used instead.

17.4.5. 404 Not Found

The server has not found anything matching the Request-URI. No indication is given of whether the condition is temporary or permanent. The 410 (Gone) status code SHOULD be used if the server knows, through some internally configurable mechanism, that an old resource is permanently unavailable and has no forwarding address. This status code is commonly used when the server does not wish to reveal exactly why the request has been refused, or when no other response is applicable.

17.4.6. 405 Method Not Allowed

The method specified in the request is not allowed for the resource identified by the Request-URI. The response MUST include an Allow header containing a list of valid methods for the requested resource. This status code is also to be used if a request attempts to use a method not indicated during SETUP.

17.4.7. 406 Not Acceptable

The resource identified by the request is only capable of generating response message bodies which have content characteristics not acceptable according to the Accept headers sent in the request.

The response SHOULD include a message body containing a list of available message body characteristics and location(s) from which the user or user agent can choose the one most appropriate. The message body format is specified by the media type given in the Content-Type header field. Depending upon the format and the capabilities of the user agent, selection of the most appropriate choice MAY be performed automatically. However, this specification does not define any standard for such automatic selection.

If the response could be unacceptable, a user agent SHOULD temporarily stop receipt of more data and query the user for a decision on further actions.

17.4.8. 407 Proxy Authentication Required

This code is similar to 401 (Unauthorized) (Section 17.4.2), but indicates that the client must first authenticate itself with the proxy. The proxy MUST return a Proxy-Authenticate header field (Section 18.34) containing a challenge applicable to the proxy for the requested resource.

17.4.9. 408 Request Timeout

The client did not produce a request within the time that the server was prepared to wait. The client MAY repeat the request without modifications at any later time.

17.4.10. 410 Gone

The requested resource is no longer available at the server and the forwarding address is not known. This condition is expected to be considered permanent. If the server does not know, or has no facility to determine, whether or not the condition is permanent, the status code 404 (Not Found) SHOULD be used instead. This response is cacheable unless indicated otherwise.

The 410 response is primarily intended to assist the task of repository maintenance by notifying the recipient that the resource is intentionally unavailable and that the server owners desire that remote links to that resource be removed. Such an event is common for limited-time, promotional services and for resources belonging to individuals no longer working at the server's site. It is not necessary to mark all permanently unavailable resources as "gone" or to keep the mark for any length of time -- that is left to the discretion of the owner of the server.

17.4.11. 411 Length Required

This error code is not defined for RTSP. This as the use of Content-Length (Section 18.17) is always required when message bodies are included in RTSP messages.

17.4.12. 412 Precondition Failed

The precondition given in one or more of the 'if-' request-header fields evaluated to false when it was tested on the server. See these sections for the 'if-' headers: If-Match Section 18.24, If-

Modified-Since Section 18.25, and If-None-Match Section 18.26. This response code allows the client to place preconditions on the current resource meta information (header field data) and thus prevent the requested method from being applied to a resource other than the one intended.

17.4.13. 413 Request Message Body Too Large

The server is refusing to process a request because the request message body is larger than the server is willing or able to process. The server MAY close the connection to prevent the client from continuing the request.

If the condition is temporary, the server SHOULD include a Retry-After header field to indicate that it is temporary and after what time the client MAY try again.

17.4.14. 414 Request-URI Too Long

The server is refusing to service the request because the Request-URI is longer than the server is willing to interpret. This rare condition is only likely to occur when a client has used a request with long query information, when the client has descended into a URI "black hole" of redirection (e.g., a redirected URI prefix that points to a suffix of itself), or when the server is under attack by a client attempting to exploit security holes present in some servers using fixed-length buffers for reading or manipulating the Request-URI.

17.4.15. 415 Unsupported Media Type

The server is refusing to service the request because the message body of the request is in a format not supported by the requested resource for the requested method.

17.4.16. 451 Parameter Not Understood

The recipient of the request does not support one or more parameters contained in the request. When returning this error message the sender SHOULD return a message body containing the offending parameter(s).

17.4.17. 452 reserved

This status code MUST NOT be used in RTSP 2.0. However, it was allowed in RTSP 1.0 [RFC2326].

17.4.18. 453 Not Enough Bandwidth

The request was refused because there was insufficient bandwidth. This may, for example, be the result of a resource reservation failure.

17.4.19. 454 Session Not Found

The RTSP session identifier in the Session header is missing, invalid, or has timed out.

17.4.20. 455 Method Not Valid in This State

The client or server cannot process this request in its current state. The response MUST contain an Allow header to make error recovery possible.

17.4.21. 456 Header Field Not Valid for Resource

The server could not act on a required request-header. For example, if PLAY contains the Range header field but the stream does not allow seeking. This error message may also be used for specifying when the time format in Range is impossible for the resource. In that case the Accept-Ranges header MUST be returned to inform the client of which format(s) that are allowed.

17.4.22. 457 Invalid Range

The Range value given is out of bounds, e.g., beyond the end of the presentation.

17.4.23. 458 Parameter Is Read-Only

The parameter to be set by SET_PARAMETER can be read but not modified. When returning this error message the sender SHOULD return a message body containing the offending parameter(s).

17.4.24. 459 Aggregate Operation Not Allowed

The requested method may not be applied on the URI in question since it is an aggregate (presentation) URI. The method may be applied on a media URI.

17.4.25. 460 Only Aggregate Operation Allowed

The requested method may not be applied on the URI in question since it is not an aggregate control (presentation) URI. The method may be applied on the aggregate control URI.

17.4.26. 461 Unsupported Transport

The Transport field did not contain a supported transport specification.

17.4.27. 462 Destination Unreachable

The data transmission channel could not be established because the client address could not be reached. This error will most likely be the result of a client attempt to place an invalid dest_addr parameter in the Transport field.

17.4.28. 463 Destination Prohibited

The data transmission channel was not established because the server prohibited access to the client address. This error is most likely the result of a client attempt to redirect media traffic to another destination with a dest_addr parameter in the Transport header.

17.4.29. 464 Data Transport Not Ready Yet

The data transmission channel to the media destination is not yet ready for carrying data. However, the responding agent still expects that the data transmission channel will be established at some point in time. Note, however, that this may result in a permanent failure like 462 "Destination Unreachable".

An example when this error may occur is in the case a client sends a PLAY request to a server prior to ensuring that the TCP connections negotiated for carrying media data was successfully established (In violation of this specification). The server would use this error code to indicate that the requested action could not be performed due to the failure of completing the connection establishment.

17.4.30. 465 Notification Reason Unknown

This indicates that the client has received a PLAY_NOTIFY (Section 13.5) with a Notify-Reason header (Section 18.32) unknown to the client.

17.4.31. 466 Key Management Error

This indicates that there has been an error in a Key Management function used in conjunction with a request. For example usage of MIKEY [RFC3830] according to Appendix C.1.4.1 may result in this error.

17.4.32. 470 Connection Authorization Required

The secured connection attempt needs user or client authorization before proceeding. The next hop's certificate is included in this response in the Accept-Credentials header.

17.4.33. 471 Connection Credentials not accepted

When performing a secure connection over multiple connections, an intermediary has refused to connect to the next hop and carry out the request due to unacceptable credentials for the used policy.

17.4.34. 472 Failure to establish secure connection

A proxy fails to establish a secure connection to the next hop RTSP agent. This is primarily caused by a fatal failure at the TLS handshake, for example due to server not accepting any cipher suites.

17.5. Server Error 5xx

Response status codes beginning with the digit "5" indicate cases in which the server is aware that it has erred or is incapable of performing the request. The server SHOULD include a message body containing an explanation of the error situation, and whether it is a temporary or permanent condition. User agents SHOULD display any included message body to the user. These response codes are applicable to any request method.

17.5.1. 500 Internal Server Error

The server encountered an unexpected condition which prevented it from fulfilling the request.

17.5.2. 501 Not Implemented

The server does not support the functionality required to fulfill the request. This is the appropriate response when the server does not recognize the request method and is not capable of supporting it for any resource.

17.5.3. 502 Bad Gateway

The server, while acting as a gateway or proxy, received an invalid response from the upstream server it accessed in attempting to fulfill the request.

17.5.4. 503 Service Unavailable

The server is currently unable to handle the request due to a temporary overloading or maintenance of the server. The implication is that this is a temporary condition which will be alleviated after some delay. If known, the length of the delay MAY be indicated in a Retry-After header. If no Retry-After is given, the client SHOULD handle the response as it would for a 500 response. The client MUST honor the length, if given in the Retry-After header.

Note: The existence of the 503 status code does not imply that a server must use it when becoming overloaded. Some servers may wish to simply refuse the connection.

The response scope is dependent on the Request. If the request was in relation to an existing RTSP session, the scope of the overload response is to this individual RTSP session. If the request was non-session specific or intended to form a RTSP session it applies to the RTSP server identified by the host name in the request URI.

17.5.5. 504 Gateway Timeout

The server, while acting as a proxy, did not receive a timely response from the upstream server specified by the URI or some other auxiliary server (e.g., DNS) it needed to access in attempting to complete the request.

17.5.6. 505 RTSP Version Not Supported

The server does not support, or refuses to support, the RTSP protocol version that was used in the request message. The server is indicating that it is unable or unwilling to complete the request using the same major version as the client other than with this error message. The response SHOULD contain a message body describing why that version is not supported and what other protocols are supported by that server.

17.5.7. 551 Option not supported

A feature-tag given in the Require or the Proxy-Require fields was not supported. The Unsupported header MUST be returned stating the feature for which there is no support.

17.5.8. 553 Proxy Unavailable

The proxy is currently unable to handle the request due to a temporary overloading or maintenance of the proxy. The implication is that this is a temporary condition which will be alleviated after

some delay. If known, the length of the delay MAY be indicated in a Retry-After header. If no Retry-After is given, the client SHOULD handle the response as it would for a 500 response. The client MUST honor the length, if given in the Retry-After header.

Note: The existence of the 553 status code does not imply that a proxy must use it when becoming overloaded. Some proxies may wish to simply refuse the connection.

The response scope is dependent on the Request. If the request was in relation to an existing RTSP session, the scope of the overload response is to this individual RTSP session. If the request was non-session specific or intended to form a RTSP session it applies to all such requests to this proxy.

18. Header Field Definitions

method	direction	object	acronym	Body
DESCRIBE	C -> S	P,S	DES	r
GET_PARAMETER	C -> S, S -> C	P,S	GPR	R,r
OPTIONS	C -> S, S -> C	P,S	OPT	
PAUSE	C -> S	P,S	PSE	
PLAY	C -> S	P,S	PLY	
PLAY_NOTIFY	S -> C	P,S	PNY	R
REDIRECT	S -> C	P,S	RDR	
SETUP	C -> S	S	STP	
SET_PARAMETER	C -> S, S -> C	P,S	SPR	R,r
TEARDOWN	C -> S	P,S	TRD	
	S -> C	P	TRD	

Table 8: Overview of RTSP methods, their direction, and what objects (P: presentation, S: stream) they operate on. Body notes if a method is allowed to carry body and in which direction, R = Request, r=response. Note: All error messages for statuses 4xx and 5xx are allowed to carry a body

The general syntax for header fields is covered in Section 5.2. This section lists the full set of header fields along with notes on meaning, and usage. The syntax definition for header fields are present in Section 20.2.3. Throughout this section, [HX.Y] is used to reference Section X.Y of the HTTP/1.1 specification RFC 2616 [RFC2616]. Examples of each header field are given.

Information about header fields in relation to methods and proxy processing is summarized in Table 9, Table 10, Table 11, and Table 12.

The "where" column describes the request and response types in which the header field can be used. Values in this column are:

R: header field may only appear in requests;

r: header field may only appear in responses;

2xx, 4xx, etc.: A numerical value or range indicates response codes with which the header field can be used;

c: header field is copied from the request to the response.

G: header field is a general-header and may be present in both requests and responses.

Note: General headers does not always use the "G" value in the where column. This is due to differences when the header may be applied in requests compared to responses. When such differences exist they are expressed using two different rows, one with where being "R" and one with it being "r".

The "proxy" column describes the operations a proxy may perform on a header field. An empty proxy column indicates that the proxy MUST NOT do any changes to that header, all allowed operations are explicitly stated:

a: A proxy can add or concatenate the header field if not present.

m: A proxy can modify an existing header field value.

d: A proxy can delete a header field value.

r: A proxy needs to be able to read the header field, and thus this header field cannot be encrypted.

The rest of the columns relate to the presence of a header field in a method. The method names when abbreviated, are according to Table 8:

- c: Conditional; requirements on the header field depend on the context of the message.
- m: The header field is mandatory.
- m*: The header field SHOULD be sent, but clients/servers need to be prepared to receive messages without that header field.
- o: The header field is optional.
- *: The header field MUST be present if the message body is not empty. See Section 18.17, Section 18.19 and Section 5.3 for details.
- : The header field is not applicable.

"Optional" means that a Client/Server MAY include the header field in a request or response. The Client/Server behavior when receiving such headers varies, for some it may ignore the header field, in other cases it is a request to process the header. This is regulated by the method and header descriptions. Example of headers that require processing are the Require and Proxy-Require header fields discussed in Section 18.43 and Section 18.37. A "mandatory" header field MUST be present in a request, and MUST be understood by the Client/Server receiving the request. A mandatory response-header field MUST be present in the response, and the header field MUST be understood by the Client/Server processing the response. "Not applicable" means that the header field MUST NOT be present in a request. If one is placed in a request by mistake, it MUST be ignored by the Client/Server receiving the request. Similarly, a header field labeled "not applicable" for a response means that the Client/Server MUST NOT place the header field in the response, and the Client/Server MUST ignore the header field in the response.

An RTSP agent MUST ignore extension headers that are not understood.

The From and Location header fields contain a URI. If the URI contains a comma, or semicolon, the URI MUST be enclosed in double quotes ("). Any URI parameters are contained within these quotes. If the URI is not enclosed in double quote, any semicolon-delimited parameters are header-parameters, not URI parameters.

Header	Where	Proxy	DE	OP	STP	PLY	PSE	TRD
Accept	R		o	-	-	-	-	-

Accept-Credentials	R	rm	o	o	o	o	o	o
Accept-Encoding	R	r	o	-	-	-	-	-
Accept-Language	R	r	o	-	-	-	-	-
Accept-Ranges	G	r	-	-	m	-	-	-
Accept-Ranges	456	r	-	-	-	m	-	-
Allow	r	am	c	c	c	-	-	-
Allow	405	am	m	m	m	m	m	m
Authentication-Info	r		o	o	o	o	o	o/-
Authorization	R		o	o	o	o	o	o
Bandwidth	R		o	o	o	o	-	-
Blocksize	R		o	-	o	o	-	-
Cache-Control	G	r	o	-	o	-	-	-
Connection	G	ad	o	o	o	o	o	o
Connection-Credentials	470, 407	ar	o	o	o	o	o	o
Content-Base	r		o	-	-	-	-	-
Content-Base	4xx, 5xx		o	o	o	o	o	o
Content-Encoding	R	r	-	-	-	-	-	-
Content-Encoding	r	r	o	-	-	-	-	-
Content-Encoding	4xx, 5xx	r	o	o	o	o	o	o
Content-Language	R	r	-	-	-	-	-	-
Content-Language	r	r	o	-	-	-	-	-
Content-Language	4xx,	r	o	o	o	o	o	o

	5xx								
Content-Length	r	r	*	-	-	-	-	-	-
Content-Length	4xx, 5xx	r	*	*	*	*	*	*	*
Content-Location	r	r	o	-	-	-	-	-	-
Content-Location	4xx, 5xx	r	o	o	o	o	o	o	o
Content-Type	r	r	*	-	-	-	-	-	-
Content-Type	4xx, 5xx	ar	*	*	*	*	*	*	*
CSeq	Gc	rm	m	m	m	m	m	m	m
Date	G	am	o/ *	o/ *	o/*	o/*	o/*	o/*	o/*
Expires	r	r	o	-	o	-	-	-	-
From	R	r	o	o	o	o	o	o	o
If-Match	R	r	-	-	o	-	-	-	-
If-Modified-Since	R	r	o	-	o	-	-	-	-
If-None-Match	R	r	o	-	o	-	-	-	-
Last-Modified	r	r	o	-	o	-	-	-	-
Location	3rr		o	o	o	o	o	o	o

Table 9: Overview of RTSP header fields (A-L) related to methods
DESCRIBE, OPTIONS, SETUP, PLAY, PAUSE, and TEARDOWN.

Header	Where	Pro xy	DE S	OP T	ST P	PLY	PSE	TRD
Media- Properties	G		-	-	m	m	m	-
Media-Range	G		-	-	m	m	m	-

MTag	r	r	o	-	o	-	-	-
Pipelined-Requests	G	amd r	-	o	o	o	o	o
Proxy-Authenticate	407	amr	m	m	m	m	m	m
Proxy-Authentication-Info	r	amd r	o	o	o	o	o	o/-
Proxy-Authorization	R	rd	o	o	o	o	o	o
Proxy- Require	R	ar	o	o	o	o	o	o
Proxy- Require	r	r	c	c	c	c	c	c
Proxy- Supported	R	amr	c	c	c	c	c	c
Proxy- Supported	r		c	c	c	c	c	c
Public	r	amr	-	m	-	-	-	-
Public	501	amr	m	m	m	m	m	m
Range	R		-	-	-	o	-	-
Range	r		-	-	c	m	m	-
Referrer	R		o	o	o	o	o	o
Request- Status	R		-	-	-	-	-	-
Require	R		o	o	o	o	o	o
Retry-After	3rr,503 ,553		o	o	o	o	o	-
Retry-After	413		o	-	-	-	-	-
RTP-Info	r		-	-	c	c	-	-
Scale	R	r	-	-	-	o	-	-
Scale	r	amr	-	-	-	c	-	-

Seek-Style	R		-	-	-	o	-	-
Seek-Style	r		-	-	-	m	-	-
Server	R	r	-	o	-	-	-	o
Server	r	r	o	o	o	o	o	o
Session	R	r	-	o	o	m	m	m
Session	r	r	-	c	m	m	m	o
Speed	R	adm r	-	-	-	o	-	-
Speed	r	adm r	-	-	-	c	-	-
Supported	R	amr	o	o	o	o	o	o
Supported	r	amr	c	c	c	c	c	c
Terminate-Reason	R	r	-	-	-	-	-	-
Timestamp	R	adm r	o	o	o	o	o	o
Timestamp	c	adm r	m	m	m	m	m	m
Transport	G	mr	-	-	m	-	-	-
Unsupported	r		c	c	c	c	c	c
User-Agent	R		m*	m*	m*	m*	m*	m*
Via	R	amr	o	o	o	o	o	o
Via	c	dr	m	m	m	m	m	m
WWW-Authenticate	401		m	m	m	m	m	m

Table 10: Overview of RTSP header fields (M-W) related to methods DESCRIBE, OPTIONS, SETUP, PLAY, PAUSE, and TEARDOWN.

Header	Where	Proxy	GPR	SPR	RDR	PNY
Accept	R	arm	o	o	-	-
Accept-Credentials	R	rm	o	o	o	-
Accept-Encoding	R	r	o	o	o	-
Accept-Language	R	r	o	o	o	-
Accept-Ranges	G	rm	o	-	-	-
Allow	405	amr	m	m	m	-
Authentication-Info	r		o/-	o/-	-	-
Authorization	R		o	o	o	-
Bandwidth	R		-	o	-	-
Blocksize	R		-	o	-	-
Cache-Control	G	r	o	o	-	-
Connection	G		o	o	o	o
Connection-Credentials	470, 407	ar	o	o	o	-
Content-Base	R		o	o	-	-
Content-Base	r		o	o	-	-
Content-Base	4xx, 5xx		o	o	o	o
Content-Encoding	R	r	o	o	-	-
Content-Encoding	r	r	o	o	-	-
Content-Encoding	4xx, 5xx	r	o	o	o	o
Content-Language	R	r	o	o	-	-
Content-Language	r	r	o	o	-	-

Content-Language	4xx, 5xx	r	o	o	o	o
Content-Length	R	r	*	*	-	-
Content-Length	r	r	*	*	-	-
Content-Length	4xx, 5xx	r	*	*	*	*
Content-Location	R		o	o	-	-
Content-Location	r		o	o	-	-
Content-Location	4xx, 5xx		o	o	o	o
Content-Type	R		*	*	-	-
Content-Type	r		*	*	-	-
Content-Type	4xx, 5xx		*	*	*	*
CSeq	R,c	mr	m	m	m	m
Date	R	a	o	o	m	o
Date	r	am	o	o	o	o
Expires	r	r	-	-	-	-
From	R	r	o	o	o	-
If-Match	R	r	-	-	-	-
If-Modified-Since	R	am	o	-	-	-
If-None-Match	R	am	o	-	-	-
Last-Modified	R	r	-	-	-	-
Last-Modified	r	r	o	-	-	-
Location	3rr		o	o	o	-
Location	R		-	-	m	-

Media-Properties	R	amr	o	-	-	c
Media-Properties	r	mr	c	-	-	-
Media-Range	R		o	-	-	c
Media-Range	r		c	-	-	-
MTag	r	r	o	-	-	-
Notify-Reason	R		-	-	-	m
Pipelined-Requests	R	amdr	o	o	-	-
Proxy-Authenticate	407	amdr	m	m	m	-
Proxy-Authentication-Info	r	amdr	o/-	o/-	-	-
Proxy-Authorization	R	amdr	o	o	o	-
Proxy-Require	R	ar	o	o	o	-
Proxy-Supported	R	amr	c	c	c	-
Proxy-Supported	r		c	c	c	-
Public	501	admr	m	m	m	-

Table 11: Overview of RTSP header fields (A-P) related to methods GET_PARAMETER, SET_PARAMETER, REDIRECT, and PLAY_NOTIFY.

Header	Where	Proxy	GPR	SPR	RDR	PNY
Range	R		o	-	o	m
Referrer	R		o	o	o	-
Request-Status	R		-	-	-	c
Require	R	r	o	o	o	-
Retry-After	3rr,503		o	o	-	-
Retry-After	413		o	o	-	-
RTP-Info	R	r	o	-	-	C

RTP-Info	r	r	c	-	-	-
Scale	G		-	-	-	c
Seek-Style	G		-	-	-	-
Server	R	r	o	o	o	o
Server	r	r	o	o	-	-
Session	R	r	o	o	o	m
Session	r	r	c	c	o	m
Speed	G		-	-	-	-
Supported	R	adrm	o	o	o	-
Supported	r	adrm	c	c	c	-
Terminate-Reason	R	r	-	-	m	-
Timestamp	R	adrm	o	o	o	-
Timestamp	c	adrm	m	m	m	-
Transport	G	mr	-	-	-	-
Unsupported	r	arm	c	c	c	-
User-Agent	R	r	m*	m*	-	-
User-Agent	r	r	m*	m*	m*	m*
Via	R	amr	o	o	o	-
Via	c	dr	m	m	m	-
WWW-Authenticate	401		m	m	m	-

Table 12: Overview of RTSP header fields (R-W) related to methods GET_PARAMETER, SET_PARAMETER, REDIRECT, and PLAY_NOTIFY.

18.1. Accept

The Accept request-header field can be used to specify certain presentation description and parameter media types [RFC6838] which are acceptable for the response to DESCRIBE and GET_PARAMETER requests.

See Section 20.2.3 for the syntax.

The asterisk "*" character is used to group media types into ranges, with "*/*" indicating all media types and "type/*" indicating all subtypes of that type. The media-range MAY include media type parameters that are applicable to that range.

Each media-range MAY be followed by one or more accept-params, beginning with the "q" parameter for indicating a relative quality factor. The first "q" parameter (if any) separates the media-range parameter(s) from the accept-params. Quality factors allow the user or user agent to indicate the relative degree of preference for that media-range, using the qvalue scale from 0 to 1 (section 3.9). The default value is q=1.

Example of use:

```
Accept: application/example ;q=0.7, application/sdp
```

Indicates that the requesting agent prefers the media type application/sdp through the default 1.0 rating but also accepts the application/example media type with a 0.7 quality rating.

If no Accept header field is present, then it is assumed that the client accepts all media types. If an Accept header field is present, and if the server cannot send a response which is acceptable according to the combined Accept field value, then the server SHOULD send a 406 (not acceptable) response.

18.2. Accept-Credentials

The Accept-Credentials header is a request-header used to indicate to any trusted intermediary how to handle further secured connections to proxies or servers. See Section 19 for the usage of this header. It MUST NOT be included in server to client requests.

In a request the header MUST contain the method (User, Proxy, or Any) for approving credentials selected by the requester. The method MUST NOT be changed by any proxy, unless it is "Proxy" when a proxy MAY change it to "user" to take the role of user approving each further hop. If the method is "User" the header contains zero or more of

credentials that the client accepts. The header may contain zero credentials in the first RTSP request to a RTSP server when using the "User" method. This is because the client has not yet received any credentials to accept. Each credential MUST consist of one URI identifying the proxy or server, the hash algorithm identifier, and the hash over that agent's ASN.1 distinguished encoding rules (DER) encoded certificate [RFC5280] in BASE64, according to Section 4 of [RFC4648] and where the padding bits are set to zero. All RTSP clients and proxies MUST implement the SHA-256[FIPS-pub-180-2] algorithm for computation of the hash of the DER encoded certificate. The SHA-256 algorithm is identified by the token "sha-256".

The intention with allowing for other hash algorithms is to enable the future retirement of algorithms that are not implemented somewhere else than here. Thus the definition of future algorithms for this purpose is intended to be extremely limited. A feature tag can be used to ensure that support for the replacement algorithm exists.

Example:

```
Accept-Credentials:User
  "rtsp://proxy2.example.com/";sha-256;exaIl9VMbQMOFGClx5rXnPJKVNI=,
  "rtsp://server.example.com/";sha-256;lurbjj5khbB0NhIuOXtt4bBRH1M=
```

18.3. Accept-Encoding

The Accept-Encoding request-header field is similar to Accept, but restricts the content-codings (see Section 18.15), i.e., transformation codings of the message body, such as gzip compression, that are acceptable in the response.

A server tests whether a content-coding is acceptable, according to an Accept-Encoding field, using these rules:

1. If the content-coding is one of the content-codings listed in the Accept-Encoding field, then it is acceptable, unless it is accompanied by a qvalue of 0. (As defined in [H3.9], a qvalue of 0 means "not acceptable.")
2. The special "*" symbol in an Accept-Encoding field matches any available content-coding not explicitly listed in the header field.
3. If multiple content-codings are acceptable, then the acceptable content-coding with the highest non-zero qvalue is preferred.

4. The "identity" content-coding is always acceptable, i.e., no transformation at all, unless specifically refused because the Accept-Encoding field includes "identity;q=0", or because the field includes "*;q=0" and does not explicitly include the "identity" content-coding. If the Accept-Encoding field-value is empty, then only the "identity" encoding is acceptable.

If an Accept-Encoding field is present in a request, and if the server cannot send a response which is acceptable according to the Accept-Encoding header, then the server SHOULD send an error response with the 406 (Not Acceptable) status code.

If no Accept-Encoding field is present in a request, the server MAY assume that the client will accept any content coding. In this case, if "identity" is one of the available content-codings, then the server SHOULD use the "identity" content-coding, unless it has additional information that a different content-coding is meaningful to the client.

18.4. Accept-Language

The Accept-Language request-header field is similar to Accept, but restricts the set of natural languages that are preferred as a response to the request. Note that the language specified applies to the presentation description and any reason phrases, but not the media content.

A language tag identifies a natural language spoken, written, or otherwise conveyed by human beings for communication of information to other human beings. Computer languages are explicitly excluded. The syntax and registry of RTSP 2.0 language tags is the same as that defined by [RFC5646].

Each language-range MAY be given an associated quality value which represents an estimate of the user's preference for the languages specified by that range. The quality value defaults to "q=1". For example:

Accept-Language: da, en-gb;q=0.8, en;q=0.7

would mean: "I prefer Danish, but will accept British English and other types of English." A language-range matches a language-tag if it exactly equals the full tag, or if it exactly equals a prefix of the tag, i.e., the primary-tag in the ABNF, such that the character following primary-tag is "-". The special range "*", if present in the Accept-Language field, matches every tag not matched by any other range present in the Accept-Language field.

Note: This use of a prefix matching rule does not imply that language tags are assigned to languages in such a way that it is always true that if a user understands a language with a certain tag, then this user will also understand all languages with tags for which this tag is a prefix. The prefix rule simply allows the use of prefix tags if this is the case.

In the process of selecting a language, each language-tag is assigned a qualification factor, i.e., if a language being supported by the client is actually supported by the server and what "preference" level the language achieves. The quality value (q-value) of the longest language-range in the field that matches the language-tag is assigned as the qualification factor for a particular language-tag. If no language-range in the field matches the tag, the language qualification factor assigned is 0. If no Accept-Language header is present in the request, the server SHOULD assume that all languages are equally acceptable. If an Accept-Language header is present, then all languages which are assigned a qualification factor greater than 0 are acceptable.

18.5. Accept-Ranges

The Accept-Ranges general-header field allows indication of the format supported in the Range header. The client MUST include the header in SETUP requests to indicate which formats are acceptable when received in PLAY and PAUSE responses, and REDIRECT requests. The server MUST include the header in SETUP and 456 error responses to indicate the formats supported for the resource indicated by the request URI. The header MAY be included in GET_PARAMETER request and response pairs. The GET_PARAMETER request MUST contain a Session header to identify the session context the request is related to. The requester and responder will indicate their capabilities regarding Range formats respectively.

Accept-Ranges: npt, smpte, clock

The syntax is defined in Section 20.2.3.

18.6. Allow

The Allow message-body header field lists the methods supported by the resource identified by the Request-URI. The purpose of this field is to inform the recipient of the complete set of valid methods associated with the resource. An Allow header field MUST be present in a 405 (Method Not Allowed) response. The Allow header MUST also be present in all OPTIONS responses where the content of the header will not include exactly the same methods as listed in the Public header.

The Allow message-body header MUST also be included in SETUP and DESCRIBE responses, if the methods allowed for the resource are different from the complete set of methods defined in this memo.

Example of use:

Allow: SETUP, PLAY, SET_PARAMETER, DESCRIBE

18.7. Authentication-Info

The Authentication-Info response-header is used by the server to communicate some information regarding the successful authentication in the response message. This usage of this header is specified in [RFC2617] with some RTSP clarification in Section 19.1. This header MUST only be used in response messages related to client to server requests.

18.8. Authorization

An RTSP client that wishes to authenticate itself with a server using authentication mechanism from HTTP [RFC2617], usually, but not necessarily, after receiving a 401 response, does so by including an Authorization request-header field with the request. The Authorization field value consists of credentials containing the authentication information of the user agent for the realm of the resource being requested. This header MUST only be used in client to server requests.

If a request is authenticated and a realm specified, the same credentials SHOULD be valid for all other requests within this realm (assuming that the authentication scheme itself does not require otherwise, such as credentials that vary according to a challenge value or using synchronized clocks). Each client to server request MUST be individually authorized by including the Authorization header with the information.

When a shared cache (see Section 16) receives a request containing an Authorization field, it MUST NOT return the corresponding response as a reply to any other request, unless one of the following specific exceptions holds:

1. If the response includes the "max-age" cache-control directive, the cache MAY use that response in replying to a subsequent request. But (if the specified maximum age has passed) a proxy cache MUST first revalidate it with the origin server, using the request-headers from the new request to allow the origin server to authenticate the new request. (This is the defined behavior

for max-age.) If the response includes "max-age=0", the proxy MUST always revalidate it before re-using it.

2. If the response includes the "must-revalidate" cache-control directive, the cache MAY use that response in replying to a subsequent request. But if the response is stale, all caches MUST first revalidate it with the origin server, using the request-headers from the new request to allow the origin server to authenticate the new request.
3. If the response includes the "public" cache-control directive, it MAY be returned in reply to any subsequent request.

18.9. Bandwidth

The Bandwidth request-header field describes the estimated bandwidth available to the client, expressed as a positive integer and measured in kilobits per second. The bandwidth available to the client may change during an RTSP session, e.g., due to mobility, congestion, etc.

Clients may not be able to accurately determine the available bandwidth, for example because the first hop is not a bottleneck. For example most local area networks (LAN) will not be a bottleneck if the server is not in the same LAN. Thus link speeds of WLAN or Ethernet networks are normally not a basis for estimating the available bandwidth. Cellular devices or other devices directly connected to a modem or connection enabling device may more accurately estimate the bottleneck bandwidth and what is a reasonable share of it for RTSP controlled media. The client will also need to take into account other traffic sharing the bottleneck. For example by only assigning a certain fraction to RTSP and its media streams. It is RECOMMENDED that only clients that have accurate and explicit information about bandwidth bottlenecks uses this header.

This header is not a substitute for proper congestion control. It is only a method providing an initial estimate and coarsely determines if the selected content can be delivered at all.

Example:

Bandwidth: 62360

18.10. Blocksize

The Blocksize request-header field is sent from the client to the media server asking the server for a particular media packet size. This packet size does not include lower-layer headers such as IP,

UDP, or RTP. The server is free to use a blocksize which is lower than the one requested. The server MAY truncate this packet size to the closest multiple of the minimum, media-specific block size, or override it with the media-specific size if necessary. The block size MUST be a positive decimal number, measured in octets. The server only returns an error (4xx) if the value is syntactically invalid.

18.11. Cache-Control

The Cache-Control general-header field is used to specify directives that MUST be obeyed by all caching mechanisms along the request/response chain.

Cache directives MUST be passed through by a proxy or gateway application, regardless of their significance to that application, since the directives may be applicable to all recipients along the request/response chain. It is not possible to specify a cache-directive for a specific cache.

Cache-Control should only be specified in a DESCRIBE, GET_PARAMETER, SET_PARAMETER and SETUP request and its response. Note: Cache-Control does not govern just the caching of responses as for HTTP, instead it also applies to the media stream identified by the SETUP request. The RTSP requests are generally not cacheable, for further information see Section 16. Below are the descriptions of the cache directives that can be included in the Cache-Control header.

no-cache: Indicates that the media stream or RTSP response MUST NOT be cached anywhere. This allows an origin server to prevent caching even by caches that have been configured to return stale responses to client requests. Note, there is no security function preventing the caching of content.

public: Indicates that the media stream or RTSP response is cacheable by any cache.

private: Indicates that the media stream or RTSP response is intended for a single user and MUST NOT be cached by a shared cache. A private (non-shared) cache may cache the media streams.

no-transform: An intermediate cache (proxy) may find it useful to convert the media type of a certain stream. A proxy might, for example, convert between video formats to save cache space or to reduce the amount of traffic on a slow link. Serious operational problems may occur, however, when these transformations have been applied to streams intended for

certain kinds of applications. For example, applications for medical imaging, scientific data analysis and those using end-to-end authentication all depend on receiving a stream that is bit-for-bit identical to the original media stream or RTSP response. Therefore, if a response includes the no-transform directive, an intermediate cache or proxy **MUST NOT** change the encoding of the stream or response. Unlike HTTP, RTSP does not provide for partial transformation at this point, e.g., allowing translation into a different language.

only-if-cached: In some cases, such as times of extremely poor network connectivity, a client may want a cache to return only those media streams or RTSP responses that it currently has stored, and not to receive these from the origin server. To do this, the client may include the only-if-cached directive in a request. If it receives this directive, a cache **SHOULD** either respond using a cached media stream or response that is consistent with the other constraints of the request, or respond with a 504 (Gateway Timeout) status. However, if a group of caches is being operated as a unified system with good internal connectivity, such a request **MAY** be forwarded within that group of caches.

max-stale: Indicates that the client is willing to accept a media stream or RTSP response that has exceeded its expiration time. If max-stale is assigned a value, then the client is willing to accept a response that has exceeded its expiration time by no more than the specified number of seconds. If no value is assigned to max-stale, then the client is willing to accept a stale response of any age.

min-fresh: Indicates that the client is willing to accept a media stream or RTSP response whose freshness lifetime is no less than its current age plus the specified time in seconds. That is, the client wants a response that will still be fresh for at least the specified number of seconds.

must-revalidate: When the must-revalidate directive is present in a SETUP response received by a cache, that cache **MUST NOT** use the cache entry after it becomes stale to respond to a subsequent request without first revalidating it with the origin server. That is, the cache is required to do an end-to-end revalidation every time, if, based solely on the origin server's Expires, the cached response is stale.

proxy-revalidate: The proxy-revalidate directive has the same meaning as the must-revalidate directive, except that it does not apply to non-shared user agent caches. It can be used on a

response to an authenticated request to permit the user's cache to store and later return the response without needing to revalidate it (since it has already been authenticated once by that user), while still requiring proxies that service many users to revalidate each time (in order to make sure that each user has been authenticated). Note that such authenticated responses also need the public cache control directive in order to allow them to be cached at all.

max-age: When an intermediate cache is forced, by means of a **max-age=0** directive, to revalidate its own cache entry, and the client has supplied its own validator in the request, the supplied validator might differ from the validator currently stored with the cache entry. In this case, the cache MAY use either validator in making its own request without affecting semantic transparency.

However, the choice of validator might affect performance. The best approach is for the intermediate cache to use its own validator when making its request. If the server replies with 304 (Not Modified), then the cache can return its now validated copy to the client with a 200 (OK) response. If the server replies with a new message body and cache validator, however, the intermediate cache can compare the returned validator with the one provided in the client's request, using the strong comparison function. If the client's validator is equal to the origin server's, then the intermediate cache simply returns 304 (Not Modified). Otherwise, it returns the new message body with a 200 (OK) response.

18.12. Connection

The Connection general-header field allows the sender to specify options that are desired for that particular connection. It MUST NOT be communicated by proxies over further connections.

RTSP 2.0 proxies MUST parse the Connection header field before a message is forwarded and, for each connection-token in this field, remove any header field(s) from the message with the same name as the connection-token. Connection options are signaled by the presence of a connection-token in the Connection header field, not by any corresponding additional header field(s), since the additional header field may not be sent if there are no parameters associated with that connection option.

Message headers listed in the Connection header MUST NOT include end-to-end headers, such as Cache-Control.

RTSP 2.0 defines the "close" connection option for the sender to signal that the connection will be closed after completion of the response. For example, Connection: close in either the request or the response-header fields indicates that the connection SHOULD NOT be considered 'persistent' (Section 10.2) after the current request/response is complete.

The use of the connection option "close" in RTSP messages SHOULD be limited to error messages when the server is unable to recover and therefore sees it necessary to close the connection. The reason is that the client has the choice of continuing using a connection indefinitely, as long as it sends valid messages.

18.13. Connection-Credentials

The Connection-Credentials response-header is used to carry the chain of credentials for any next hop that needs to be approved by the requester. It MUST only be used in server to client responses.

The Connection-Credentials header in an RTSPresponse MUST, if included, contain the credential information (in form of a list of certificates providing the chain of certification) of the next hop that an intermediary needs to securely connect to. The header MUST include the URI of the next hop (proxy or server) and a BASE64 (according to Section 4 of [RFC4648] and where the padding bits are set to zero) encoded binary structure containing a sequence of DER encoded X.509v3 certificates [RFC5280].

The binary structure starts with the number of certificates (NR_CERTS) included as a 16 bit unsigned integer. This is followed by NR_CERTS number of 16 bit unsigned integers providing the size in octets of each DER encoded certificate. This is followed by NR_CERTS number of DER encoded X.509v3 certificates in a sequence (chain). This format is exemplified in Figure 2. The proxy or server's certificate must come first in the structure. Each following certificate must directly certify the one preceding it. Because certificate validation requires that root keys be distributed independently, the self-signed certificate which specifies the root certificate authority may optionally be omitted from the chain, under the assumption that the remote end must already possess it in order to validate it in any case.

Example:

Connection-Credentials:"rtsps://proxy2.example.com/";MIIDNTCC...

Where MIIDNTCC... is a Base64 encoding of the following structure:

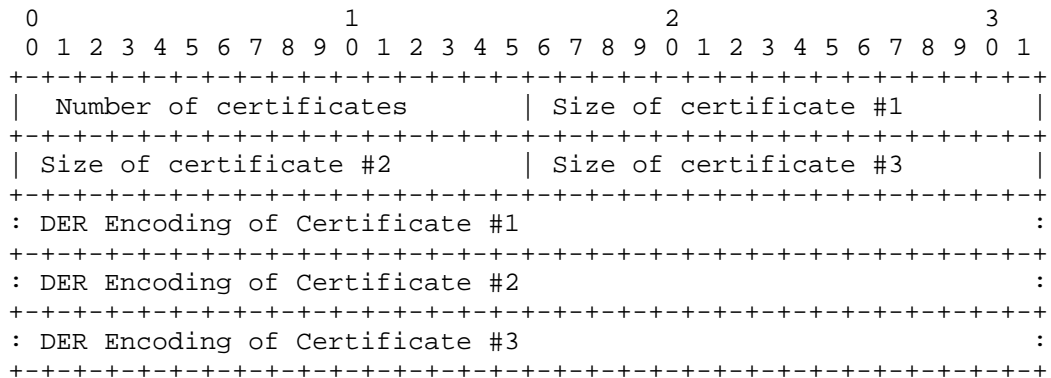


Figure 2: Connection-Credentials header's Certificate Format Example

18.14. Content-Base

The Content-Base message-body header field may be used to specify the base URI for resolving relative URIs within the message body.

Content-Base: rtsp://media.example.com/movie/twister/

If no Content-Base field is present, the base URI of an message body is defined either by its Content-Location (if that Content-Location URI is an absolute URI) or the URI used to initiate the request, in that order of precedence. Note, however, that the base URI of the contents within the message-body may be redefined within that message-body.

18.15. Content-Encoding

The Content-Encoding message-body header field is used as a modifier to the media-type. When present, its value indicates what additional content codings have been applied to the message body, and thus what decoding mechanisms must be applied in order to obtain the media-type referenced by the Content-Type header field. Content-Encoding is primarily used to allow a document to be compressed without losing the identity of its underlying media type.

The content-coding is a characteristic of the message body identified by the Request-URI. Typically, the message body is stored with this encoding and is only decoded before rendering or analogous usage. However, an RTSP proxy MAY modify the content-coding if the new coding is known to be acceptable to the recipient, unless the "no-transform" cache-control directive is present in the message.

If the content-coding of a message body is not "identity", then the message **MUST** include a Content-Encoding Message-body header that lists the non-identity content-coding(s) used.

If the content-coding of a message body in a request message is not acceptable to the origin server, the server **SHOULD** respond with a status code of 415 (Unsupported Media Type).

If multiple encodings have been applied to a message body, the content codings **MUST** be listed in the order in which they were applied, first to last from left to right. Additional information about the encoding parameters **MAY** be provided by other header fields not defined by this specification.

18.16. Content-Language

The Content-Language message-body header field describes the natural language(s) of the intended audience for the enclosed message body. Note that this might not be equivalent to all the languages used within the message body.

Language tags are mentioned in Section 18.4. The primary purpose of Content-Language is to allow a user to identify and differentiate entities according to the user's own preferred language. Thus, if the body content is intended only for a Danish-literate audience, the appropriate field is

Content-Language: da

If no Content-Language is specified, the default is that the content is intended for all language audiences. This might mean that the sender does not consider it to be specific to any natural language, or that the sender does not know for which language it is intended.

Multiple languages **MAY** be listed for content that is intended for multiple audiences. For example, a rendition of the "Treaty of Waitangi," presented simultaneously in the original Maori and English versions, would call for

Content-Language: mi, en

However, just because multiple languages are present within a message body does not mean that it is intended for multiple linguistic audiences. An example would be a beginner's language primer, such as "A First Lesson in Latin," which is clearly intended to be used by an English-literate audience. In this case, the Content-Language would properly only include "en".

Content-Language MAY be applied to any media type -- it is not limited to textual documents.

18.17. Content-Length

The Content-Length message-body header field contains the length of the message body of the RTSP message (i.e., after the double CRLF following the last header). Unlike HTTP, it MUST be included in all messages that carry a message body beyond the header portion of the RTSP message. If it is missing, a default value of zero is assumed. Any Content-Length greater than or equal to zero is a valid value.

18.18. Content-Location

The Content-Location message-body header field MAY be used to supply the resource location for the message body enclosed in the message when that body is accessible from a location separate from the requested resource's URI. A server SHOULD provide a Content-Location for the variant corresponding to the response message body; especially in the case where a resource has multiple variants associated with it, and those entities actually have separate locations by which they might be individually accessed, the server SHOULD provide a Content-Location for the particular variant which is returned.

As example, if an RTSP client performs a DESCRIBE request on a given resource, e.g., "rtsp://a.example.com/movie/Plan9FromOuterSpace", then the server may use additional information, such as the User-Agent header, to determine the capabilities of the agent. The server will then return a media description tailored to that class of RTSP agents. To indicate which specific description the agent receives the resource identifier ("rtsp://a.example.com/movie/Plan9FromOuterSpace/FullHD.sdp") is provided in Content-Location, while the description is still a valid response for the generic resource identifier. Thus enabling both debugging and cache operation as discussed below.

The Content-Location value is not a replacement for the original requested URI; it is only a statement of the location of the resource corresponding to this particular variant at the time of the request. Future requests MAY specify the Content-Location URI as the request URI if the desire is to identify the source of that particular variant. This is useful if the RTSP agent desires to verify if the resource variant is current through a conditional request.

A cache cannot assume that a message body with a Content-Location different from the URI used to retrieve it can be used to respond to later requests on that Content-Location URI. However, the Content-

Location can be used to differentiate between multiple variants retrieved from a single requested resource.

If the Content-Location is a relative URI, the relative URI is interpreted relative to the Request-URI.

Note, that Content-Location can be used in some cases to derive the base-URI for relative URI(s) present in session description formats. This needs to be taken into account when Content-Location is used. The easiest way to avoid needing to consider that issue is to include the Content-Base whenever the Content-Location is included.

Note also, when using Media Tags in conjunction with Content-Location it is important that the different versions have different MTags, even if provided under different Content-Location URIs. This as they have still been provided under the same request URI.

Note also, as in most cases the URI used in the DESCRIBE and the SETUP requests are different, the URI provided in a DESCRIBE Content-Location response can't directly be used in a SETUP request. Instead the extra step of resolving URIs combined with the media descriptions indication, like with SDP's a=control attribute.

18.19. Content-Type

The Content-Type message-body header indicates the media type of the message body sent to the recipient. Note that the content types suitable for RTSP are likely to be restricted in practice to presentation descriptions and parameter-value types.

18.20. CSeq

The CSeq general-header field specifies the sequence number (integer) for an RTSP request-response pair. This field **MUST** be present in all requests and responses. RTSP agents maintain a sequence number series for each responder to which they have an open message transport channel. For each new RTSP request an agent originates on a particular RTSP message transport the CSeq value **MUST** be incremented by one. The initial sequence number can be any number, however, it is **RECOMMENDED** to start at 0. Each sequence number series is unique between each requester and responder, i.e., the client has one series for its requests to a server and the server has another when sending requests to the client. Each requester and responder is identified by its socket address (IP address and port number), i.e., per direction of a TCP connection. Any retransmitted request **MUST** contain the same sequence number as the original, i.e., the sequence number is not incremented for retransmissions of the same request. The RTSP agent receiving requests **MUST** process the

requests arriving on a particular transport in the order of the sequence numbers. Responses are sent in the order that they are generated. The RTSP response **MUST** have the same sequence number as was present in the corresponding request. A RTSP Agent receiving a response **MAY** receive the responses out of order compared to the order of the requests it sent. Thus, the agent **MUST** use the sequence number in the response to pair it with the corresponding request.

The main purpose of the sequence number is to map responses to requests.

The requirement to use a sequence number increment of one for each new request is to support any future specification of RTSP message transport over a protocol that does not provide in order delivery or is unreliable.

The above rules relating to the initial sequence number may appear unnecessarily loose. The reason is to cater for some common behavior of existing implementations: When using multiple reliable connections in sequence it may still be easiest to use a single sequence number series for a client connecting with a particular server. Thus, the initial sequence number may be arbitrary depending on the number of previous requests. For any unreliable transport a stricter definition or other solution will be required to enable detection of any loss of the first request.

When using multiple sequential transport connections, there is no protocol mechanism to ensure in order processing as the sequence number is scoped on the individual transport connection and its five tuple. Thus, there are potential issues with opening a new transport connection to the same host for which there already exists a transport connection with outstanding requests and previously dispatched requests related to the same RTSP session.

RTSP Proxies also need to follow the above rules. This implies that proxies that aggregate requests from multiple clients onto a single transport towards a server or a next hop proxy need to renumber these requests to form a unified sequence on that transport, fulfilling the above rules. A proxy capable of fulfilling some agent's request without emitting its own request (e.g., a caching proxy that fulfils a request from its cache), also causes a need to renumber as the number of received requests with a particular target, may not be the same as the number of emitted requests towards that target agent. A proxy that needs to renumber, needs to perform the corresponding renumbering back to the original sequence number for any received response before forwarding it back to the originator of the request.

A client connected to a proxy, and using that transport to send requests to multiple servers creates a situation where it is quite likely to receive the responses out of order. This is because the proxy will establish separate transports from the proxy to the servers on which to forward the client's requests. When the responses arrive from the different servers they will be forwarded to the client in the order they arrive at the proxy and can be processed, not the order of the client's original sequence numbers. This is intentional to avoid some session's requests being blocked by another server's slow processing of requests.

18.21. Date

The Date general-header field represents the date and time at which the message was originated. The inclusion of the Date header in RTSP message follows these rules:

- o An RTSP message, sent either by the client or the server, containing a body **MUST** include a Date header, if the sending host has a clock;
- o Clients and servers are **RECOMMENDED** to include a Date header in all other RTSP messages, if the sending host has a clock;
- o If the server does not have a clock that can provide a reasonable approximation of the current time, its responses **MUST NOT** include a Date header field. In this case, this rule **MUST** be followed: Some origin server implementations might not have a clock available. An origin server without a clock **MUST NOT** assign Expires or Last-Modified values to a response, unless these values were associated with the resource by a system or user with a reliable clock. It **MAY** assign an Expires value that is known, at or before server configuration time, to be in the past (this allows "pre-expiration" of responses without storing separate Expires values for each resource).

A received message that does not have a Date header field **MUST** be assigned one by the recipient if the message will be cached by that recipient. An RTSP implementation without a clock **MUST NOT** cache responses without revalidating them on every use. An RTSP cache, especially a shared cache, **SHOULD** use a mechanism, such as Network Time Protocol (NTP) [RFC5905], to synchronize its clock with a reliable external standard.

The RTSP-date, a full date as specified by Section 3.3 of [RFC5322], sent in a Date header **SHOULD NOT** represent a date and time subsequent to the generation of the message. It **SHOULD** represent the best available approximation of the date and time of message generation,

unless the implementation has no means of generating a reasonably accurate date and time. In theory, the date ought to represent the moment just before the message body is generated. In practice, the date can be generated at any time during the message origination without affecting its semantic value.

Note: The RTSP 2.0 date format is defined to be the RFC 5322 full date format. This format is more flexible than the RFC 1123 date format used by RTSP 1.0. Thus implementations should use single spaces as recommended by RFC 5322 as separators and support receiving the obsolete format.

Further Note that the syntax allow for a comment to be added at the end of the date.

[RFC Editor please remove this note in brackets: Prior to version 37 of the draft, rfc2326bis envisaged sticking with the RFC 1123 format.]

18.22. Expires

The Expires message-body header field gives a date and time after which the description or media-stream should be considered stale. The interpretation depends on the method:

DESCRIBE response: The Expires header indicates a date and time after which the presentation description (body) SHOULD be considered stale.

SETUP response: The Expires header indicate a date and time after which the media stream SHOULD be considered stale.

A stale cache entry may not normally be returned by a cache (either a proxy cache or an user agent cache) unless it is first validated with the origin server (or with an intermediate cache that has a fresh copy of the message body). See Section 16 for further discussion of the expiration model.

The presence of an Expires field does not imply that the original resource will change or cease to exist at, before, or after that time.

The format is an absolute date and time as defined by RTSP-date. An example of its use is

Expires: Wed, 23 Jan 2013 15:36:52 +0000

RTSP/2.0 clients and caches MUST treat other invalid date formats, especially including the value "0", as having occurred in the past (i.e., already expired).

To mark a response as "already expired," an origin server should use an Expires date that is equal to the Date header value. To mark a response as "never expires," an origin server SHOULD use an Expires date approximately one year from the time the response is sent. RTSP/2.0 servers SHOULD NOT send Expires dates more than one year in the future.

18.23. From

The From request-header field, if given, SHOULD contain an Internet e-mail address for the human user who controls the requesting user agent. The address SHOULD be machine-usable, as defined by "mailbox" in [RFC1123].

This header field MAY be used for logging purposes and as a means for identifying the source of invalid or unwanted requests. It SHOULD NOT be used as an insecure form of access protection. The interpretation of this field is that the request is being performed on behalf of the person given, who accepts responsibility for the method performed. In particular, robot agents SHOULD include this header so that the person responsible for running the robot can be contacted if problems occur on the receiving end.

The Internet e-mail address in this field MAY be separate from the Internet host which issued the request. For example, when a request is passed through a proxy the original issuer's address SHOULD be used.

The client SHOULD NOT send the From header field without the user's approval, as it might conflict with the user's privacy interests or their site's security policy. It is strongly recommended that the user be able to disable, enable, and modify the value of this field at any time prior to a request.

18.24. If-Match

The If-Match request-header field is especially useful for ensuring the integrity of the presentation description, independent of how the presentation description was received. The presentation description can be fetched via means external to RTSP (such as HTTP) or via the DESCRIBE message. In the case of retrieving the presentation description via RTSP, the server implementation is guaranteeing the integrity of the description between the time of the DESCRIBE message and the SETUP message. By including the MTag given in or with the

session description in an If-Match header part of the SETUP request, the client ensures that resources set up are matching the description. A SETUP request with the If-Match header for which the MTag validation check fails, MUST generate a response using 412 (Precondition Failed).

This validation check is also very useful if a session has been redirected from one server to another.

18.25. If-Modified-Since

The If-Modified-Since request-header field is used with the DESCRIBE and SETUP methods to make them conditional. If the requested variant has not been modified since the time specified in this field, a description will not be returned from the server (DESCRIBE) or a stream will not be set up (SETUP). Instead, a 304 (Not Modified) response MUST be returned without any message-body.

An example of the field is:

If-Modified-Since: Sat, 29 Oct 1994 19:43:31 GMT

18.26. If-None-Match

This request-header can be used with one or several message body tags to make DESCRIBE requests conditional. A client that has one or more message bodies previously obtained from the resource, can verify that none of those entities is current by including a list of their associated message body tags in the If-None-Match header field. The purpose of this feature is to allow efficient updates of cached information with a minimum amount of transaction overhead. As a special case, the value "*" matches any current entity of the resource.

If any of the message body tags match the message body tag of the message body that would have been returned in the response to a similar DESCRIBE request (without the If-None-Match header) on that resource, or if "*" is given and any current entity exists for that resource, then the server MUST NOT perform the requested method, unless required to do so because the resource's modification date fails to match that supplied in an If-Modified-Since header field in the request. Instead, if the request method was DESCRIBE, the server SHOULD respond with a 304 (Not Modified) response, including the cache-related header fields (particularly MTag) of one of the message bodies that matched. For all other request methods, the server MUST respond with a status of 412 (Precondition Failed).

See Section 16.1.3 for rules on how to determine if two message body tags match.

If none of the message body tags match, then the server MAY perform the requested method as if the If-None-Match header field did not exist, but MUST also ignore any If-Modified-Since header field(s) in the request. That is, if no message body tags match, then the server MUST NOT return a 304 (Not Modified) response.

If the request would, without the If-None-Match header field, result in anything other than a 2xx or 304 status, then the If-None-Match header MUST be ignored. (See Section 16.1.4 for a discussion of server behavior when both If-Modified-Since and If-None-Match appear in the same request.)

The result of a request having both an If-None-Match header field and an If-Match header field is unspecified and MUST be considered an illegal request.

18.27. Last-Modified

The Last-Modified message-body header field indicates the date and time at which the origin server believes the presentation description or media stream was last modified. For the method DESCRIBE, the header field indicates the last modification date and time of the description, for SETUP that of the media stream.

An origin server MUST NOT send a Last-Modified date which is later than the server's time of message origination. In such cases, where the resource's last modification would indicate some time in the future, the server MUST replace that date with the message origination date.

An origin server SHOULD obtain the Last-Modified value of the message body as close as possible to the time that it generates the Date value of its response. This allows a recipient to make an accurate assessment of the message body's modification time, especially if the message body changes near the time that the response is generated.

RTSP servers SHOULD send Last-Modified whenever feasible.

18.28. Location

The Location response-header field is used to redirect the recipient to a location other than the Request-URI for completion of the request or identification of a new resource. For 3rr responses, the location SHOULD indicate the server's preferred URI for automatic

redirection to the resource. The field value consists of a single absolute URI.

Note: The Content-Location header field (Section 18.18) differs from Location in that the Content-Location identifies the original location of the message body enclosed in the request. It is therefore possible for a response to contain header fields for both Location and Content-Location. Also, see Section 16.2 for cache requirements of some methods.

18.29. Media-Properties

This general-header is used in SETUP response or PLAY_NOTIFY requests to indicate the media's properties that currently are applicable to the RTSP session. PLAY_NOTIFY MAY be used to modify these properties at any point. However, the client SHOULD have received the update prior to any action related to the new media properties taking effect. For aggregated sessions, the Media-Properties header will be returned in each SETUP response. The header received in the latest response is the one that applies on the whole session from this point until any future update. The header MAY be included without value in GET_PARAMETER requests to the server with a Session header included to query the current Media-Properties for the session. The responder MUST include the current session's media properties.

The media properties expressed by this header is the one applicable to all media in the RTSP session. For aggregated sessions, the header expressed the combined media-properties. As a result, aggregation of media MAY result in a change of the media properties, and thus the content of the Media-Properties header contained in subsequent SETUP responses.

The header contains a list of property values that are applicable to the currently setup media or aggregate of media as indicated by the RTSP URI in the request. No ordering is enforced within the header. Property values should be grouped into a single group that handles a particular orthogonal property. Values or groups that express multiple properties SHOULD NOT be used. The list of properties that can be expressed MAY be extended at any time. Unknown property values MUST be ignored.

This specification defines the following 4 groups and their property values:

Random Access:

Random-Access: Indicates that random access is possible. May optionally include a floating point value in seconds indicating

the longest duration between any two random access points in the media.

Beginning-Only: Seeking is limited to the beginning only.

No-Seeking: No seeking is possible.

Content Modifications:

Immutable: The content will not be changed during the life-time of the RTSP session.

Dynamic: The content may be changed based on external methods or triggers

Time-Progressing: The media accessible progresses as wallclock time progresses.

Retention:

Unlimited: Content will be retained for the duration of the life-time of the RTSP session.

Time-Limited: Content will be retained at least until the specified wallclock time. The time must be provided in the absolute time format specified in Section 4.4.3.

Time-Duration: Each individual media unit is retained for at least the specified time duration. This definition allows for retaining data with a time based sliding window. The time duration is expressed as floating point number in seconds. 0.0 is a valid value as this indicates that no data is retained in a time-progressing session.

Supported Scale:

Scales: A quoted comma separated list of one or more decimal values or ranges of scale values supported by the content in arbitrary order. A range has a start and stop value separated by a colon. A range indicates that the content supports fine grained selection of scale values. Fine grained allows for steps at least as small as one tenth of a scale value. A content is considered to support fine grained selection when the server in response to a given scale value can produce content with an actual scale that is less than 1 tenth of scale unit, i.e., 0.1, from the requested value. Negative values are supported. The value 0 has no meaning and MUST NOT be used.

Examples of this header for on-demand content and a live stream without recording are:

On-demand:

Media-Properties: Random-Access=2.5, Unlimited, Immutable,
 Scales="-20, -10, -4, 0.5:1.5, 4, 8, 10, 15, 20"

Live stream without recording/timeshifting:

Media-Properties: No-Seeking, Time-Progressing, Time-Duration=0.0

18.30. Media-Range

The Media-Range general-header is used to give the range of the media at the time of sending the RTSP message. This header MUST be included in SETUP response, and PLAY and PAUSE response for media that are Time-Progressing, and PLAY and PAUSE response after any change for media that are Dynamic, and in PLAY_NOTIFY request that are sent due to Media-Property-Update. Media-Range header without any range specifications MAY be included in GET_PARAMETER requests to the server to request the current range. The server MUST in this case include the current range at the time of sending the response.

The header MUST include range specifications for all time formats supported for the media, as indicated in Accept-Ranges header (Section 18.5) when setting up the media. The server MAY include more than one range specification of any given time format to indicate media that has non-continuous range. The range specifications SHALL be ordered with the range with the lowest value or earliest start time first, followed by ranges with increasingly higher values or later start time.

For media that has the Time-Progressing property, the Media-Range values will only be valid for the particular point in time when it was issued. As wallclock progresses so will also the media range. However, it shall be assumed that media time progresses in direct relationship to wallclock time (with the exception of clock skew) so that a reasonably accurate estimation of the media range can be calculated.

18.31. MTag

The MTag response-header MAY be included in DESCRIBE, GET_PARAMETER or SETUP responses. The message body tags (Section 4.6) returned in a DESCRIBE response, and the one in SETUP refers to the presentation, i.e., both the returned session description and the media stream. This allows for verification that one has the right session description to a media resource at the time of the SETUP request.

However, it has the disadvantage that a change in any of the parts results in invalidation of all the parts.

If the MTag is provided both inside the message body, e.g., within the "a=mtag" attribute in SDP, and in the response message, then both tags MUST be identical. It is RECOMMENDED that the MTag is primarily given in the RTSP response message, to ensure that caches can use the MTag without requiring content inspection. However, for session descriptions that are distributed outside of RTSP, for example using HTTP, etc. it will be necessary to include the message body tag in the session description as specified in Appendix D.1.9.

SETUP and DESCRIBE requests can be made conditional upon the MTag using the headers If-Match (Section 18.24) and If-None-Match (Section 18.26).

18.32. Notify-Reason

The Notify-Reason response-header is solely used in the PLAY_NOTIFY method. It indicates the reason why the server has sent the asynchronous PLAY_NOTIFY request (see Section 13.5).

18.33. Pipelined-Requests

The Pipelined-Requests general-header is used to indicate that a request is to be executed in the context created by a previous request(s). The primary usage of this header is to allow pipelining of SETUP requests so that any additional SETUP request after the first one does not need to wait for the session ID to be sent back to the requesting agent. The header contains a unique identifier that is scoped by the persistent connection used to send the requests.

Upon receiving a request with the Pipelined-Requests the responding agent MUST look up if there exists a binding between this Pipelined-Requests identifier for the current persistent connection and an RTSP session ID. If that exists then the received request is processed the same way as if it contained the Session header with the found session ID. If there does not exist a mapping and no Session header is included in the request, the responding agent MUST create a binding upon the successful completion of a session creating request, i.e., SETUP. A binding MUST NOT be created, if the request failed to create an RTSP session. In case the request contains both a Session header and the Pipelined-Requests header the Pipelined-Requests MUST be ignored.

Note: Based on the above definition at least the first request containing a new unique Pipelined-Requests will be required to be a SETUP request (unless the protocol is extended with new methods of

creating a session). After that first one, additional SETUP requests or requests of any type using the RTSP session context may include the Pipelined-Requests header.

When responding to any request that contained the Pipelined-Requests header the server MUST also include the Session header when a binding to a session context exists. An RTSP agent that knows the session identifier SHOULD NOT use the Pipelined-Requests header in any request and only use the Session header. This as the Session identifier is persistent across transport contexts, like TCP connections, which the Pipelined-Requests identifier is not.

The RTSP agent sending the request with a Pipelined-Requests header has the responsibility for using a unique and previously unused identifier within the transport context. Currently only a TCP connection is defined as such transport context. A server MUST delete the Pipelined-Requests identifier and its binding to a session upon the termination of that session. Despite the previous mandate, RTSP agents are RECOMMENDED to not reuse identifiers to allow for better error handling and logging.

RTSP Proxies may need to translate Pipelined-Requests identifier values from incoming requests to outgoing to allow for aggregation of requests onto a persistent connection.

18.34. Proxy-Authenticate

The Proxy-Authenticate response-header field MUST be included as part of a 407 (Proxy Authentication Required) response. The field value consists of a challenge that indicates the authentication scheme and parameters applicable to the proxy for this Request-URI.

The HTTP access authentication process is described in [RFC2617]. Unlike WWW-Authenticate, the Proxy-Authenticate header field applies only to the current connection and SHOULD NOT be passed on to downstream agents. This header MUST only be used in response messages related to client to server requests.

18.35. Proxy-Authentication-Info

The Proxy-Authentication-Info response-header is used by the proxy to communicate some information regarding the successful authentication to the proxy in the message response. The content and usage of this header is described in the HTTP access authentication [RFC2617] that is also used by RTSP and clarified in Section 19.1. This header MUST only be used in response messages related to client to server requests. This header has hop by hop scope.

18.36. Proxy-Authorization

The Proxy-Authorization request-header field allows the client to identify itself (or its user) to a proxy which requires authentication. The Proxy-Authorization field value consists of credentials containing the authentication information of the user agent for the proxy and/or realm of the resource being requested.

The HTTP access authentication process is described in [RFC2617]. Unlike Authorization, the Proxy-Authorization header field applies only to the next hop proxy. This header MUST only be used in client to server requests.

18.37. Proxy-Require

The Proxy-Require request-header field is used to indicate proxy-sensitive features that MUST be supported by the proxy. Any Proxy-Require header features that are not supported by the proxy MUST be negatively acknowledged by the proxy to the client using the Unsupported header. The proxy MUST use the 551 (Option Not Supported) status code in the response. Any feature-tag included in the Proxy-Require does not apply to the end-point (server or client). To ensure that a feature is supported by both proxies and servers the tag needs to be included in also a Require header.

See Section 18.43 for more details on the mechanics of this message and a usage example. See discussion in the proxies section (Section 15.1) about when to consider that a feature requires proxy support.

Example of use:

Proxy-Require: play.basic

18.38. Proxy-Supported

The Proxy-Supported general-header field enumerates all the extensions supported by the proxy using feature-tags. The header carries the intersection of extensions supported by the forwarding proxies. The Proxy-Supported header MAY be included in any request by a proxy. It MUST be added by any proxy if the Supported header is present in a request. When present in a request, the receiver MUST in the response copy the received Proxy-Supported header.

The Proxy-Supported header field contains a list of feature-tags applicable to proxies, as described in Section 4.5. The list is the intersection of all feature-tags understood by the proxies. To achieve an intersection, the proxy adding the Proxy-Supported header

includes all proxy feature-tags it understands. Any proxy receiving a request with the header, MUST check the list and removes any feature-tag(s) it does not support. A Proxy-Supported header present in the response MUST NOT be modified by the proxies. These feature tags are the ones the proxy chain support in general, and is not specific to the request resource.

Example:

```
C->P1: OPTIONS rtsp://example.com/ RTSP/2.0
      Supported: foo, bar, blech
      User-Agent: PhonyClient/1.2

P1->P2: OPTIONS rtsp://example.com/ RTSP/2.0
      Supported: foo, bar, blech
      Proxy-Supported: proxy-foo, proxy-bar, proxy-blech
      Via: 2.0 pro.example.com

P2->S:  OPTIONS rtsp://example.com/ RTSP/2.0
      Supported: foo, bar, blech
      Proxy-Supported: proxy-foo, proxy-blech
      Via: 2.0 pro.example.com, 2.0 prox2.example.com

S->C:  RTSP/2.0 200 OK
      Supported: foo, bar, baz
      Proxy-Supported: proxy-foo, proxy-blech
      Public: OPTIONS, SETUP, PLAY, PAUSE, TEARDOWN
      Via: 2.0 pro.example.com, 2.0 prox2.example.com
```

18.39. Public

The Public response-header field lists the set of methods supported by the response sender. This header applies to the general capabilities of the sender and its only purpose is to indicate the sender's capabilities to the recipient. The methods listed may or may not be applicable to the Request-URI; the Allow header field (Section 18.6) MAY be used to indicate methods allowed for a particular URI.

Example of use:

```
Public: OPTIONS, SETUP, PLAY, PAUSE, TEARDOWN
```

In the event that there are proxies between the sender and the recipient of a response, each intervening proxy MUST modify the Public header field to remove any methods that are not supported via that proxy. The resulting Public header field will contain an

intersection of the sender's methods and the methods allowed through by the intervening proxies.

In general, proxies should allow all methods to transparently pass through from the sending RTSP agent to the receiving RTSP agent, but there may be cases where this is not desirable for a given proxy. Modification of the Public response-header field by the intervening proxies ensures that the request sender gets an accurate response indicating the methods that can be used on the target agent via the proxy chain.

18.40. Range

The Range general-header specifies a time range in PLAY (Section 13.4), PAUSE (Section 13.6), SETUP (Section 13.3), REDIRECT (Section 13.10), and PLAY_NOTIFY (Section 13.5) requests and responses. It MAY be included in GET_PARAMETER requests from the client to the server with only a Range format and no value to request the current media position, whether the session is in Play or Ready state in the included format. The server SHALL, if supporting the range format, respond with the current playing point or pause point as the start of the range. If an explicit stop point was used in the previous PLAY request, then that value shall be included as stop point. Note that if the server is currently under any type of media playback manipulation affecting the interpretation of Range, like Scale, that is also required to be included in any GET_PARAMETER response to provide complete information.

The range can be specified in a number of units. This specification defines smpte (Section 4.4.1), npt (Section 4.4.2), and clock (Section 4.4.3) range units. While octet ranges (Byte Ranges) [H14.35.1] and other extended units MAY be used, their behavior is unspecified since they are not normally meaningful in RTSP. Servers supporting the Range header MUST understand the NPT range format and SHOULD understand the SMPTE range format. If the Range header is sent in a time format that is not understood, the recipient SHOULD return 456 (Header Field Not Valid for Resource) and include an Accept-Ranges header indicating the supported time formats for the given resource.

Example:

Range: clock=19960213T143205Z-

The Range header contains a range of one single range format. A range is a half-open interval with a start and an end point, including the start point, but excluding the end point. A range may either be fully specified with explicit values for start point and

end point, or have either start or end point be implicit. An implicit start point indicates the session's pause point, and if no pause point is set the start of the content. An implicit end point indicates the end of the content. The usage of both implicit start and end point is not allowed in the same range header, however, the exclusion of the range header has that meaning, i.e., from pause point (or start) until end of content.

Regarding the half-open intervals; a range of A-B starts exactly at time A, but ends just before B. Only the start time of a media unit such as a video or audio frame is relevant. For example, assume that video frames are generated every 40 ms. A range of 10.0-10.1 would include a video frame starting at 10.0 or later time and would include a video frame starting at 10.08, even though it lasted beyond the interval. A range of 10.0-10.08, on the other hand, would exclude the frame at 10.08.

Please note the difference between NPT time scales' "now" and an implicit start value. Implicit value reference the current pause-point. While "now" is the currently ongoing time. In a time-progressing session with recording (retention for some or full time) the pause point may be 2 min into the session while now could be 1 hour into the session.

By default, range intervals increase, where the second point is larger than the first point.

Example:

Range: npt=10-15

However, range intervals can also decrease if the Scale header (see Section 18.46) indicates a negative scale value. For example, this would be the case when a playback in reverse is desired.

Example:

Scale: -1
Range: npt=15-10

Decreasing ranges are still half open intervals as described above. Thus, for range A-B, A is closed and B is open. In the above example, 15 is closed and 10 is open. An exception to this rule is the case when B=0 in a decreasing range. In this case, the range is closed on both ends, as otherwise there would be no way to reach 0 on a reverse playback for formats that have such a notion, like NPT and SMPTE.

Example:

Scale: -1
Range: npt=15-0

In this range both 15 and 0 are closed.

A decreasing range interval without a corresponding negative Scale header is not valid.

18.41. Referrer

The Referrer request-header field allows the client to specify, for the server's benefit, the address (URI) of the resource from which the Request-URI was obtained. The URI refers to that of the presentation description, typically retrieved via HTTP. The Referrer request-header allows a server to generate lists of back-links to resources for interest, logging, optimized caching, etc. It also allows obsolete or mistyped links to be traced for maintenance. The Referrer field **MUST NOT** be sent if the Request-URI was obtained from a source that does not have its own URI, such as input from the user keyboard.

If the field value is a relative URI, it **SHOULD** be interpreted relative to the Request-URI. The URI **MUST NOT** include a fragment identifier.

Because the source of a link might be private information or might reveal an otherwise private information source, it is strongly recommended that the user be able to select whether or not the Referrer field is sent. For example, a streaming client could have a toggle switch for openly/anonymously, which would respectively enable/disable the sending of Referrer and From information.

Clients **SHOULD NOT** include a Referrer header field in a (non-secure) RTSP request if the referring page was transferred with a secure protocol.

18.42. Request-Status

This request-header is used to indicate the end result for requests that take time to complete, such as PLAY (Section 13.4). It is sent in PLAY_NOTIFY (Section 13.5) with the end-of-stream reason to report how the PLAY request concluded, either in success or in failure. The header carries a reference to the request it reports on using the CSeq number for the session indicated by the Session header in the request. It provides both a numerical status code (according to Section 8.1.1) and a human readable reason phrase.

Example:

Request-Status: cseq=63 status=500 reason="Media data unavailable"

18.43. Require

The Require request-header field is used by agents to ensure that the other end-point supports features that are required in respect to this request. It can also be used to query if the other end-point supports certain features, however, the use of the Supported general-header (Section 18.51) is much more effective in this purpose. In case any of the feature-tags listed by the Require header are not supported by the server or client receiving the request, it MUST respond to the request using the error code 551 (Option Not Supported) and include the Unsupported header listing those feature-tags which are NOT supported. This header does not apply to proxies, for the same functionality in respect to proxies see Proxy-Require header (Section 18.37) with the exception of media modifying proxies. Media modifying proxies, due to their nature of handling media in a way that is very similar to a server, do need to understand also the server's features to correctly serve the client.

This is to make sure that the client-server interaction will proceed without delay when all features are understood by both sides, and only slow down if features are not understood (as in the example below). For a well-matched client-server pair, the interaction proceeds quickly, saving a round-trip often required by negotiation mechanisms. In addition, it also removes state ambiguity when the client requires features that the server does not understand.

Example (Not complete):

```
C->S:  SETUP rtsp://server.com/foo/bar/baz.rm RTSP/2.0
       CSeq: 302
       Require: funky-feature
       Funky-Parameter: funkystuff

S->C:  RTSP/2.0 551 Option not supported
       CSeq: 302
       Unsupported: funky-feature
```

In this example, "funky-feature" is the feature-tag which indicates to the client that the fictional Funky-Parameter field is required. The relationship between "funky-feature" and Funky-Parameter is not communicated via the RTSP exchange, since that relationship is an immutable property of "funky-feature" and thus should not be transmitted with every exchange.

Proxies and other intermediary devices MUST ignore this header. If a particular extension requires that intermediate devices support it, the extension should be tagged in the Proxy-Require field instead (see Section 18.37). See discussion in the proxies section (Section 15.1) about when to consider that a feature requires proxy support.

18.44. Retry-After

The Retry-After response-header field can be used with a 503 (Service Unavailable) or 553 (Proxy Unavailable) response to indicate how long the service is expected to be unavailable to the requesting client. This field MAY also be used with any 3rr (Redirection) response to indicate the minimum time the user-agent is asked to wait before issuing the redirected request. The value of this field can be either an RTSP-date or an integer number of seconds (in decimal) after the time of the response.

Example:

```
Retry-After: Fri, 31 Dec 1999 23:59:59 GMT
Retry-After: 120
```

In the latter example, the delay is 2 minutes.

18.45. RTP-Info

The RTP-Info general-header field is used to set RTP-specific parameters in the PLAY and GET_PARAMETER responses or a PLAY_NOTIFY and GET_PARAMETER requests. For streams using RTP as transport protocol the RTP-Info header SHOULD be part of a 200 response to PLAY.

The exclusion of the RTP-Info in a PLAY response for RTP transported media will result in a client needing to synchronize the media streams using RTCP. This may have negative impact as the RTCP can be lost, and does not need to be particularly timely in its arrival. Also functionality that informs the client from which packet a seek has occurred is affected.

The RTP-Info MAY be included in SETUP responses to provide synchronization information when changing transport parameters, see Section 13.3. The RTP-Info header and the Range header MAY be included in a GET_PARAMETER request from client to server without any values to request the current playback point and corresponding RTP synchronization information. When the RTP-Info header is included in a Request the Range header MUST also be included (Note, Range header only MAY be used). The server response SHALL include both the Range

header and the RTP-Info header. If the session is in Play state, then the value of the Range header SHALL be filled in with the current playback point and with the corresponding RTP-Info values. If the server is another state, no values are included in the RTP-Info header. The header is included in PLAY_NOTIFY requests with the Notify-Reason of end-of-stream to provide RTP information about the end of the stream.

The header can carry the following parameters:

url: Indicates the stream URI for which the following RTP parameters correspond, this URI MUST be the same as used in the SETUP request for this media stream. Any relative URI MUST use the Request-URI as base URI. This parameter MUST be present.

ssrc: The Synchronization source (SSRC) that the RTP timestamp and sequence number provided applies to. This parameter MUST be present.

seq: Indicates the sequence number of the first packet of the stream that is direct result of the request. This allows clients to gracefully deal with packets when seeking. The client uses this value to differentiate packets that originated before the seek from packets that originated after the seek. Note that a client may not receive the packet with the expressed sequence number, and instead packets with a higher sequence number, due to packet loss or reordering. This parameter is RECOMMENDED to be present.

rtptime: MUST indicate the RTP timestamp value corresponding to the start time value in the Range response-header, or if not explicitly given the implied start point. The client uses this value to calculate the mapping of RTP time to NPT or other media timescale. This parameter SHOULD be present to ensure inter-media synchronization is achieved. There exists no requirement that any received RTP packet will have the same RTP timestamp value as the one in the parameter used to establish synchronization.

A mapping from RTP timestamps to Network Time Protocol (NTP) format timestamps (wallclock) is available via RTCP. However, this information is not sufficient to generate a mapping from RTP timestamps to media clock time (NPT, etc.). Furthermore, in order to ensure that this information is available at the necessary time (immediately at startup or after a seek), and that it is delivered reliably, this mapping is placed in the RTSP control channel.

In order to compensate for drift for long, uninterrupted presentations, RTSP clients should additionally map NPT to NTP, using initial RTCP sender reports to do the mapping, and later reports to check drift against the mapping.

Example:

```
Range:npt=3.25-15
RTP-Info:url="rtsp://example.com/foo/audio" ssrc=0A13C760:seq=45102;
           rtptime=12345678,url="rtsp://example.com/foo/video"
           ssrc=9A9DE123:seq=30211;rtptime=29567112
```

Lets assume that Audio uses a 16kHz RTP timestamp clock and Video a 90kHz RTP timestamp clock. Then the media synchronization is depicted in the following way.

```
NPT      3.0---3.1---3.2-X-3.3---3.4---3.5---3.6
Audio                PA A
Video                V    PV
```

X: NPT time value = 3.25, from Range header.
A: RTP timestamp value for Audio from RTP-Info header (12345678).
V: RTP timestamp value for Video from RTP-Info header (29567112).
PA: RTP audio packet carrying an RTP timestamp of 12344878. Which corresponds to NPT = $(12344878 - A) / 16000 + 3.25 = 3.2$
PV: RTP video packet carrying an RTP timestamp of 29573412. Which corresponds to NPT = $(29573412 - V) / 90000 + 3.25 = 3.32$

18.46. Scale

The Scale general-header indicates the requested or used view rate for the media resource being played back. A scale value of 1 indicates normal play at the normal forward viewing rate. If not 1, the value corresponds to the rate with respect to normal viewing rate. For example, a ratio of 2 indicates twice the normal viewing rate ("fast forward") and a ratio of 0.5 indicates half the normal viewing rate. In other words, a ratio of 2 has content time increase at twice the playback time. For every second of elapsed (wallclock) time, 2 seconds of content time will be delivered. A negative value indicates reverse direction. For certain media transports this may require certain considerations to work consistent, see Appendix C.1 for description on how RTP handles this.

The transmitted data rate SHOULD NOT be changed by selection of a different scale value. The resulting bit-rate should be reasonably close to the nominal bit-rate of the content for Scale = 1. The server has to actively manipulate the data when needed to meet the bitrate constraints. Implementation of scale changes depends on the

server and media type. For video, a server may, for example, deliver only key frames or selected frames. For audio, it may time-scale the audio while preserving pitch or, less desirably, deliver fragments of audio, or completely mute the audio.

The server and content may restrict the range of scale values that it supports. The supported values are indicated by the Media-Properties header (Section 18.29). The client SHOULD only indicate request values to be supported. However, as the values may change as the content progresses a requested value may no longer be valid when the request arrives. Thus, a non-supported value in a request does not generate an error, only forces the server to choose the closest value. The response MUST always contain the actual scale value chosen by the server.

If the server does not implement the possibility to scale, it will not return a Scale header. A server supporting Scale operations for PLAY MUST indicate this with the use of the "play.scale" feature-tag.

When indicating a negative scale for a reverse playback, the Range header MUST indicate a decreasing range as described in Section 18.40.

Example of playing in reverse at 3.5 times normal rate:

```
Scale: -3.5
Range: npt=15-10
```

18.47. Seek-Style

When a client sends a PLAY request with a Range header to perform a random access to the media, the client does not know if the server will pick the first media samples or the first random access point prior to the request range. Depending on use case, the client may have a strong preference. To express this preference and provide the client with information on how the server actually acted on that preference the Seek-Style general-header is defined.

Seek-Style is a general-header that MAY be included in any PLAY request to indicate the client's preference for any media stream that has random access properties. The server MUST always include the header in any PLAY response for media with random access properties to indicate what policy was applied. A server that receives an unknown Seek-Style policy MUST ignore it and select the server default policy. A client receiving an unknown policy MUST ignore it and use the Range header and any media synchronization information as basis to determine what the server did.

This specification defines the following seek policies that may be requested (see also Section 4.7.1):

RAP: Random Access Point (RAP) is the behavior of requesting the server to locate the closest previous random access point that exists in the media aggregate and deliver from that. By requesting a RAP, media quality will be the best possible as all media will be delivered from a point where full media state can be established in the media decoder.

CoRAP: Conditional Random Access Point (CoRAP) is a variant of the above RAP behavior. This policy is primarily intended for cases where there is larger distance between the random access points in the media. CoRAP is conditioned on that there is a Random Access Point closer to the requested start point than to the current pause point. This policy assumes that the media state existing prior to the pause is usable if delivery is continued. If the client or server knows that this is not the fact the RAP policy should be used. In other words: in most cases when the client requests a start point prior to the current pause point, a valid decoding dependency chain from the media delivered prior to the pause and to the requested media unit will not exist. If the server searched to a random access point the server MUST return the CoRAP policy in the Seek-Style header and adjust the Range header to reflect the position of the picked RAP. In case the random access point is further away and the server selects to continue from the current pause point it MUST include the "Next" policy in the Seek-Style header and adjust the Range header start point to the current pause point.

First-Prior: The first-prior policy will start delivery with the media unit that has a playout time first prior to the requested time. For discrete media that would only include media units that would still be rendered at the request time. For continuous media that is media that will be rendered during the requested start time of the range.

Next: The next media units after the provided start time of the range. For continuous framed media that would mean the first next frame after the provided time. For discrete media the first unit that is to be rendered after the provided time. The main usage for this case is when the client knows it has all media up to a certain point and would like to continue delivery so that a complete non-interrupted media playback can be achieved. Example of such scenarios include switching from a broadcast/multicast delivery to a unicast based delivery. This policy MUST only be used on the client's explicit request.

Please note that these expressed preferences exist for optimizing the startup time or the media quality. The "Next" policy breaks the normal definition of the Range header to enable a client to request media with minimal overlap, although some may still occur for aggregated sessions. RAP and First-Prior both fulfill the requirement of providing media from the requested range and forward. However, unless RAP is used, the media quality for many media codecs using predictive methods can be severely degraded unless additional data is available as, for example, already buffered, or through other side channels.

18.48. Server

The Server general-header field contains information about the software used by the origin server to create or handle the request. The field can contain multiple product tokens and comments identifying the server and any significant subproducts. The product tokens are listed in order of their significance for identifying the application.

Example:

Server: PhonyServer/1.0

If the response is being forwarded through a proxy, the proxy application **MUST NOT** modify the Server response-header. Instead, it **SHOULD** include a Via field (Section 18.57). If the response is generated by the proxy, the proxy application **MUST** return the Server response-header as previously returned by the server.

18.49. Session

The Session general-header field identifies an RTSP session. An RTSP session is created by the server as a result of a successful SETUP request and in the response the session identifier is given to the client. The RTSP session exists until destroyed by a TEARDOWN, REDIRECT or timed out by the server.

The session identifier is chosen by the server (see Section 4.3) and **MUST** be returned in the SETUP response. Once a client receives a session identifier, it **MUST** be included in any request related to that session. This means that the Session header **MUST** be included in a request, using the following methods: PLAY, PAUSE, and TEARDOWN, and **MAY** be included in SETUP, OPTIONS, SET_PARAMETER, GET_PARAMETER, and REDIRECT, and **MUST NOT** be included in DESCRIBE. The Session header **MUST NOT** be included in the following methods, if these requests are pipelined and if the session identifier is not yet

known: PLAY, PAUSE, TEARDOWN, SETUP, OPTIONS SET_PARAMETER, and GET_PARAMETER.

In an RTSP response the session header MUST be included in methods, SETUP, PLAY, and PAUSE, and MAY be included in methods, TEARDOWN, and REDIRECT, and if included in the request of the following methods it MUST also be included in the response, OPTIONS, GET_PARAMETER, and SET_PARAMETER, and MUST NOT be included in DESCRIBE responses.

Note that a session identifier identifies an RTSP session across transport sessions or connections. RTSP requests for a given session can use different URIs (Presentation and media URIs). Note, that there are restrictions depending on the session which URIs that are acceptable for a given method. However, multiple "user" sessions for the same URI from the same client will require use of different session identifiers.

The session identifier is needed to distinguish several delivery requests for the same URI coming from the same client.

The response 454 (Session Not Found) MUST be returned if the session identifier is invalid.

The header MAY include a parameter for session timeout period. If not explicitly provided this value is set to 60 seconds. As this affects how often session keep-alives are needed values smaller than 30 seconds are not recommended. However, larger than default values can be useful in applications of RTSP that have inactive but established sessions for longer time periods.

60 seconds was chosen as session timeout value due to: Resulting in not too frequent keep-alive messages and having low sensitivity to variations in request response timing. If one reduces the timeout value to below 30 seconds the corresponding request response timeout becomes a significant part of the session timeout. 60 seconds also allows for reasonably rapid recovery of committed server resources in case of client failure.

18.50. Speed

The Speed general-header field requests the server to deliver specific amounts of nominal media time per unit of delivery time, contingent on the server's ability and desire to serve the media stream at the given speed. The client requests the delivery speed to be within a given range with a lower and upper bound. The server SHALL deliver at the highest possible speed within the range, but not faster than the upper-bound, for which the underlying network path can support the resulting transport data rates. As long as any speed

value within the given range can be provided the server SHALL NOT modify the media quality. Only if the server is unable to deliver media at the speed value provided by the lower bound shall it reduce the media quality.

Implementation of the Speed functionality by the server is OPTIONAL. The server can indicate its support through a feature-tag, play.speed. The lack of a Speed header in the response is an indication of lack of support of this functionality.

The speed parameter values are expressed as a positive decimal value, e.g., a value of 2.0 indicates that data is to be delivered twice as fast as normal. A speed value of zero is invalid. The range is specified in the form "lower bound - upper bound". The lower bound value may be smaller or equal to the upper bound. All speeds may not be possible to support. Therefore the server MAY modify the requested values to the closest supported. The actual supported speed MUST be included in the response. Note, however, that the use cases may vary and that Speed value ranges such as 0.7 - 0.8, 0.3-2.0, 1.0-2.5, 2.5-2.5 all have their usage.

Example:

Speed: 1.0-2.5

Use of this header changes the bandwidth used for data delivery. It is meant for use in specific circumstances where delivery of the presentation at a higher or lower rate is desired. The main use cases are buffer operations or local scale operations. Implementors should keep in mind that bandwidth for the session may be negotiated beforehand (by means other than RTSP), and therefore re-negotiation may be necessary. To perform Speed operations the server needs to ensure that the network path can support the resulting bit-rate. Thus the media transport needs to support feedback so that the server can react and adapt to the available bitrate.

18.51. Supported

The Supported general-header enumerates all the extensions supported by the client or server using feature tags. The header carries the extensions supported by the message sending client or server. The Supported header MAY be included in any request. When present in a request, the receiver MUST respond with its corresponding Supported header. Note that the Supported header is also included in 4xx and 5xx responses.

The Supported header contains a list of feature-tags, described in Section 4.5, that are understood by the client or server. These

feature tags are the ones the server or client support in general, and is not specific to the request resource.

Example:

```
C->S:  OPTIONS rtsp://example.com/ RTSP/2.0
      Supported: foo, bar, blech
      User-Agent: PhonyClient/1.2
```

```
S->C:  RTSP/2.0 200 OK
      Supported: bar, blech, baz
```

18.52. Terminate-Reason

The Terminate-Reason request-header allows the server when sending a REDIRECT or TEARDOWN request to provide a reason for the session termination and any additional information. This specification identifies three reasons for Redirections and may be extended in the future:

Server-Admin: The server needs to be shutdown for some administrative reason.

Session-Timeout: A client's session has been kept alive for extended periods of time and the server has determined that it needs to reclaim the resources associated with this session.

Internal-Error An internal error that is impossible to recover from has occurred forcing the server to terminate the session.

The Server may provide additional parameters containing information around the redirect. This specification defines the following ones.

time: Provides a wallclock time when the server will stop providing any service.

user-msg: An UTF-8 text string with a message from the server to the user. This message SHOULD be displayed to the user.

18.53. Timestamp

The Timestamp general-header describes when the agent sent the request. The value of the timestamp is of significance only to the agent and may use any timescale. The responding agent MUST echo the exact same value and MAY, if it has accurate information about this, add a floating point number indicating the number of seconds that has elapsed since it has received the request. The timestamp can be used by the agent to compute the round-trip time to the responding agent

so that it can adjust the timeout value for retransmissions when running over an unreliable protocol. It also resolves retransmission ambiguities for unreliable transport of RTSP.

Note that the present specification provides only for reliable transport of RTSP messages. The Timestamp general-header is specified in case the protocol is extended in the future to use unreliable transport.

18.54. Transport

The Transport general-header indicates which transport protocol is to be used and configures its parameters such as destination address, compression, multicast time-to-live and destination port for a single stream. It sets those values not already determined by a presentation description.

A Transport request-header MAY contain a list of transport options acceptable to the client, in the form of multiple transport specification entries. Transport specifications are comma separated, listed in decreasing order of preference. Each transport specification consists of a transport protocol identifier, followed by any number of parameters, each parameter separated by a semicolon. A Transport request-header MAY contain multiple transport specifications using the same transport protocol Identifier. The server MUST return a Transport response-header in the response to indicate the values actually chosen if any. If no transport specification is supported, no transport header is returned and the response MUST use the status code 461 (Unsupported Transport) (Section 17.4.26). In case more than one transport specification was present in the request, the server MUST return the single transport specification (transport-spec) which was actually chosen, if any. The number of transport-spec entries is expected to be limited as the client will receive guidance on what configurations that are possible from the presentation description.

The Transport header MAY also be used in subsequent SETUP requests to change transport parameters. A server MAY refuse to change parameters of an existing stream.

The transport protocol identifier defines for each transport specification which transport protocol to use and any related rules. Each transport protocol identifier defines the parameters that are required to occur; additional optional parameters MAY occur. This flexibility is provided as parameters may be different and provide different options to the RTSP Agent. A transport specification may only contain one of any given parameter within it. A parameter consists of a name and optionally a value string. Parameters MAY be

given in any order. Additionally, a transport specification may only contain either of the unicast or the multicast transport type parameter. The transport protocol identifier and all parameters need to be understood in a transport specification; if not, the transport specification MUST be ignored. An RTSP proxy of any type that uses or modifies the transport specification, e.g., access proxy or security proxy, MUST remove specifications with unknown parameters before forwarding the RTSP message. If that results in no remaining transport specification the proxy SHALL send a 461 (Unsupported Transport) (Section 17.4.26) response without any Transport header.

The Transport header is restricted to describing a single media stream. (RTSP can also control multiple streams as a single entity.) Making it part of RTSP rather than relying on a multitude of session description formats greatly simplifies designs of firewalls.

The general syntax for the transport protocol identifier is a list of slash separated tokens:

Value1/Value2/Value3...

Which for RTP transports take the form:

RTP/profile/lower-transport.

The default value for the "lower-transport" parameters is specific to the profile. For RTP/AVP, the default is UDP.

There are two different methods for how to specify where the media should be delivered for unicast transport:

dest_addr: The presence of this parameter and its values indicates the destination address or addresses (host address and port pairs for IP flows) necessary for the media transport.

No dest_addr: The lack of the dest_addr parameter indicates that the server MUST send media to the same address from which the RTSP messages originates.

The choice of method for indicating where the media is to be delivered depends on the use case. In some cases the only allowed method will be to use no explicit address indication and have the server deliver media to the source of the RTSP messages.

For Multicast there is several methods for specifying addresses but they are different in how they work compared with unicast:

dest_addr with client picked address: The address and relevant parameters, like TTL (scope), for the actual multicast group to deliver the media to. There are security implications (Section 21) with this method that need to be addressed if using this method because a RTSP server can be used as a Denial of Service (DoS) attacker on an existing multicast group.

dest_addr using Session Description Information: The information included in the transport header can all be coming from the session description, e.g., the SDP c= and m= line. This mitigates some of the security issues of the previous methods as it is the session provider that picks the multicast group and scope. The client MUST include the information if it is available in the session description.

No dest_addr: The behavior when no explicit multicast group is present in a request is not defined.

An RTSP proxy will need to take care. If the media is not desired to be routed through the proxy, the proxy will need to introduce the destination indication.

Below are the configuration parameters associated with transport:

General parameters:

unicast / multicast: This parameter is a mutually exclusive indication of whether unicast or multicast delivery will be attempted. One of the two values MUST be specified. Clients that are capable of handling both unicast and multicast transmission need to indicate such capability by including two full transport-specs with separate parameters for each.

layers: The number of multicast layers to be used for this media stream. The layers are sent to consecutive addresses starting at the dest_addr address. If the parameter is not included, it defaults to a single layer.

dest_addr: A general destination address parameter that can contain one or more address specifications. Each combination of protocol/profile/lower transport needs to have the format and interpretation of its address specification defined. For RTP/AVP/UDP and RTP/AVP/TCP, the address specification is a tuple containing a host address and port. Note, only a single destination parameter per transport spec is intended. The usage of multiple destinations to distribute a single media to multiple entities is unspecified.

The client originating the RTSP request MAY specify the destination address of the stream recipient with the host address part of the tuple. When the destination address is specified, the recipient may be a different party than the originator of the request. To avoid becoming the unwitting perpetrator of a remote-controlled denial-of-service attack, a server MUST perform security checks (see Section 21.2.1) and SHOULD log such attempts before allowing the client to direct a media stream to a recipient address not chosen by the server. Implementations cannot rely on TCP as reliable means of client identification. If the server does not allow the host address part of the tuple to be set, it MUST return 463 (Destination Prohibited).

The host address part of the tuple MAY be empty, for example ":58044", in cases when it is desired to specify only the destination port. Responses to requests including the Transport header with a dest_addr parameter SHOULD include the full destination address that is actually used by the server. The server MUST NOT remove address information present already in the request when responding unless the protocol requires it.

src_addr: A general source address parameter that can contain one or more address specifications. Each combination of protocol/profile/lower transport needs to have the format and interpretation of its address specification defined. For RTP/AVP/UDP and RTP/AVP/TCP, the address specification is a tuple containing a host address and port.

This parameter MUST be specified by the server if it transmits media packets from another address than the one RTSP messages are sent to. This will allow the client to verify source address and give it a destination address for its RTCP feedback packets, if RTP is used. The address or addresses indicated in the src_addr parameter SHOULD be used both for sending and receiving of the media stream's data packets. The main reasons are threefold: First, indicating the port and source address(s) lets the receiver know where from the packets is expected to originate. Secondly, traversal of NATs is greatly simplified when traffic is flowing symmetrically over a NAT binding. Thirdly, certain NAT traversal mechanisms, needs to know to which address and port to send so called "binding packets" from the receiver to the sender, thus creating an address binding in the NAT that the sender to receiver packet flow can use.

This information may also be available through SDP. However, since this is more a feature of transport than media initialization, the authoritative source for this information should be in the SETUP response.

mode: The mode parameter indicates the methods to be supported for this session. Currently defined valid values are "PLAY". If not provided, the default is "PLAY". The "RECORD" value was defined in RFC 2326 and is in this specification unspecified but reserved. RECORD and other values may be specified in the future.

interleaved: The interleaved parameter implies mixing the media stream with the control stream in whatever protocol is being used by the control stream, using the mechanism defined in Section 14. The argument provides the channel number to be used in the \$ block (see Section 14) and MUST be present. This parameter MAY be specified as an interval, e.g., interleaved=4-5 in cases where the transport choice for the media stream requires it, e.g., for RTP with RTCP. The channel number given in the request is only a guidance from the client to the server on what channel number(s) to use. The server MAY set any valid channel number in the response. The declared channel(s) are bi-directional, so both end-parties MAY send data on the given channel. One example of such usage is the second channel used for RTCP, where both server and client send RTCP packets on the same channel.

This allows RTP/RTCP to be handled similarly to the way that it is done with UDP, i.e., one channel for RTP and the other for RTCP.

MIKEY: This parameter is used in conjunction with transport specifications that can utilize MIKEY [RFC3830] for security context establishment. So far only the SRTP based RTP profiles SAVP and SAVPF can utilize MIKEY and this is defined in Appendix C.1.4.1. This parameter can be included both in request and response messages. The binary MIKEY message SHALL be BASE64 [RFC4648] encoded before being included in the value part of the parameter, where the encoding adheres to the definition in Section 4 of RFC 4648 and where the padding bits are set to zero.

Multicast-specific:

ttl: multicast time-to-live for IPv4. When included in requests the value indicate the TTL value that the client requests the server to use. In a response, the value actually being used by the server is returned. A server will need to consider what values that are reasonable and also the authority of the user to set this value. Corresponding functions are not needed for IPv6 as the scoping is part of the IPv6 multicast address [RFC4291].

RTP-specific:

These parameters MAY only be used if the media transport protocol is RTP.

ssrc: The ssrc parameter, if included in a SETUP response, indicates the RTP SSRC [RFC3550] value(s) that will be used by the media server for RTP packets within the stream. It is expressed as an eight digit hexadecimal value.

The ssrc parameter MUST NOT be specified in requests. The functionality of specifying the ssrc parameter in a SETUP request is deprecated as it is incompatible with the specification of RTP in RFC 3550[RFC3550]. If the parameter is included in the Transport header of a SETUP request, the server SHOULD ignore it, and choose appropriate SSRCS for the stream. The server SHOULD set the ssrc parameter in the Transport header of the response.

RTCP-mux: Use to negotiate the usage of RTP and RTCP multiplexing [RFC5761] on a single underlying transport stream / flow. The presence of this parameter in a SETUP request indicates the client's support and requires the server to use RTP and RTCP multiplexing. The client SHALL only include one transport stream in the Transport header specification. To provide the server with a choice between using RTP/RTCP multiplexing or not, two different transport header specifications must be included.

The parameters setup and connection defined below MAY only be used if the media transport protocol of the lower-level transport is connection-oriented (such as TCP). However, these parameters MUST NOT be used when interleaving data over the RTSP connection.

setup: Clients use the setup parameter on the Transport line in a SETUP request, to indicate the roles it wishes to play in a TCP connection. This parameter is adapted from [RFC4145]. The use of this parameter in RTP/AVP/TCP non-interleaved transport is discussed in Appendix C.2.2; the discussion below is limited to

syntactic issues. Clients may specify the following values for the setup parameter:

active: The client will initiate an outgoing connection.

passive: The client will accept an incoming connection.

actpass: The client is willing to accept an incoming connection or to initiate an outgoing connection.

If a client does not specify a setup value, the "active" value is assumed.

In response to a client SETUP request where the setup parameter is set to "active", a server's 2xx reply MUST assign the setup parameter to "passive" on the Transport header line.

In response to a client SETUP request where the setup parameter is set to "passive", a server's 2xx reply MUST assign the setup parameter to "active" on the Transport header line.

In response to a client SETUP request where the setup parameter is set to "actpass", a server's 2xx reply MUST assign the setup parameter to "active" or "passive" on the Transport header line.

Note that the "holdconn" value for setup is not defined for RTSP use, and MUST NOT appear on a Transport line.

connection: Clients use the connection parameter in a transport specification part of the Transport header in a SETUP request, to indicate the client's preference for either reusing an existing connection between client and server (in which case the client sets the "connection" parameter to "existing"), or requesting the creation of a new connection between client and server (in which case the client sets the "connection" parameter to "new"). Typically, clients use the "new" value for the first SETUP request for a URL, and "existing" for subsequent SETUP requests for a URL.

If a client SETUP request assigns the "new" value to "connection", the server response MUST also assign the "new" value to "connection" on the Transport line.

If a client SETUP request assigns the "existing" value to "connection", the server response MUST assign a value of "existing" or "new" to "connection" on the Transport line, at its discretion.

The default value of "connection" is "existing", for all SETUP requests (initial and subsequent).

The combination of transport protocol, profile and lower transport needs to be defined. A number of combinations are defined in the Appendix C.

Below is a usage example, showing a client advertising the capability to handle multicast or unicast, preferring multicast. Since this is a unicast-only stream, the server responds with the proper transport parameters for unicast.

```
C->S: SETUP rtsp://example.com/foo/bar/baz.rm RTSP/2.0
      CSeq: 302
      Transport: RTP/AVP;multicast;mode="PLAY",
                 RTP/AVP;unicast;dest_addr="192.0.2.5:3456"/
                 "192.0.2.5:3457";mode="PLAY"
      Accept-Ranges: npt, smpte, clock
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 302
      Date: Fri, 20 Dec 2013 10:20:32 +0000
      Session: 47112344
      Transport: RTP/AVP;unicast;dest_addr="192.0.2.5:3456"/
                 "192.0.2.5:3457";src_addr="192.0.2.224:6256"/
                 "192.0.2.224:6257";mode="PLAY"
      Accept-Ranges: npt
      Media-Properties: Random-Access=0.6, Dynamic,
                       Time-Limited=20081128T165900
```

18.55. Unsupported

The Unsupported response-header lists the features not supported by the responding RTSP agent. In the case where the feature was specified via the Proxy-Require field (Section 18.37), if there is a proxy on the path between the client and the server, the proxy MUST send a response message with a status code of 551 (Option Not Supported). The request MUST NOT be forwarded.

See Section 18.43 for a usage example.

18.56. User-Agent

The User-Agent general-header field contains information about the user agent originating the request or producing a response. This is for statistical purposes, the tracing of protocol violations, and automated recognition of user agents for the sake of tailoring

responses to avoid particular user agent limitations. User agents SHOULD include this field with requests. The field can contain multiple product tokens and comments identifying the agent and any subproducts which form a significant part of the user agent. By convention, the product tokens are listed in order of their significance for identifying the application.

Example:

User-Agent: PhonyClient/1.2

18.57. Via

The Via general-header field MUST be used by proxies to indicate the intermediate protocols and recipients between the user agent and the server on requests, and between the origin server and the client on responses. The field is intended to be used for tracking message forwards, avoiding request loops, and identifying the protocol capabilities of all senders along the request/response chain.

Multiple Via field values represents each proxy that has forwarded the message. Each recipient MUST append its information such that the end result is ordered according to the sequence of forwarding applications.

Proxies (e.g., Access Proxy or Translator Proxy) SHOULD NOT, by default, forward the names and ports of hosts within the private/protected region. This information SHOULD only be propagated if explicitly enabled. If not enabled, the via-received of any host behind the firewall/NAT SHOULD be replaced by an appropriate pseudonym for that host.

For organizations that have strong privacy requirements for hiding internal structures, a proxy MAY combine an ordered subsequence of Via header field entries with identical sent-protocol values into a single such entry. Applications MUST NOT combine entries which have different received-protocol values.

18.58. WWW-Authenticate

The WWW-Authenticate response-header field MUST be included in 401 (Unauthorized) response messages. The field value consists of at least one challenge that indicates the authentication scheme(s) and parameters applicable to the Request-URI. This header MUST only be used in response messages related to client to server requests.

The HTTP access authentication process is described in [RFC2617] with some clarification in Section 19.1. User agents are advised to take

special care in parsing the WWW-Authenticate field value as it might contain more than one challenge, or if more than one WWW-Authenticate header field is provided, the contents of a challenge itself can contain a comma-separated list of authentication parameters.

19. Security Framework

The RTSP security framework consists of two high level components: the pure authentication mechanisms based on HTTP authentication, and the message transport protection based on TLS, which is independent of RTSP. Because of the similarity in syntax and usage between RTSP servers and HTTP servers, the security for HTTP is re-used to a large extent.

19.1. RTSP and HTTP Authentication

RTSP and HTTP share common authentication schemes, and thus follow the same usage guidelines as specified in [RFC2617] with the additions for digest authentication specified below in Section 19.1.1. Servers SHOULD implement both basic and digest [RFC2617] authentication. Clients MUST implement both basic and digest authentication [RFC2617] so that a server that requires the client to authenticate can trust that the capability is present.

It should be stressed that using the HTTP authentication alone does not provide full control message security. Therefore, in environments requiring tighter security for the control messages, TLS SHOULD be used, see Section 19.2. Any RTSP message containing an Authorization header using basic authorization MUST be using a TLS connection with confidentiality protection enabled, i.e., no NULL encryption.

In cases where there is a chain of proxies between the client and the server, each proxy may individually request the client or previous proxy to authenticate itself. This is done using the Proxy-Authenticate (Section 18.34), the Proxy-Authorization (Section 18.36) and the Proxy-Authentication-Info (Section 18.35) headers. These headers are hop-by-hop headers and are only scoped to the current connection and hop. Thus if a proxy chain exists, a proxy connecting to another proxy will have to act as a client to authorize itself towards the next proxy. The WWW-Authenticate (Section 18.58), Authorization (Section 18.8) and Authentication-Info (Section 18.7) headers are end-to-end and must not be modified by proxies.

This authentication mechanism works only for client to server requests as currently defined. This leaves server to client request outside of the context of TLS based communication more vulnerable to message injection attacks on the client. Based on the server to

client methods that exist, the potential risks are various; hijacking (REDIRECT), denial of service (TEARDOWN and PLAY_NOTIFY) or attacks with uncertain results (SET_PARAMETER).

19.1.1.1. Digest Authentication

This section describes the modifications and clarifications required to apply the HTTP Digest authentication scheme to RTSP. The RTSP scheme usage is almost completely identical to that for HTTP [RFC2617]. These are based on the procedures defined for SIP 2.0 [RFC3261].

The rules for Digest authentication follow those defined in [RFC2617], with "HTTP/1.1" replaced by "RTSP/2.0" in addition to the following differences:

1. Use the ABNF specified in this document, rather than the one in [RFC2617]. Consequently the following is assured:
 - * Using the right RTSP URIs allowed in the challenge as well as in the digest.
 - * Resolved the error in the "uri" parameter of the Authorization header in [RFC2617].
2. If MTags are used then the example procedure for choosing a nonce based on Etag can work based on replacing ETag with the MTag.
3. As a clarification to the calculation of the A2 value for message integrity assurance in the Digest authentication scheme, implementers should assume, when the entity-body is empty (that is, when the RTSP messages have no message body) that the hash of the message-body resolves to the MD5 hash of an empty string, or: $H(\text{entity-body}) = \text{MD5}("") = \text{"d41d8cd98f00b204e9800998ecf8427e"}$.
4. RFC 2617 notes that a cnonce value MUST NOT be sent in an Authorization (and by extension Proxy-Authorization) header field if no qop directive has been sent. Therefore, any algorithms that have a dependency on the cnonce (including "MD5-Sess") require that the qop directive be sent. Use of the "qop" parameter is optional in RFC 2617 for the purposes of backwards compatibility with RFC 2069; since this specification defines RTSP 2.0 there is no backwards compatibility issue with mandating. Thus, all RTSP agents MUST implement qop-options.

19.2. RTSP over TLS

RTSP agents MUST implement RTSP over TLS as defined in this section and the next Section 19.3. RTSP MUST follow the same guidelines with regards to TLS [RFC5246] usage as specified for HTTP, see [RFC2818]. RTSP over TLS is separated from unsecured RTSP both on the URI level and the port level. Instead of using the "rtsp" scheme identifier in the URI, the "rtsps" scheme identifier MUST be used to signal RTSP over TLS. If no port is given in a URI with the "rtsps" scheme, port 322 MUST be used for TLS over TCP/IP.

When a client tries to setup an insecure channel to the server (using the "rtsp" URI), and the policy for the resource requires a secure channel, the server MUST redirect the client to the secure service by sending a 301 redirect response code together with the correct Location URI (using the "rtsps" scheme). A user or client MAY upgrade a non secured URI to a secured by changing the scheme from "rtsp" to "rtsps". A server implementing support for "rtsps" MUST allow this.

It should be noted that TLS allows for mutual authentication (when using both server and client certificates). Still, one of the more common ways TLS is used is to only provide server side authentication (often to avoid client certificates). TLS is then used in addition to HTTP authentication, providing transport security and server authentication, while HTTP Authentication is used to authenticate the client.

RTSP includes the possibility to keep a TCP session up between the client and server, throughout the RTSP session lifetime. It may be convenient to keep the TCP session, not only to save the extra setup time for TCP, but also the extra setup time for TLS (even if TLS uses the resume function, there will be almost two extra round trips). Still, when TLS is used, such behavior introduces extra active state in the server, not only for TCP and RTSP, but also for TLS. This may increase the vulnerability to DoS attacks.

There exists a potential security vulnerability when reusing TCP and TLS state for different resources (URIs). If two different host names point at the same IP address it can be desirable to re-use the TCP/TLS connection to that server. In that case the RTSP agent having the TCP/TLS connection MUST verify that the server certificate associated with the connection has a SubjectAltName matching the host name present in the URI for the resource an RTSP request is to be issued for.

In addition to these recommendations, Section 19.3 gives further recommendations of TLS usage with proxies.

19.3. Security and Proxies

The nature of a proxy is often to act as a "man-in-the-middle", while security is often about preventing the existence of a "man-in-the-middle". This section provides clients with the possibility to use proxies even when applying secure transports (TLS) between the RTSP agents. The TLS proxy mechanism allows for server and proxy identification using certificates. However, the client cannot be identified based on certificates. The client needs to select between using the procedure specified below or using a TLS connection directly (by-passing any proxies) to the server. The choice may be dependent on policies.

There are in general two categories of proxies, the transparent proxies (of which the client is not aware) and the non-transparent proxies (of which the client is aware). This memo specifies only non-transparent RTSP proxies, i.e., proxies visible to the RTSP client and RTSP server. An infrastructure based on proxies requires that the trust model is such that both client and servers can trust the proxies to handle the RTSP messages correctly. To be able to trust a proxy, the client and server also need to be aware of the proxy. Hence, transparent proxies cannot generally be seen as trusted and will not work well with security (unless they work only at the transport layer). In the rest of this section any reference to proxy will be to a non-transparent proxy, which inspects or manipulates the RTSP messages.

HTTP Authentication is built on the assumption of proxies and can provide user-proxy authentication and proxy-proxy/server authentication in addition to the client-server authentication.

When TLS is applied and a proxy is used, the client will connect to the proxy's address when connecting to any RTSP server. This implies that for TLS, the client will authenticate the proxy server and not the end server. Note that when the client checks the server certificate in TLS, it MUST check the proxy's identity (URI or possibly other known identity) against the proxy's identity as presented in the proxy's Certificate message.

The problem is that for a proxy accepted by the client, the proxy needs to be provided information on which grounds it should accept the next-hop certificate. Both the proxy and the user may have rules for this, and the user should have the possibility to select the desired behavior. To handle this case, the Accept-Credentials header (See Section 18.2) is used, where the client can request the proxy/proxies to relay back the chain of certificates used to authenticate any intermediate proxies as well as the server. The assumption that the proxies are viewed as trusted, gives the user a possibility to

enforce policies to each trusted proxy of whether it should accept the next agent in the chain. However, it should be noted that not all deployments will return the chain of certificates used to authenticate any intermediate proxies as well as the server. An operator of such a deployment may want to hide its topology from the client. It should be noted well that the client does not have any insight into the proxy's operation. Even if the proxy is trusted, it can still return an incomplete chain of certificates.

A proxy MUST use TLS for the next hop if the RTSP request includes a "rtsp" URI. TLS MAY be applied on intermediate links (e.g., between client and proxy, or between proxy and proxy), even if the resource and the end server are not required to use it. The chain of proxies used by a client to reach a server and their TLS sessions MUST have commensurate security. Therefore a proxy MUST, when initiating the next hop TLS connection, use the incoming TLS connections cipher suite list, only modified by removing any cipher suites that the proxy does not support. In case a proxy fails to establish a TLS connection due to cipher suite mismatch between proxy and next hop proxy or server, this is indicated using error code 472 (Failure to establish secure connection).

19.3.1. Accept-Credentials

The Accept-Credentials header can be used by the client to distribute simple authorization policies to intermediate proxies. The client includes the Accept-Credentials header to dictate how the proxy treats the server/next proxy certificate. There are currently three methods defined:

Any: which means that the proxy (or proxies) MUST accept whatever certificate is presented. This is of course not a recommended option to use, but may be useful in certain circumstances (such as testing).

Proxy: which means that the proxy (or proxies) MUST use its own policies to validate the certificate and decide whether to accept it or not. This is convenient in cases where the user has a strong trust relation with the proxy. Reasons why a strong trust relation may exist are: personal/company proxy, proxy has a out-of-band policy configuration mechanism.

User: which means that the proxy (or proxies) MUST send credential information about the next hop to the client for authorization. The client can then decide whether the proxy should accept the certificate or not. See Section 19.3.2 for further details.

If the Accept-Credentials header is not included in the RTSP request from the client, then the "Proxy" method MUST be used as default. If another method than the "Proxy" is to be used, then the Accept-Credentials header MUST be included in all of the RTSP requests from the client. This is because it cannot be assumed that the proxy always keeps the TLS state or the user's previous preference between different RTSP messages (in particular if the time interval between the messages is long).

With the "Any" and "Proxy" methods the proxy will apply the policy as defined for each method. If the policy does not accept the credentials of the next hop, the proxy MUST respond with a message using status code 471 (Connection Credentials not accepted).

An RTSP request in the direction server to client MUST NOT include the Accept-Credentials header. As for the non-secured communication, the possibility for these requests depends on the presence of a client established connection. However, if the server to client request is in relation to a session established over a TLS secured channel, it MUST be sent in a TLS secured connection. That secured connection MUST also be the one used by the last client to server request. If no such transport connection exists at the time when the server desires to send the request, the server MUST discard the message.

Further policies MAY be defined and registered, but should be done so with caution.

19.3.2. User approved TLS procedure

For the "User" method, each proxy MUST perform the following procedure for each RTSP request:

- o Setup the TLS session to the next hop if not already present (i.e., run the TLS handshake, but do not send the RTSP request).
- o Extract the peer certificate chain for the TLS session.
- o Check if a matching identity and hash of the peer certificate is present in the Accept-Credentials header. If present, send the message to the next hop, and conclude these procedures. If not, go to the next step.
- o The proxy responds to the RTSP request with a 470 or 407 response code. The 407 response code MAY be used when the proxy requires both user and connection authorization from user or client. In this message the proxy MUST include a Connection-Credentials

header, see Section 18.13 with the next hop's identity and certificate.

The client MUST upon receiving a 470 or 407 response with Connection-Credentials header take the decision on whether to accept the certificate or not (if it cannot do so, the user SHOULD be consulted). Using IP addresses in the next hop URI and certificates rather than domain names makes it very difficult for a user to determine if it should approve the next hop or not. Proxies are RECOMMENDED to use domain names to identify themselves in URIs and in the certificates. If the certificate is accepted, the client has to again send the RTSP request. In that request the client has to include the Accept-Credentials header including the hash over the DER encoded certificate for all trusted proxies in the chain.

Example:

```
C->P: SETUP rtsp://test.example.org/secret/audio RTSP/2.0
      CSeq: 2
      Transport: RTP/AVP;unicast;dest_addr="192.0.2.5:4588"/
                "192.0.2.5:4589"
      Accept-Ranges: npt, smpte, clock
      Accept-Credentials: User

P->C: RTSP/2.0 470 Connection Authorization Required
      CSeq: 2
      Connection-Credentials: "rtsp://test.example.org";
      MIIDNTCCAp...

C->P: SETUP rtsp://test.example.org/secret/audio RTSP/2.0
      CSeq: 3
      Transport: RTP/AVP;unicast;dest_addr="192.0.2.5:4588"/
                "192.0.2.5:4589"
      Accept-Credentials: User "rtsp://test.example.org";sha-256;
      dPYD7txpoGTbAqZZQJ+vaeOkYH4=
      Accept-Ranges: npt, smpte, clock

P->S: SETUP rtsp://test.example.org/secret/audio RTSP/2.0
      CSeq: 3
      Transport: RTP/AVP;unicast;dest_addr="192.0.2.5:4588"/
                "192.0.2.5:4589"
      Via: RTSP/2.0 proxy.example.org
      Accept-Credentials: User "rtsp://test.example.org";sha-256;
      dPYD7txpoGTbAqZZQJ+vaeOkYH4=
      Accept-Ranges: npt, smpte, clock
```

One implication of this process is that the connection for secured RTSP messages may take significantly more round-trip times for the

first message. A complete extra message exchange between the proxy connecting to the next hop and the client results because of the process for approval for each hop. However, if each message contains the chain of proxies that the requester accepts, the remaining message exchange should not be delayed. The procedure of including the credentials in each request rather than building state in each proxy, avoids the need for revocation procedures.

20. Syntax

The RTSP syntax is described in an Augmented Backus-Naur Form (ABNF) as defined in RFC 5234 [RFC5234]. It uses the basic definitions present in RFC 5234.

Please note that ABNF strings, e.g., "Accept", are case insensitive as specified in section 2.3 of RFC 5234.

The RTSP syntax makes use of the ISO 10646 character set in UTF-8 encoding RFC 3629 [RFC3629].

20.1. Base Syntax

RTSP header values can be folded onto multiple lines if the continuation line begins with a space or horizontal tab. All linear white space, including folding, has the same semantics as SP. A recipient MAY replace any linear white space with a single SP before interpreting the field value or forwarding the message downstream. This is intended to behave exactly as HTTP/1.1 as described in RFC 2616 [RFC2616]. The SWS construct is used when linear white space is optional, generally between tokens and separators.

To separate the header name from the rest of value, a colon is used, which, by the above rule, allows whitespace before, but no line break, and whitespace after, including a line break. The HCOLON defines this construct.

```

OCTET      = %x00-FF ; any 8-bit sequence of data
CHAR       = %x01-7F ; any US-ASCII character (octets 1 - 127)
UPALPHA    = %x41-5A ; any US-ASCII uppercase letter "A".. "Z"
LOALPHA    = %x61-7A ;any US-ASCII lowercase letter "a".. "z"
ALPHA      = UPALPHA / LOALPHA
DIGIT      = %x30-39 ; any US-ASCII digit "0".. "9"
CTL        = %x00-1F / %x7F  ; any US-ASCII control character
           ; (octets 0 - 31) and DEL (127)
CR         = %x0D ; US-ASCII CR, carriage return (13)
LF         = %x0A  ; US-ASCII LF, linefeed (10)
SP         = %x20  ; US-ASCII SP, space (32)
HT         = %x09  ; US-ASCII HT, horizontal-tab (9)
BACKSLASH  = %x5C  ; US-ASCII backslash (92)
CRLF       = CR LF

LWS        = [CRLF] 1*( SP / HT ) ; Line-breaking White Space
SWS        = [LWS] ; Separating White Space
HCOLON     = *( SP / HT ) ":" SWS
TEXT       = %x20-7E / %x80-FF  ; any OCTET except CTLs
tspecials  = "(" / ")" / "<" / ">" / "@"
           / "," / ";" / ":" / BACKSLASH / DQUOTE
           / "/" / "[" / "]" / "?" / "="
           / "{" / "}" / SP / HT
token       = 1*(%x21 / %x23-27 / %x2A-2B / %x2D-2E / %x30-39
           / %x41-5A / %x5E-7A / %x7C / %x7E)
           ; 1*<any CHAR except CTLs or tspecials>
quoted-string = ( DQUOTE *qdttext DQUOTE )
qdttext      = %x20-21 / %x23-5B / %x5D-7E / quoted-pair
           / UTF8-NONASCII
           ; No DQUOTE and no "\"
quoted-pair  = "\\\" / ( "\" DQUOTE )
ctext        = %x20-27 / %x2A-7E
           / %x80-FF  ; any OCTET except CTLs, "(" and ")"
generic-param = token [ EQUAL gen-value ]
gen-value    = token / host / quoted-string

```

```

safe           = "$" / "-" / "_" / "." / "+"
extra          = "!" / "*" / "'" / "(" / ")" / ","
rtsp-extra     = "!" / "*" / "'" / "(" / ")"

HEX            = DIGIT / "A" / "B" / "C" / "D" / "E" / "F"
               / "a" / "b" / "c" / "d" / "e" / "f"
LHEX           = DIGIT / "a" / "b" / "c" / "d" / "e" / "f"
               ; lowercase "a-f" Hex
reserved       = ";" / "/" / "?" / ":" / "@" / "&" / "="

unreserved     = ALPHA / DIGIT / safe / extra
rtsp-unreserved = ALPHA / DIGIT / safe / rtsp-extra

base64         = *base64-unit [base64-pad]
base64-unit    = 4base64-char
base64-pad     = (2base64-char "==") / (3base64-char "=")
base64-char    = ALPHA / DIGIT / "+" / "/"

SLASH          = SWS "/" SWS ; slash
EQUAL          = SWS "=" SWS ; equal
LPAREN         = SWS "(" SWS ; left parenthesis
RPAREN         = SWS ")" SWS ; right parenthesis
COMMA         = SWS "," SWS ; comma
SEMI           = SWS ";" SWS ; semicolon
COLON          = SWS ":" SWS ; colon
MINUS          = SWS "-" SWS ; minus/dash
LDQUOT        = SWS DQUOTE ; open double quotation mark
RDQUOT        = DQUOTE SWS ; close double quotation mark
RAQUOT        = ">" SWS ; right angle quote
LAQUOT        = SWS "<" ; left angle quote

TEXT-UTF8char  = %x21-7E / UTF8-NONASCII
UTF8-NONASCII = UTF8-2 / UTF8-3 / UTF8-4
UTF8-1         = <As defined in RFC 3629>
UTF8-2         = <As defined in RFC 3629>
UTF8-3         = <As defined in RFC 3629>
UTF8-4         = <As defined in RFC 3629>
UTF8-tail      = <As defined in RFC 3629>

POS-FLOAT      = 1*12DIGIT [ "." 1*9DIGIT ]
FLOAT          = [ "-" ] POS-FLOAT

```

20.2. RTSP Protocol Definition

20.2.1. Generic Protocol elements

```

RTSP-IRI      = schemes ":" IRI-rest
IRI-rest      = ihier-part [ "?" iquery ]
ihier-part    = "//" iauthority ipath-abempty
RTSP-IRI-ref  = RTSP-IRI / irelative-ref
irelative-ref = irelative-part [ "?" iquery ]
irelative-part = "//" iauthority ipath-abempty
               / ipath-absolute
               / ipath-noscheme
               / ipath-empty

iauthority    = < As defined in RFC 3987>
ipath         = ipath-abempty ; begins with "/" or is empty
               / ipath-absolute ; begins with "/" but not "//"
               / ipath-noscheme ; begins with a non-colon segment
               / ipath-rootless ; begins with a segment
               / ipath-empty ; zero characters

ipath-abempty = *( "/" isegment )
ipath-absolute = "/" [ isegment-nz *( "/" isegment ) ]
ipath-noscheme = isegment-nz-nc *( "/" isegment )
ipath-rootless = isegment-nz *( "/" isegment )
ipath-empty   = 0<ipchar>

isegment      = *ipchar [ ";" *ipchar ]
isegment-nz   = 1*ipchar [ ";" *ipchar ]
               / ";" *ipchar
isegment-nz-nc = (1*ipchar-nc [ ";" *ipchar-nc ])
               / ";" *ipchar-nc
               ; non-zero-length segment without any colon ":"
               ; No parameter (; delimited) inside path.

ipchar        = iunreserved / pct-encoded / sub-delims / ":" / "@"
ipchar-nc     = iunreserved / pct-encoded / sub-delims / "@"
               ; sub-delims is different from RFC 3987
               ; not including ";"

iquery        = < As defined in RFC 3987>
iunreserved   = < As defined in RFC 3987>
pct-encoded   = < As defined in RFC 3987>

```

```

RTSP-URI           = schemes ":" URI-rest
RTSP-REQ-URI       = schemes ":" URI-req-rest
RTSP-URI-Ref       = RTSP-URI / RTSP-Relative
RTSP-REQ-Ref       = RTSP-REQ-URI / RTSP-REQ-Rel
schemes            = "rtsp" / "rtsps" / scheme
scheme             = < As defined in RFC 3986>
URI-rest           = hier-part [ "?" query ]
URI-req-rest       = hier-part [ "?" query ]
                    ; Note fragment part not allowed in requests
hier-part          = "://" authority path-abempty

RTSP-Relative      = relative-part [ "?" query ]
RTSP-REQ-Rel       = relative-part [ "?" query ]
relative-part      = "://" authority path-abempty
                    / path-absolute
                    / path-noscheme
                    / path-empty

authority           = < As defined in RFC 3986>
query              = < As defined in RFC 3986>

path               = path-abempty      ; begins with "/" or is empty
                    / path-absolute    ; begins with "/" but not "://"
                    / path-noscheme    ; begins with a non-colon segment
                    / path-rootless    ; begins with a segment
                    / path-empty       ; zero characters

path-abempty       = *( "/" segment )
path-absolute      = "/" [ segment-nz *( "/" segment ) ]
path-noscheme      = segment-nz-nc *( "/" segment )
path-rootless      = segment-nz *( "/" segment )
path-empty         = 0<pchar>

segment            = *pchar [ ";" *pchar ]
segment-nz         = ( 1*pchar [ ";" *pchar ] ) / ( ";" *pchar )
segment-nz-nc      = ( 1*pchar-nc [ ";" *pchar-nc ] ) / ( ";" *pchar-nc )
                    ; non-zero-length segment without any colon ":"
                    ; No parameter ( ; delimited) inside path.

pchar              = unreserved / pct-encoded / sub-delims / ":" / "@"
pchar-nc           = unreserved / pct-encoded / sub-delims / "@"

sub-delims         = "!" / "$" / "&" / "'" / "(" / ")"
                    / "*" / "+" / "," / "="
                    ; sub-delims is different from RFC 3986/3987
                    ; not including ";"

```

```

smpte-range           = smpte-type [EQUAL smpte-range-spec]
                        ; See section 4.4
smpte-range-spec      = ( smpte-time "-" [ smpte-time ] )
                        / ( "-" smpte-time )
smpte-type             = "smpte" / "smpte-30-drop"
                        / "smpte-25" / smpte-type-extension
                        ; other timecodes may be added
smpte-type-extension  = "smpte" token
smpte-time             = 1*2DIGIT ":" 1*2DIGIT ":" 1*2DIGIT
                        [ ":" 1*2DIGIT [ "." 1*2DIGIT ] ]

npt-range             = "npt" [EQUAL npt-range-spec]
npt-range-spec        = ( npt-time "-" [ npt-time ] ) / ( "-" npt-time )
npt-time              = "now" / npt-sec / npt-hhmmss / npt-hhmmss-comp
npt-sec               = 1*19DIGIT [ "." 1*9DIGIT ]
npt-hhmmss            = npt-hh ":" npt-mm ":" npt-ss [ "." 1*9DIGIT ]
npt-hh                = 2*19DIGIT ; any positive number
npt-mm                = 2*2DIGIT ; 0-59
npt-ss                = 2*2DIGIT ; 0-59
npt-hhmmss-comp       = npt-hh-comp ":" npt-mm-comp ":" npt-ss-comp
                        [ "." 1*9DIGIT ] # Compatibility format
npt-hh-comp           = 1*19DIGIT ; any positive number
npt-mm-comp           = 1*2DIGIT ; 0-59
npt-ss-comp           = 1*2DIGIT ; 0-59

utc-range             = "clock" [EQUAL utc-range-spec]
utc-range-spec        = ( utc-time "-" [ utc-time ] ) / ( "-" utc-time )
utc-time              = utc-date "T" utc-clock "Z"
utc-date              = 8DIGIT
utc-clock              = 6DIGIT [ "." 1*9DIGIT ]

feature-tag           = token

session-id            = 1*256( ALPHA / DIGIT / safe )

extension-header      = header-name HCOLON header-value
header-name           = token
header-value          = *(TEXT-UTF8char / LWS)

```

20.2.2. Message Syntax

RTSP-message = Request / Response ; RTSP/2.0 messages

Request = Request-Line
 *((general-header
 / request-header
 / message-body-header) CRLF)
 CRLF
 [message-body-data]

Response = Status-Line
 *((general-header
 / response-header
 / message-body-header) CRLF)
 CRLF
 [message-body-data]

Request-Line = Method SP Request-URI SP RTSP-Version CRLF

Status-Line = RTSP-Version SP Status-Code SP Reason-Phrase CRLF

Method = "DESCRIBE"
 / "GET_PARAMETER"
 / "OPTIONS"
 / "PAUSE"
 / "PLAY"
 / "PLAY_NOTIFY"
 / "REDIRECT"
 / "SETUP"
 / "SET_PARAMETER"
 / "TEARDOWN"
 / extension-method

extension-method = token

Request-URI = "*" / RTSP-REQ-URI

RTSP-Version = "RTSP/" 1*DIGIT "." 1*DIGIT

message-body-data = 1*OCTET

Status-Code = "100" ; Continue
 / "200" ; OK
 / "301" ; Moved Permanently
 / "302" ; Found
 / "303" ; See Other
 / "304" ; Not Modified
 / "305" ; Use Proxy
 / "400" ; Bad Request
 / "401" ; Unauthorized


```

/ "402" ; Payment Required
/ "403" ; Forbidden
/ "404" ; Not Found
/ "405" ; Method Not Allowed
/ "406" ; Not Acceptable
/ "407" ; Proxy Authentication Required
/ "408" ; Request Time-out
/ "410" ; Gone
/ "412" ; Precondition Failed
/ "413" ; Request Message Body Too Large
/ "414" ; Request-URI Too Large
/ "415" ; Unsupported Media Type
/ "451" ; Parameter Not Understood
/ "452" ; reserved
/ "453" ; Not Enough Bandwidth
/ "454" ; Session Not Found
/ "455" ; Method Not Valid in This State
/ "456" ; Header Field Not Valid for Resource
/ "457" ; Invalid Range
/ "458" ; Parameter Is Read-Only
/ "459" ; Aggregate operation not allowed
/ "460" ; Only aggregate operation allowed
/ "461" ; Unsupported Transport
/ "462" ; Destination Unreachable
/ "463" ; Destination Prohibited
/ "464" ; Data Transport Not Ready Yet
/ "465" ; Notification Reason Unknown
/ "466" ; Key Management Error
/ "470" ; Connection Authorization Required
/ "471" ; Connection Credentials not accepted
/ "472" ; Failure to establish secure connection
/ "500" ; Internal Server Error
/ "501" ; Not Implemented
/ "502" ; Bad Gateway
/ "503" ; Service Unavailable
/ "504" ; Gateway Time-out
/ "505" ; RTSP Version not supported
/ "551" ; Option not supported
/ extension-code

extension-code = 3DIGIT

Reason-Phrase = 1*(TEXT-UTF8char / HT / SP)
```

```
rtsp-header      = general-header
                  / request-header
                  / response-header
                  / message-body-header

general-header    = Accept-Ranges
                  / Cache-Control
                  / Connection
                  / CSeq
                  / Date
                  / Media-Properties
                  / Media-Range
                  / Pipelined-Requests
                  / Proxy-Supported
                  / Range
                  / RTP-Info
                  / Scale
                  / Seek-Style
                  / Server
                  / Session
                  / Speed
                  / Supported
                  / Timestamp
                  / Transport
                  / User-Agent
                  / Via
                  / extension-header

request-header    = Accept
                  / Accept-Credentials
                  / Accept-Encoding
                  / Accept-Language
                  / Authorization
                  / Bandwidth
                  / Blocksize
                  / From
                  / If-Match
                  / If-Modified-Since
                  / If-None-Match
                  / Notify-Reason
                  / Proxy-Authorization
                  / Proxy-Require
                  / Referrer
                  / Request-Status
                  / Require
                  / Terminate-Reason
                  / extension-header
```

```

response-header = Authentication-Info
                  / Connection-Credentials
                  / Location
                  / MTag
                  / Proxy-Authenticate
                  / Proxy-Authentication-Info
                  / Public
                  / Retry-After
                  / Unsupported
                  / WWW-Authenticate
                  / extension-header

```

```

message-body-header = Allow
                     / Content-Base
                     / Content-Encoding
                     / Content-Language
                     / Content-Length
                     / Content-Location
                     / Content-Type
                     / Expires
                     / Last-Modified
                     / extension-header

```

20.2.3. Header Syntax

```

Accept           = "Accept" HCOLON
                  [ accept-range *(COMMA accept-range) ]
accept-range     = media-type-range [SEMI accept-params]
media-type-range = ( "*"/*"
                  / ( m-type SLASH "*" )
                  / ( m-type SLASH m-subtype )
                  ) *( SEMI m-parameter )
accept-params    = "q" EQUAL qvalue *(SEMI generic-param )
qvalue           = ( "0" [ "." *3DIGIT ] )
                  / ( "1" [ "." *3("0") ] )
Accept-Credentials = "Accept-Credentials" HCOLON cred-decision
cred-decision    = ( "User" [LWS cred-info])
                  / "Proxy"
                  / "Any"
                  / (token [LWS 1*header-value])
                  ; For future extensions
cred-info        = cred-info-data *(COMMA cred-info-data)
cred-info-data   = DQUOTE RTSP-REQ-URI DQUOTE SEMI hash-alg
                  SEMI base64
hash-alg         = "sha-256" / extension-alg
extension-alg    = token
Accept-Encoding  = "Accept-Encoding" HCOLON

```

```

[ encoding *(COMMA encoding) ]
encoding          = codings [SEMI accept-params]
codings           = content-coding / "*"
content-coding    = "identity" / token
Accept-Language   = "Accept-Language" HCOLON
                  language *(COMMA language)
language          = language-range [SEMI accept-params]
language-range    = language-tag / "*"
language-tag      = primary-tag *( "-" subtag )
primary-tag       = 1*8ALPHA
subtag            = 1*8ALPHA
Accept-Ranges     = "Accept-Ranges" HCOLON acceptable-ranges
acceptable-ranges = (range-unit *(COMMA range-unit))
range-unit        = "npt" / "smpte" / "smpte-30-drop" / "smpte-25"
                  / "clock" / extension-format
extension-format  = token
Allow             = "Allow" HCOLON Method *(COMMA Method)
Authentication-Info = "Authentication-Info" HCOLON auth-info
auth-info         = auth-info-entry *(COMMA auth-info-entry)
auth-info-entry   = nextnonce
                  / message-qop
                  / response-auth
                  / cnonce
                  / nonce-count
nextnonce         = "nextnonce" EQUAL nonce-value
response-auth     = "rspauth" EQUAL response-digest
response-digest   = DQUOTE *LHEX DQUOTE
Authorization     = "Authorization" HCOLON credentials
credentials       = basic-credential
                  / digest-credential
                  / other-response
basic-credential  = "Basic" LWS basic-credentials
basic-credentials = base64 ; Base64 encoding of user-password
user-password     = basic-username ":" password
basic-username    = *CF-TEXT
CF-TEXT          = %x20-39 / %x3B-7E / %x80-FF ; TEXT without :
password          = *TEXT
digest-credential = ("Digest" LWS digest-response)
digest-response   = dig-resp *(COMMA dig-resp)
dig-resp          = username / realm / nonce / digest-uri
                  / dresponse / algorithm / cnonce
                  / opaque / message-qop
                  / nonce-count / auth-param
username          = "username" EQUAL username-value
username-value    = quoted-string
digest-uri        = "uri" EQUAL LDQUOT digest-uri-value RDQUOT
digest-uri-value  = RTSP-REQ-URI
message-qop       = "qop" EQUAL qop-value

```

```

cnonce           = "cnonce" EQUAL cnonce-value
cnonce-value     = nonce-value
nonce-count      = "nc" EQUAL nc-value
nc-value         = 8LHEX
dresponse        = "response" EQUAL request-digest
request-digest   = LDQUOTE 32LHEX RDQUOTE
auth-param       = auth-param-name EQUAL
                  ( token / quoted-string )
auth-param-name  = token
other-response   = auth-scheme LWS auth-param
                  *(COMMA auth-param)
auth-scheme      = token

Bandwidth        = "Bandwidth" HCOLON 1*19DIGIT

Blocksize        = "Blocksize" HCOLON 1*9DIGIT

Cache-Control    = "Cache-Control" HCOLON cache-directive
                  *(COMMA cache-directive)
cache-directive  = cache-rqst-directive
                  / cache-rspns-directive

cache-rqst-directive = "no-cache"
                    / "max-stale" [EQUAL delta-seconds]
                    / "min-fresh" EQUAL delta-seconds
                    / "only-if-cached"
                    / cache-extension

cache-rspns-directive = "public"
                    / "private"
                    / "no-cache"
                    / "no-transform"
                    / "must-revalidate"
                    / "proxy-revalidate"
                    / "max-age" EQUAL delta-seconds
                    / cache-extension

cache-extension  = token [EQUAL (token / quoted-string)]
delta-seconds    = 1*19DIGIT

Connection       = "Connection" HCOLON connection-token
                  *(COMMA connection-token)
connection-token = "close" / token

Connection-Credentials = "Connection-Credentials" HCOLON cred-chain
cred-chain          = DQUOTE RTSP-REQ-URI DQUOTE SEMI base64

Content-Base      = "Content-Base" HCOLON RTSP-URI

```

```

Content-Encoding      = "Content-Encoding" HCOLON
                        content-coding *(COMMA content-coding)
Content-Language      = "Content-Language" HCOLON
                        language-tag *(COMMA language-tag)
Content-Length        = "Content-Length" HCOLON 1*19DIGIT
Content-Location      = "Content-Location" HCOLON RTSP-REQ-Ref
Content-Type          = "Content-Type" HCOLON media-type
media-type            = m-type SLASH m-subtype *(SEMI m-parameter)
m-type                = discrete-type / composite-type
discrete-type         = "text" / "image" / "audio" / "video"
                        / "application" / extension-token
composite-type        = "message" / "multipart" / extension-token
extension-token       = ietf-token / x-token
ietf-token            = token
x-token               = "x-" token
m-subtype             = extension-token / iana-token
iana-token            = token
m-parameter           = m-attribute EQUAL m-value
m-attribute           = token
m-value               = token / quoted-string

CSeq                  = "CSeq" HCOLON cseq-nr
cseq-nr               = 1*9DIGIT
Date                  = "Date" HCOLON RTSP-date
RTSP-date             = date-time ;
date-time             = <As defined in RFC 5322>
Expires               = "Expires" HCOLON RTSP-date
From                  = "From" HCOLON from-spec
from-spec             = ( name-addr / addr-spec ) *( SEMI from-param )
name-addr             = [ display-name ] LAQUOT addr-spec RAQUOT
addr-spec             = RTSP-REQ-URI / absolute-URI
absolute-URI          = < As defined in RFC 3986>
display-name          = *(token LWS) / quoted-string
from-param            = tag-param / generic-param
tag-param             = "tag" EQUAL token
If-Match              = "If-Match" HCOLON ("*" / message-tag-list)
message-tag-list      = message-tag *(COMMA message-tag)
message-tag           = [ weak ] opaque-tag
weak                  = "W/"
opaque-tag            = quoted-string
If-Modified-Since     = "If-Modified-Since" HCOLON RTSP-date
If-None-Match         = "If-None-Match" HCOLON ("*" / message-tag-list)
Last-Modified         = "Last-Modified" HCOLON RTSP-date
Location              = "Location" HCOLON RTSP-REQ-URI
Media-Properties      = "Media-Properties" HCOLON [media-prop-list]
media-prop-list       = media-prop-value *(COMMA media-prop-value)
media-prop-value      = ( "Random-Access" [EQUAL POS-FLOAT])
                        / "Beginning-Only"

```

```

/ "No-Seeking"
/ "Immutable"
/ "Dynamic"
/ "Time-Progressing"
/ "Unlimited"
/ ("Time-Limited" EQUAL utc-time)
/ ("Time-Duration" EQUAL POS-FLOAT)
/ ("Scales" EQUAL scale-value-list)
/ media-prop-ext
media-prop-ext = token [EQUAL (1*rtsp-unreserved / quoted-string)]
scale-value-list = DQUOTE scale-entry *(COMMA scale-entry) DQUOTE
scale-entry = scale-value / (scale-value COLON scale-value)
scale-value = FLOAT
Media-Range = "Media-Range" HCOLON [ranges-list]
ranges-list = ranges-spec *(COMMA ranges-spec)
MTag = "MTag" HCOLON message-tag
Notify-Reason = "Notify-Reason" HCOLON Notify-Reas-val
Notify-Reas-val = "end-of-stream"
/ "media-properties-update"
/ "scale-change"
/ Notify-Reason-extension
Notify-Reason-extension = token
Pipelined-Requests = "Pipelined-Requests" HCOLON startup-id
startup-id = 1*8DIGIT
```

```

Proxy-Authenticate = "Proxy-Authenticate" HCOLON challenge-list
challenge-list     = challenge *(COMMA challenge)
challenge          = ("Digest" LWS digest-cln *(COMMA digest-cln))
                   / ("Basic" LWS realm)
                   / other-challenge
other-challenge    = auth-scheme LWS auth-param
                   *(COMMA auth-param)
digest-cln        = realm / domain / nonce
                   / opaque / stale / algorithm
                   / qop-options / auth-param
realm              = "realm" EQUAL realm-value
realm-value        = quoted-string
domain             = "domain" EQUAL LDQUOT RTSP-REQ-Ref
                   *(1*SP RTSP-REQ-Ref ) RDQUOT
nonce              = "nonce" EQUAL nonce-value
nonce-value        = quoted-string
opaque             = "opaque" EQUAL quoted-string
stale              = "stale" EQUAL ( "true" / "false" )
algorithm          = "algorithm" EQUAL ( "MD5" / "MD5-sess" / token)
qop-options        = "qop" EQUAL LDQUOT qop-value
                   *("," qop-value) RDQUOT
qop-value          = "auth" / "auth-int" / token
Proxy-Authentication-Info = "Proxy-Authentication-Info" HCOLON auth-info
Proxy-Authorization = "Proxy-Authorization" HCOLON credentials
Proxy-Require       = "Proxy-Require" HCOLON feature-tag-list
feature-tag-list    = feature-tag *(COMMA feature-tag)
Proxy-Supported      = "Proxy-Supported" HCOLON [feature-tag-list]

Public              = "Public" HCOLON Method *(COMMA Method)

Range               = "Range" HCOLON ranges-spec

ranges-spec         = npt-range / utc-range / smpte-range
                   / range-ext
range-ext           = extension-format [EQUAL range-value]
range-value         = 1*(rtsp-unreserved / quoted-string / ":" )

Referrer            = "Referrer" HCOLON (absolute-URI / RTSP-URI-Ref)
Request-Status      = "Request-Status" HCOLON req-status-info
req-status-info     = cseq-info LWS status-info LWS reason-info
cseq-info           = "cseq" EQUAL cseq-nr
status-info         = "status" EQUAL Status-Code
reason-info         = "reason" EQUAL DQUOTE Reason-Phrase DQUOTE
Require            = "Require" HCOLON feature-tag-list

```



```

RTP-Info          = "RTP-Info" HCOLON [rtsp-info-spec
                        *(COMMA rtsp-info-spec)]
rtsp-info-spec    = stream-url 1*ssrc-parameter
stream-url        = "url" EQUAL DQUOTE RTSP-REQ-Ref DQUOTE
ssrc-parameter    = LWS "ssrc" EQUAL ssrc HCOLON
                        ri-parameter *(SEMI ri-parameter)
ri-parameter      = ("seq" EQUAL 1*5DIGIT)
                        / ("rtptime" EQUAL 1*10DIGIT)
                        / generic-param

Retry-After       = "Retry-After" HCOLON (RTSP-date / delta-seconds)
Scale             = "Scale" HCOLON scale-value
Seek-Style        = "Seek-Style" HCOLON Seek-S-values
Seek-S-values     = "RAP"
                        / "CoRAP"
                        / "First-Prior"
                        / "Next"
                        / Seek-S-value-ext
Seek-S-value-ext  = token

Server            = "Server" HCOLON ( product / comment )
                        *(LWS (product / comment))
product           = token [SLASH product-version]
product-version   = token
comment           = LPAREN *( ctext / quoted-pair) RPAREN

Session          = "Session" HCOLON session-id
                        [ SEMI "timeout" EQUAL delta-seconds ]

Speed            = "Speed" HCOLON lower-bound MINUS upper-bound
lower-bound      = POS-FLOAT
upper-bound      = POS-FLOAT

Supported         = "Supported" HCOLON [feature-tag-list]

```

```
Terminate-Reason      = "Terminate-Reason" HCOLON TR-Info
TR-Info               = TR-Reason *(SEMI TR-Parameter)
TR-Reason             = "Session-Timeout"
                      / "Server-Admin"
                      / "Internal-Error"
                      / token
TR-Parameter          = TR-time / TR-user-msg / generic-param
TR-time               = "time" EQUAL utc-time
TR-user-msg           = "user-msg" EQUAL quoted-string

Timestamp             = "Timestamp" HCOLON timestamp-value [LWS delay]
timestamp-value       = *19DIGIT [ "." *9DIGIT ]
delay                 = *9DIGIT [ "." *9DIGIT ]

Transport             = "Transport" HCOLON transport-spec
                      *(COMMA transport-spec)
transport-spec        = transport-id *trns-parameter
transport-id          = trans-id-rtp / other-trans
trans-id-rtp          = "RTP/" profile [ "/" lower-transport ]
                      ; no LWS is allowed inside transport-id
other-trans           = token *( "/" token)
```

```

profile           = "AVP" / "SAVP" / "AVPF" / "SAVPF" / token
lower-transport   = "TCP" / "UDP" / token
trns-parameter    = (SEMI ( "unicast" / "multicast" ))
                  / (SEMI "interleaved" EQUAL channel ["-" channel])
                  / (SEMI "ttl" EQUAL ttl)
                  / (SEMI "layers" EQUAL 1*DIGIT)
                  / (SEMI "ssrc" EQUAL ssrc *(SLASH ssrc))
                  / (SEMI "mode" EQUAL mode-spec)
                  / (SEMI "dest_addr" EQUAL addr-list)
                  / (SEMI "src_addr" EQUAL addr-list)
                  / (SEMI "setup" EQUAL contrans-setup)
                  / (SEMI "connection" EQUAL contrans-con)
                  / (SEMI "RTCP-mux")
                  / (SEMI "MIKEY" EQUAL MIKEY-Value)
                  / (SEMI trn-param-ext)
contrans-setup     = "active" / "passive" / "actpass"
contrans-con       = "new" / "existing"
trn-param-ext      = par-name [EQUAL trn-par-value]
par-name           = token
trn-par-value      = *(rtsp-unreserved / quoted-string)
ttl                = 1*3DIGIT ; 0 to 255
ssrc               = 8HEX
channel            = 1*3DIGIT ; 0 to 255
MIKEY-Value        = base64
mode-spec          = ( DQUOTE mode *(COMMA mode) DQUOTE )
mode               = "PLAY" / token
addr-list          = quoted-addr *(SLASH quoted-addr)
quoted-addr        = DQUOTE (host-port / extension-addr) DQUOTE
host-port          = ( host [":" port] )
                  / ( ":" port )
extension-addr     = 1*qdttext
host                = < As defined in RFC 3986>
port               = < As defined in RFC 3986>

```

```

Unsupported      = "Unsupported" HCOLON feature-tag-list
User-Agent       = "User-Agent" HCOLON ( product / comment )
                  *(LWS (product / comment))
Via              = "Via" HCOLON via-parm *(COMMA via-parm)
via-parm         = sent-protocol LWS sent-by *( SEMI via-params )
via-params       = via-ttl / via-maddr
                  / via-received / via-extension
via-ttl          = "ttl" EQUAL ttl
via-maddr        = "maddr" EQUAL host
via-received     = "received" EQUAL (IPv4address / IPv6address)
IPv4address      = < As defined in RFC 3986>
IPv6address      = < As defined in RFC 3986>
via-extension    = generic-param
sent-protocol    = protocol-name SLASH protocol-version
                  SLASH transport-prot
protocol-name    = "RTSP" / token
protocol-version = token
transport-prot   = "UDP" / "TCP" / "TLS" / other-transport
other-transport  = token
sent-by          = host [ COLON port ]

WWW-Authenticate = "WWW-Authenticate" HCOLON challenge-list

```

20.3. SDP extension Syntax

This section defines in ABNF the SDP extensions defined for RTSP. See Appendix D for the definition of the extensions in text.

```

control-attribute = "a=control:" *SP RTSP-REQ-Ref CRLF

a-range-def       = "a=range:" ranges-spec CRLF

a-mtag-def        = "a=mtag:" message-tag CRLF

```

21. Security Considerations

The security considerations and threats around RTSP and its usage can be divided into considerations around the signaling protocol itself and the issues related to the media stream delivery. However, when it comes to mitigations of security threats, a threat depending on the media stream delivery may in fact be mitigated by a mechanism in the signaling protocol.

There are several chapters and an appendix in this document that define security solutions for the protocol. These sections will be referenced when discussing the threats below. But the reader should take special notice of the Security Framework (Section 19) and the

specification of how to use SRTP and its key-mangement
(Appendix C.1.4) to achieve certain aspects of the media security.

21.1. Signaling Protocol Threats

This section focuses on issues related to the signaling protocol. Because of the similarity in syntax and usage between RTSP servers and HTTP servers, the security considerations outlined in [H15] apply also.

Specifically, please note the following:

Abuse of Server Log Information: A server is in the position to save personal data about a user's requests which might identify their media consumption patterns or subjects of interest. This information is clearly confidential in nature and its handling can be constrained by law in certain countries. RTSP servers will presumably have similar logging mechanisms to HTTP, and thus should be equally guarded in protecting the contents of those logs, thus protecting the privacy of the users of the servers. People using the RTSP protocol to provide media are responsible for ensuring that logging material is not distributed without the permission of any individuals that are identifiable by the published results.

Transfer of Sensitive Information: There is no reason to believe that information transferred in RTSP message, such as the URI and the content of headers, especially the Server, Via, Referrer and From headers, may be any less sensitive than when used in HTTP. Therefore, all of the precautions regarding the protection of data privacy and user privacy apply to implementors of RTSP clients, servers, and proxies. See [H15.1.2] for further details.

The RTSP methods defined in this document is primarily used to establish and control the delivery of the media data represented by the URI, thus the RTSP message bodies are generally less sensitive than the ones in HTTP. Where HTTP bodies could contain for example your medical records, in RTSP the sensitive video of your medical operation would be in the media stream over the media transport protocol, not in the RTSP message. Still one have to take note of what potential sensitive informative are included in the RTSP protocol. The protection of the media data is separate, can be applied directly between client and server, and is dependent on the media transport protocol in use. See Section 21.2 for further discussion. This possibility for separation of security between media resource content and the signalling protocol

mitigates the risk of exposing the media content when using hop-by-hop security for RTSP signaling using proxies (Section 19.3).

Attacks Based On File and Path Names: Though RTSP URIs are opaque handles that do not necessarily have file system semantics, it is anticipated that many implementations will translate portions of the Request-URIs directly to file system calls. In such cases, file systems SHOULD follow the precautions outlined in [H15.2], such as checking for ".." in path components.

Personal Information: RTSP clients are often privy to the same information that HTTP clients are (user name, location, etc.) and thus should be equally sensitive. See [H15.1] for further recommendations.

Privacy Issues Connected to Accept Headers: Since similar usages of the "Accept" headers exist in RTSP as in HTTP, the same caveats outlined in [H15.1.4] with regards to their use should be followed.

DNS Spoofing: Presumably, given the longer connection times typically associated with RTSP sessions relative to HTTP sessions, RTSP client DNS optimizations should be less prevalent. Nonetheless, the recommendations provided in [H15.3] are still relevant to any implementation which attempts to rely on a DNS-to-IP mapping to hold beyond a single use of the mapping.

Location Headers and Spoofing: If a single server supports multiple organizations that do not trust each another, then it MUST check the values of the Content-Location header fields in responses that are generated under control of said organizations to make sure that they do not attempt to invalidate resources over which they have no authority. ([H15.4])

In addition to the recommendations in the current HTTP specification (RFC 2616 [RFC2616], as of this writing) and also of the previous RFC 2068 [RFC2068], future HTTP specifications may provide additional guidance on security issues.

The following are added considerations for RTSP implementations.

Session hijacking: Since there is no or little relation between a transport layer connection and an RTSP session, it is possible for a malicious client to issue requests with random session identifiers which could affect other clients of an unsuspecting

server. To mitigate this the server SHALL use a large, random and non-sequential session identifier to minimize the possibility of this kind of attack. However, unless the RTSP signaling is always confidentiality protected, e.g., using TLS, an on-path attacker will be able to hijack a session. Another choice for preventing session hijacking is to use client authentication and only allow the authenticated client creating the session to access that session.

Authentication: Servers SHOULD implement both basic and digest [RFC2617] authentication. In environments requiring tighter security for the control messages, the transport layer mechanism TLS [RFC5246] SHOULD be used.

Suspicious behavior: RTSP servers upon detecting instances of behavior which is deemed a security risk SHOULD return error code 403 (Forbidden). RTSP servers SHOULD also be aware of attempts to probe the server for weaknesses and entry points and MAY arbitrarily disconnect and ignore further requests from clients which are deemed to be in violation of local security policy.

TLS through proxies: If one uses the possibility to connect TLS in multiple legs (Section 19.3) one really needs to be aware of the trust model. That procedure requires full faith and trust in all proxies, which will be identified, that one allows to connect through. They are men in the middle and have access to all that goes on over the TLS connection. Thus it is important to consider if that trust model is acceptable in the actual application. Further discussion of the actual trust model is in Section 19.3. It is important to note what difference in security properties, if any, that may exist with the used media transport protocol and its security mechanism. Using SRTP and the MIKEY based key-establishment defined in Appendix C.1.4.1, enables to media key-establishment to done end-to-end without revealing the keys to the proxies.

Resource Exhaustion: As RTSP is a stateful protocol and establishes resource usage on the server there is a clear possibility to attack the server by trying to overbook these resources to perform a denial of service attack. This attack can be both against ongoing sessions and to prevent others from establishing sessions. RTSP agents will need to have mechanisms to prevent single peers from consuming extensive amounts of resources. The methods for guarding against this are varied and depends on the agent's role and capabilities and policies. Each implementation has to carefully consider their methods and policies to mitigate this threat. For example

regarding handling of connections there are recommendations in Section 10.7.

The above threats and considerations have resulted in a set of security functions and mechanisms built into or used by the protocol. The signaling protocol relies on two security features defined in the Security Framework (Section 19) namely client authentication using HTTP authentication and TLS based transport protection of the signaling messages. Both of these mechanisms are required to be implemented by any RTSP agent.

A number of different security mitigations have been designed into the protocol and will be instantiated if the specification is implemented as written, for example by ensuring sufficient amount of entropy in the randomly generated session identifiers when not using client authentication to minimize the risk of session hijacking. When client authentication is used the protection against hijacking will be greatly improved by scoping the accessible sessions to the one this client identity has created. Some of the above threats are such that the implementation of the RTSP functionality itself needs to consider which policy and strategy it uses to mitigate them.

21.2. Media Stream Delivery Threats

The fact that RTSP establishes and controls a media stream delivery results in a set of security issues related to the media streams. This section will attempt to analyze general threats, however the choice of media stream transport protocol, such as RTP will result in some differences in threats and what mechanisms exist to mitigate them. Thus it becomes important that each specification of a new media stream transport and delivery protocol usable by RTSP requires its own security analysis. This section includes one for RTP.

The set of general threats from or by the media stream delivery itself are:

Concentrated denial-of-service attack: The protocol offers the opportunity for a remote-controlled denial-of-service (DoS) attack, where the media stream is the hammer in that DoS attack. See Section 21.2.1.

Media Confidentiality: The media delivery may contain content of any type and it is not possible in general to determine how sensitive this content is from a confidentiality point. Thus it is a strong requirement that any media delivery protocol provides a method for providing confidentiality of the actual media content. In addition to the media level confidentiality it becomes critical that no resource identifiers used in the signaling are exposed to

an attacker as they may have human understandable names, or may be also available to the attacker so they can determine the content the user was delivered. Thus the signaling protocol must also provide confidentiality protection of any information related to the media resource.

Media Integrity and Authentication: There are several reasons, such as discrediting the target, misinformation of the target, why an attacker will be interested in substituting the media stream sent out from the RTSP server with one of the attacker's creation or selection. Therefore it is important that the media protocol provides mechanisms to verify the source authentication, integrity and prevent replay attacks on the media stream.

Scope of Multicast: If RTSP is used to control the transmission of media onto a multicast network the scope of the delivery must be considered. RTSP supports the TTL Transport header parameter to indicate this scope for IPv4. IPv6 has a different mechanism for scope boundary. However, such scope control has risks, as it may be set too large and distribute media beyond the intended scope.

Below (Section 21.2.2) a protocol specific analysis of security considerations for RTP based media transport is done. In that section it is also made clear the requirements on implementing security functions for RTSP agents supporting media delivery over RTP.

21.2.1. Remote Denial of Service Attack

The attacker may initiate traffic flows to one or more IP addresses by specifying them as the destination in SETUP requests. While the attacker's IP address may be known in this case, this is not always useful in prevention of more attacks or ascertaining the attacker's identity. Thus, an RTSP server **MUST** only allow client-specified destinations for RTSP-initiated traffic flows if the server has ensured that the specified destination address accepts receiving media through different security mechanisms. Security mechanisms that are acceptable in order of increasing generality are:

- o Verification of the client's identity against a database of known users using RTSP authentication mechanisms (preferably digest authentication or stronger)
- o A list of addresses that have consented to be media destinations, especially considering user identity
- o Media path based verification

The server SHOULD NOT allow the destination field to be set unless a mechanism exists in the system to authorize the request originator to direct streams to the recipient. It is preferred that this authorization be performed by the media recipient (destination) itself and the credentials passed along to the server. However, in certain cases, such as when the recipient address is a multicast group, or when the recipient is unable to communicate with the server in an out-of-band manner, this may not be possible. In these cases the server may choose another method such as a server-resident authorization list to ensure that the request originator has the proper credentials to request stream delivery to the recipient.

One solution that performs the necessary verification of acceptance of media suitable for unicast based delivery is the Interactive Connectivity Establishment (ICE) [RFC5245] based NAT traversal method described in [I-D.ietf-mmusic-rtsp-nat]. This mechanism uses random passwords and a username so that the probability of unintended indication as a valid media destination is very low. In addition the server includes in its Session Traversal Utilities for NAT (STUN) [RFC5389] requests a cookie (consisting of random material) that the destination echoes back, thus the solution also safe-guards against having an off-path attacker being able to spoof the STUN checks. This leaves this solution vulnerable only to on-path attackers that can see the STUN requests go to the target of attack and thus forge a response.

For delivery to multicast addresses there is a need for another solution which is not specified in this memo.

21.2.2. RTP Security analysis

RTP is a commonly used media transport protocol and has been the most common choice for RTSP 1.0 implementations. The core RTP protocol has been in use for a long time and it has well-known security properties and the RTP security consideration (Section 9 of [RFC3550]) needs to be reviewed. In perspective of the usage of RTP in context of RTSP the following properties should be noted:

Stream Additions: RTP has support for multiple simultaneous media streams in each RTP session. As some use cases require support for non-synchronized adding and removal of media streams and their identifiers an attacker can easily insert additional media streams into a session context that according to protocol design is intended to be played out. Another threat vector is one of denial of service by exhausting the resources of the RTP session receiver, for example by using a large number of SSRC identifiers simultaneously. The strong mitigation of this is to ensure that one cryptographically authenticates any incoming packet flow to

the RTP session. Weak mitigations like blocking additional media streams in session contexts easily lead to a denial of service vulnerability in addition to preventing certain RTP extensions or use cases which rely on multiple media streams, such as RTP retransmission [RFC4588] to function.

Forged Feedback: The built in RTP control Protocol (RTCP) also offers a large attack surface for a couple of different types of attacks. One venue is to send RTCP feedback to the media sender indicating large amounts of packet loss and thus trigger a media bit-rate adaptation response from the sender resulting in lowered media quality and potentially shut down of the media stream. Another attack is to perform a resource exhaustion attack on the receiver by using many SSRC identifiers to create large state tables and increase the RTCP related processing demands.

RTP/RTCP Extensions: RTP and RTCP extensions generally provide additional and sometimes extremely powerful tools to do denial of service or service disruption. For example the Code Control Message [RFC5104] RTCP extensions enables both locking down the bit-rate to low values and disruption of video quality by requesting Intra frames.

Taking into account the above general discussion in Section 21.2 and the RTP specific discussion in this section it is clear that it is necessary that a strong security mechanism is supported to protect RTP. Therefore this specification has the following requirements on RTP security functions for all RTSP agents that handles media streams and where the media stream transport is done using RTP.

RTSP agents supporting RTP MUST implement Secure RTP (SRTP) [RFC3711] and thus the SAVP profile. In addition the secure AVP profile (SAVPF) [RFC5124] MUST also be supported if the AVPF profile is implemented. This specification requires no additional cryptographic transforms or configuration values beyond those specified as mandatory to implement in RFC3711, i.e., AES-CM and HMAC-SHA1. The default key-management mechanism which MUST be implemented is the one defined in the MIKEY Key Establishment (Appendix C.1.4.1). The MIKEY implementation MUST implement the necessary functions for MIKEY-RSA-R mode [RFC4738] and in addition the SRTP parameter negotiation necessary to negotiate the supported SRTP transforms and parameters.

22. IANA Considerations

This section sets up a number of registries for RTSP 2.0 that should be maintained by IANA. These registries are separate from any registries existing for RTSP 1.0. For each registry there is a description of what it is required to contain, what specification is

needed when adding an entry with IANA, and finally the entries that this document needs to register. See also the Section 2.7 "Extending RTSP". There is also an IANA registration of three SDP attributes.

Registries or entries in registries which have been made for RTSP 1.0 are not moved to RTSP 2.0. The registries and entries in registries of RTSP 1.0 and RTSP 2.0 are independent. If any registry or entry in a registry is also required in RTSP 2.0, it MUST follow the procedure defined below to allocate the registry or entry in a registry.

The sections describing how to register an item uses some of the registration policies described in RFC 5226 [RFC5226], namely "First Come, First Served", "Expert Review", "Specification Required", and "Standards Action".

RFC-Editor Note: Please replace all occurrences of RFCXXXX with the RFC number this specification receives when published.

In case a registry requires a contact person, the authors, with Magnus Westerlund (magnus.westerlund@ericsson.com) as primary, are the contact persons for any entries created by this document.

IANA will request the following information for any registration request:

- o A name of the item to register according to the rules specified by the intended registry.
- o Indication of who has change control over the feature (for example, IETF, ISO, ITU-T, other international standardization bodies, a consortium, a particular company or group of companies, or an individual);
- o A reference to a further description, if available, for example (in decreasing order of preference) an RFC, a published standard, a published paper, a patent filing, a technical report, documented source code or a computer manual;
- o For proprietary features, contact information (postal and email address);

22.1. Feature-tags

22.1.1.1. Description

When a client and server try to determine what part and functionality of the RTSP specification and any future extensions that its counterpart implements there is need for a namespace. This registry contains named entries representing certain functionality.

The usage of feature-tags is explained in Section 11 and Section 13.1.

22.1.1.2. Registering New Feature-tags with IANA

The registering of feature-tags is done on a First Come, First Served [RFC5226] basis.

The registry entry for a feature-tag has the following information:

- o The name of the feature-tag
 - * If the registrant indicates that the feature is proprietary, IANA should request a vendor "prefix" portion of the name. The name will then be the vendor prefix followed by a "." followed by the rest of the provided feature name.
 - * If the feature is not proprietary, then IANA need not collect a prefix for the name.
- o A one paragraph description of what the feature-tag represents
- o The applicability (server, client, proxy, or some combination)
- o A reference to a specification, if applicable

Feature-tag names (including the vendor prefix) may contain any non-space and non-control characters. There is no length limit on feature-tags.

Examples for a vendor tag describing a proprietary feature are:

vendorA.specfeat01

vendorA.specfeat02

22.1.1.3. Registered entries

The following feature-tags are defined in this specification and hereby registered. The change control belongs to the IETF.

play.basic: The implementation for delivery and playback operations according to the core RTSP specification, as defined in this memo. Applies for both clients, servers and proxies. See Section 11.1.

play.scale: Support of scale operations for media playback. Applies only for servers. See Section 18.46.

play.speed: Support of the speed functionality for media delivery. Applies only for servers. See Section 18.50.

setup.rtp.rtcp.mux Support of the RTP and RTCP multiplexing as discussed in Appendix C.1.6.4. Applies for both client and servers and any media caching proxy.

This should be represented by IANA as a table with the feature tags, contact person and their references.

22.2. RTSP Methods

22.2.1. Description

Methods are described in Section 13. Extending the protocol with new methods allow for totally new functionality.

22.2.2. Registering New Methods with IANA

A new method is registered through an IETF Standards Action [RFC5226]. The reason is that new methods may radically change the protocol's behavior and purpose.

A specification for a new RTSP method consist of the following items:

- o A method name which follows the ABNF rules for methods.
- o A clear specification what a request using the method does and what responses are expected. Which directions the method is used, C->S or S->C or both. How the use of headers, if any, modifies the behavior and effect of the method.
- o A list or table specifying which of the IANA registered headers that are allowed to be used with the method in request or/and response. The list or table SHOULD follow the format of tables in Section 18.
- o Describe how the method relates to network proxies.

22.2.3. Registered Entries

This specification, RFCXXXX, registers 10 methods: DESCRIBE, GET_PARAMETER, OPTIONS, PAUSE, PLAY, PLAY_NOTIFY, REDIRECT, SETUP, SET_PARAMETER, and TEARDOWN. The initial table of the registry is provided below.

Method	Directionality	Reference
DESCRIBE	C->S	[RFCXXXX]
GET_PARAMETER	C->S, S->C	[RFCXXXX]
OPTIONS	C->S, S->C	[RFCXXXX]
PAUSE	C->S	[RFCXXXX]
PLAY	C->S	[RFCXXXX]
PLAY_NOTIFY	S->C	[RFCXXXX]
REDIRECT	S->C	[RFCXXXX]
SETUP	C->S	[RFCXXXX]
SET_PARAMETER	C->S, S->C	[RFCXXXX]
TEARDOWN	C->S, S->C	[RFCXXXX]

22.3. RTSP Status Codes

22.3.1. Description

A status code is the three digit number used to convey information in RTSP response messages, see Section 8. The number space is limited and care should be taken not to fill the space.

22.3.2. Registering New Status Codes with IANA

A new status code registration follows the policy of IETF Review [RFC5226]. New RTSP functionality requiring Status Codes should first be registered in the range x50-x99. Only when the range is full should registrations be done in the x00-x49 range, unless it is to adopt an HTTP extension also to RTSP. The reason is to enable any HTTP extension to be adopted to RTSP without needing to renumber any related status codes. A specification for a new status code specify the following:

- o The registered number.
- o A description of what the status code means and the expected behavior of the sender and receiver of the code.

22.3.3. Registered Entries

RFCXXXX, registers the numbered status code defined in the ABNF entry "Status-Code" except "extension-code" (that defines the syntax allowed for future extensions) in Section 20.2.2.

22.4. RTSP Headers

22.4.1. Description

By specifying new headers a method(s) can be enhanced in many different ways. An unknown header will be ignored by the receiving agent. If the new header is vital for a certain functionality, a feature-tag for the functionality can be created and demanded to be used by the counter-part with the inclusion of a Require header carrying the feature-tag.

22.4.2. Registering New Headers with IANA

Registrations in the registry can be done following the Expert Review policy [RFC5226]. A specification is recommended to be provided, preferably an IETF RFC or other Standards Developing Organization specification. The minimal information in a registration request is the header name and the contact information.

The expert reviewer verifies that the registration request contain the following information:

- o The name of the header.
- o An ABNF specification of the header syntax.
- o A list or table specifying when the header may be used, encompassing all methods, their request or response, the direction (C->S or S->C).
- o How the header is to be handled by proxies.
- o A description of the purpose of the header.

22.4.3. Registered entries

All headers specified in Section 18 in RFCXXXX are to be registered. The Registry is to include header name and reference.

Furthermore the following legacy RTSP headers defined in other specifications are registered with header name, reference and description according to below list. Note: These references may not

fulfill all of the above rules for registrations due to their legacy status.

- o x-wap-profile defined in [TS-26234]. The x-wap-profile request-header contains one or more absolute URLs to the requesting agent's device capability profile.
- o x-wap-profile-diff defined in [TS-26234]. The x-wap-profile-diff request-header contains a subset of a device capability profile.
- o x-wap-profile-warning defined in [TS-26234]. The x-wap-profile-warning is a response-header that contains error codes explaining to what extent the server has been able to match the terminal request in regards to device capability profile as described using x-wap-profile and x-wap-profile-diff headers.
- o x-predecbufsize defined in [TS-26234]. This response-header provides an RTSP agent with the TS 26.234 Annex G hypothetical pre-decoder buffer size.
- o x-initpredecbufperiod defined in [TS-26234]. This response-header provides an RTSP agent with the TS 26.234 Annex G hypothetical pre-decoder buffering period.
- o x-initpostdecbufperiod defined in [TS-26234]. This response-header provides an RTSP agent with the TS 26.234 Annex G post-decoder buffering period.
- o 3gpp-videopostdecbufsize defined in [TS-26234]. This response-header provides an RTSP agent with the TS 26.234 defined post-decoder buffer size usable for H.264 (AVC) video streams.
- o 3GPP-Link-Char defined in [TS-26234]. This request-header provides the RTSP server with the RTSP client's link characteristics as determined from the radio interface. The information that can be provided are guaranteed bit-rate, maximum bit-rate and maximum transfer delay.
- o 3GPP-Adaptation defined in [TS-26234]. This general-header is part of the bit-rate adaptation solution specified for PSS. It provides the RTSP client's buffer sizes and target buffer levels to the server and responses are used to acknowledge the support and values.
- o 3GPP-QoE-Metrics defined in [TS-26234]. This general-header is used by PSS RTSP agents to negotiate the quality of experience metrics that a client should gather and report to the server.

- o 3GPP-QoE-Feedback defined in [TS-26234]. This request-header is used by RTSP clients supporting PSS to report the actual values of the metrics gathered in its quality of experience metering.

The use of "x-" is NOT RECOMMENDED but the above headers in the register list were defined prior to the clarification.

22.5. Accept-Credentials

The security framework's TLS connection mechanism has two registerable entities.

22.5.1. Accept-Credentials policies

This registry are for polices for a RTSP proxy's handling and verification of TLS certificates when establishing outbound TLS connection on clients behalf. In Section 19.3.1 three policies for how to handle certificates are specified. Further policies may be defined and registration is done through an IETF Standards Action [RFC5226]. The registration is required to contain the following information:

- o Name of the policy.
- o A describing text that explains how the policy works for handling the certificates.
- o A contact person.

This specification registers the following values:

Any: A policy requiring the proxy to accept any received certificate.

Proxy: A policy where the proxy applies its own policies to determine which certificates are accepted or not.

User: A policy where the certificate is required to be forwarded down the proxy chain to the client, thus allowing the user to decided to accept or refuse a certificate.

22.5.2. Accept-Credentials hash algorithms

The Accept-Credentials header (See Section 18.2) allows for the usage of other algorithms for hashing the DER records of accepted entities. The registration of any future algorithm is expected to be extremely rare and could also cause interoperability problems. Therefore the bar for registering new algorithms is intentionally placed high.

Any registration of a new hash algorithm requires an IETF Standards Action [RFC5226]. The registration needs to fulfill the following requirement:

- o The algorithms identifier meeting the "token" ABNF requirement.
- o Provide a definition of the algorithm.

The registered value is:

Hash Alg. Id	Reference

sha-256	[RFCXXXX]

22.6. Cache-Control Cache Directive Extensions

There exists a number of cache directives which can be sent in the Cache-Control header. A registry for these cache directives is established by IANA. New registrations in this registry requires an IETF Standards Action or IESG Approval [RFC5226]. The registration needs to contain the following information.

- o Name of the directive
- o A definition of the parameter value, if any is allowed.
- o Specification if it is a request or response directive.
- o A describing text that explains how the cache directive is used for RTSP controlled media streams.
- o A contact person.

This specification registers the following values:

no-cache:

public:

private:

no-transform:

only-if-cached:

max-stale:

min-fresh:

must-revalidate:

proxy-revalidate:

max-age:

The registry should be represented as: Name of the directive, contact person and reference.

22.7. Media Properties

22.7.1. Description

The media streams being controlled by RTSP can have many different properties. The media properties required to cover the use cases that were in mind when writing the specification are defined. However, it can be expected that further innovation will result in new use cases or media streams with properties not covered by the ones specified here. Thus new media properties can be specified. As new media properties may need a substantial amount of new definitions to correctly specify behavior for this property the bar is intended to be high.

22.7.2. Registration Rules

Registering a new media property is done following the Specification Required policy [RFC5226]. The Expert reviewer verifies that a registration request fulfill the following requirements.

- o Have an ABNF definition of the media property value name that meets "media-prop-ext" definition.
- o Define which media property group it belongs to or define a new group.
- o Description of all changes to the behavior of the RTSP protocol as result of these changes.
- o A Contact Person for the Registration.

22.7.3. Registered Values

This specification registers the ten values listed in Section 18.29. The registry should be represented as: Name of the media property, property group, contact person and reference.

22.8. Notify-Reason header

22.8.1. Description

Notify-Reason values are used for indicating the reason the notification was sent. Each reason has its associated rules on what headers and information that may or must be included in the notification. New notification behaviors need to be specified to enable interoperable usage, thus a specification of each new value is required.

22.8.2. Registration Rules

Registrations for new Notify-Reason value follows the Specification Required policy [RFC5226]. The Expert Reviewer verifies that the request fulfills the following requirements:

- o An ABNF definition of the Notify reason value name that meets "Notify-Reason-extension" definition
- o Description of which headers shall be included in the request and response, when it should be sent, and any effect it has on the server client state.
- o A Contact Person for the Registration

22.8.3. Registered Values

This specification registers 3 values defined in the Notify-Reas-val ABNF, Section 20.2.3:

end-of-stream: This Notify-Reason value indicates the end of a media stream.

media-properties-update: This Notify-Reason value allows the server to indicate that the properties of the media has changed during the playout.

scale-change: This Notify-Reason value allows the server to notify the client about a change in the Scale of the media.

The registry entries should be represented in the registry as: Name, short description, contact and reference.

22.9. Range Header Formats

22.9.1. Description

The Range header (Section 18.40) allows for different range formats. These range formats also needs an identifier to be used the Accept-Ranges header (Section 18.5). New range formats may be registered, but moderation should be applied as it makes interoperability more difficult.

22.9.2. Registration Rules

A registration follows the Specification Required policy [RFC5226]. The Expert Reviewer verifies that the request fulfills the following requirements:

- o An ABNF definition of the range format that fulfills the "range-ext" definition.
- o Define the range format identifier used in Accept-Ranges header according to the "extension-format" definition.
- o Rules for how one handles the range when using a negative Scale.
- o A Contact person for the registration.

22.9.3. Registered Values

The registry should be represented as: Range header format identifier, Name of the range format, contact person and reference. This specification registers the following values.

npt: Normal Play Time

clock: UTC Absolute Time format

smpte: SMPTE Timestamps

smpte-30-drop: SMPTE Timestamps 29.97 Frames/sec (30 Hz with Drop)

smpte-25: SMPTE Timestamps 25 Frames/sec

22.10. Terminate-Reason Header

The Terminate-Reason header (Section 18.52) has two registries for extensions.

22.10.1. Redirect Reasons

This registry contains reasons for session Termination that can be included in a Terminate-Reason header (Section 18.52). Registrations are following the policy of Expert Review [RFC5226]. The Expert Reviewer verifies that the registration contains the following information:

- o The value follows the Terminate-Reason ABNF, i.e., be a token.
- o That the specification provide a definition of what procedures are to be followed when a client receives this redirect reason.
- o A Contact Person

This specification registers three values:

- o Session-Timeout
- o Server-Admin
- o Internal-Error

The registry should be represented as: Name of the Redirect Reason, contact person and reference.

22.10.2. Terminate-Reason Header Parameters

This registry contains parameters that may be included in the Terminate-Reason header (Section 18.52) in addition to a reason. Registrations are done under the policy of Specification Required [RFC5226]. The Expert Reviewer verifies that the registration or the reference specification contains the following:

- o A Parameter Name.
- o A Parameter following the syntax allowed by the RTSP 2.0 specification.
- o A Reference to the specification.
- o A contact person.

This specification registers:

- o time
- o user-msg

The registry should be represented as: Name of the Terminate Reason, contact person and reference.

22.11. RTP-Info header parameters

22.11.1. Description

The RTP-Info header (Section 18.45) carries one or more parameter value pairs with information about a particular point in the RTP stream. RTP extensions or new usages may need new types of information. As RTP information that could be needed is likely to be generic enough and to maximize the interoperability, new registration requires Specification Required.

22.11.2. Registration Rules

Registrations for new RTP-Info values follows the policy of Specification Required [RFC5226]. The Expert Reviewer verifies that the registration and its reference contains the following information.

- o Have an ABNF definition that meets the "generic-param" definition.
- o A Reference to the specification.
- o A Contact Person for the Registration.

22.11.3. Registered Values

This specification registers the following parameter value pairs:

- o url
- o ssrc
- o seq
- o rtptime

The registry should be represented as: Name of the parameter, contact person and reference.

22.12. Seek-Style Policies

22.12.1. Description

Seek-Style Policies defines how the RTSP agent seeks in media content when given a position within the media content. New seek policies may be registered, however, a large number of these will complicate implementation substantially. The impact of unknown policies is that the server will not honor the unknown and use the server default policy instead.

22.12.2. Registration Rules

Registrations of new Seek-Style policies follows the policy of Specification Required [RFC5226]. The Expert Reviewer verifies that the registration fulfill the following requirements:

- o Have an ABNF definition of the Seek-Style policy name that meets "Seek-S-value-ext" definition
- o Short Description
- o A Contact Person for the Registration
- o Description of which headers shall be included in the request and response, when it should be sent, and any affect it has on the server client state.

22.12.3. Registered Values

This specification registers 4 values (Name - Short Description):

- o RAP - Using the closest Random Access Point prior or at the requested start position.
- o CoRAP - Conditional Random Access Point is like RAP, but only if the RAP is closer prior to the requested start position than current pause point.
- o First-Prior - The first-prior policy will start delivery with the media unit that has a playout time first prior to the requested start position.
- o Next - The next media units after the provided start position.

The registry should be represented as: Name of the Seek-Style Policy, short description, contact person and reference.

22.13. Transport Header Registries

The transport header (Section 18.54) contains a number of parameters which have possibilities for future extensions. Therefore registries for these need to be defined.

22.13.1. Transport Protocol Identifier

A Transport Protocol Specification consists of a Transport Protocol Identifier, representing some combination of transport protocols, and any number of transport header parameters required or optional to use with the identified protocol specification. This registry contains the identifiers used by registered Transport Protocol Identifiers.

A registration for the parameter transport protocol identifier follows the policy of Specification Required [RFC5226]. The expert reviewer verifies that the registration fulfill the following requirements:

- o A contact person or organization with address and email.
- o A value definition that are following the ABNF syntax definition of "transport-id" Section 20.2.3.
- o A descriptive text that explains how the registered value are used in RTSP, which underlying transport protocols that are used, and any required Transport header parameters.

The registry should be represented as: The protocol ID string, contact person and reference.

This specification registers the following values:

RTP/AVP: Use of the RTP [RFC3550] protocol for media transport in combination with the "RTP profile for audio and video conferences with minimal control" [RFC3551] over UDP. The usage is explained in RFC XXXX, Appendix C.1.

RTP/AVP/UDP: the same as RTP/AVP.

RTP/AVPF: Use of the RTP [RFC3550] protocol for media transport in combination with the "Extended RTP Profile for RTCP-based Feedback (RTP/AVPF)" [RFC4585] over UDP. The usage is explained in RFC XXXX, Appendix C.1.

RTP/AVPF/UDP: the same as RTP/AVPF.

RTP/SAVP: Use of the RTP [RFC3550] protocol for media transport in combination with the "The Secure Real-time Transport Protocol (SRTP)" [RFC3711] over UDP. The usage is explained in RFC XXXX, Appendix C.1.

RTP/SAVP/UDP: the same as RTP/SAVP.

RTP/SAVPF: Use of the RTP[RFC3550] protocol for media transport in combination with the Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF) [RFC5124] over UDP. The usage is explained in RFC XXXX, Appendix C.1.

RTP/SAVPF/UDP: the same as RTP/SAVPF.

RTP/AVP/TCP: Use of the RTP [RFC3550] protocol for media transport in combination with the "RTP profile for audio and video conferences with minimal control" [RFC3551] over TCP. The usage is explained in RFC XXXX, Appendix C.2.2.

RTP/AVPF/TCP: Use of the RTP [RFC3550] protocol for media transport in combination with the "Extended RTP Profile for RTCP-based Feedback (RTP/AVPF)" [RFC4585] over TCP. The usage is explained in RFC XXXX, Appendix C.2.2.

RTP/SAVP/TCP: Use of the RTP [RFC3550] protocol for media transport in combination with the "The Secure Real-time Transport Protocol (SRTP)" [RFC3711] over TCP. The usage is explained in RFC XXXX, Appendix C.2.2.

RTP/SAVPF/TCP: Use of the RTP [RFC3550] protocol for media transport in combination with the "Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)" [RFC5124] over TCP. The usage is explained in RFC XXXX, Appendix C.2.2.

22.13.2. Transport modes

The Transport Mode is a Transport header (Section 18.54) parameter, it is used to identify the general mode of media transport. The PLAY value registered defines a PLAYBACK mode, where media flows from Server to Client.

A registration for the transport parameter mode follows the policy of IETF Standards Action [RFC5226]. The registration needs to meet the following requirements:

- o A value definition that are following the ABNF "token" definition Section 20.2.3.
- o A describing text that explains how the registered value are used in RTSP.

This specification registers 1 value:

PLAY: See RFC XXXX.

The registry should be represented as: The Transport Mode value, contact person and reference.

22.13.3. Transport Parameters

Transport Parameters are different parameters used in a Transport Header's transport specification (Section 18.54) to provide additional information required beyond the transport protocol identifier to establish a functioning transport.

A registration for parameters that may be included in the Transport header follows the policy of Specification Required [RFC5226]. The expert reviewer verifies that the registration fulfill the following requirements:

- o A Transport Parameter Name following the "token" ABNF definition.
- o A value definition, if the parameter takes a value, that are following the ABNF definition "trn-par-value" Section 20.2.3.
- o A describing text that explains how the registered value are used in RTSP.

This specification registers all the transport parameters defined in Section 18.54. This is a copy of this list:

- o unicast
- o multicast
- o interleaved
- o ttl
- o layers
- o ssrc

- o mode
- o dest_addr
- o src_addr
- o setup
- o connection
- o RTCP-mux
- o MIKEY

The registry should be represented as: The transport parameter name, contact person and reference.

22.14. URI Schemes

This specification updates two URI schemes, one previously registered "rtsp", and one missing in the registry "rtspu", previously only defined in the RTSP 1.0 [RFC2326], one new URI scheme "rtsps" is also registered. These URI schemes are registered in an existing registry (Uniform Resource Identifier (URI) Schemes) which is not created by this memo. Registrations are following RFC 4395[RFC4395].

22.14.1. The rtsp URI Scheme

URI scheme name: rtsp

Status: Permanent

URI scheme syntax: See Section 20.2.1 of RFC XXXX.

URI scheme semantics: The rtsp scheme is used to indicate resources accessible through the usage of the Real-time Streaming Protocol (RTSP). RTSP allows different operations on the resource identified by the URI, but the primary purpose is the streaming delivery of the resource to a client. However, the operations that are currently defined are: DESCRIBE, GET_PARAMETER, OPTIONS, PLAY, PLAY_NOTIFY, PAUSE, REDIRECT, SETUP, SET_PARAMETER, and TEARDOWN.

Encoding considerations: IRIs in this scheme are defined and needs to be encoded as RTSP URIs when used within the RTSP protocol. That encoding is done according to RFC 3987.

Applications/protocols that use this URI scheme name: RTSP 1.0 (RFC 2326), RTSP 2.0 (RFC XXXX)

Interoperability considerations: The extensions in the URI syntax performed between RTSP 1.0 and 2.0 can create interoperability issues. The changes are:

Support for IPV6 literal in host part and future IP literals through RFC 3986 defined mechanism.

A new relative format to use in the RTSP protocol elements that is not required to start with "/".

The above changes should have no impact on interoperability as in detail discussed in Section 4.2 of RFCXXXX.

Security considerations: All the security threats identified in Section 7 of RFC 3986 apply also to this scheme. They need to be reviewed and considered in any implementation utilizing this scheme.

Contact: Magnus Westerlund, magnus.westerlund@ericsson.com

Author/Change controller: IETF

References: RFC 2326, RFC 3986, RFC 3987, RFC XXXX

22.14.2. The rtsp URI Scheme

URI scheme name: rtsp

Status: Permanent

URI scheme syntax: See Section 20.2.1 of RFC XXXX.

URI scheme semantics: The rtsp scheme is used to indicate resources accessible through the usage of the Real-time Streaming Protocol (RTSP) over TLS. RTSP allows different operations on the resource identified by the URI, but the primary purpose is the streaming delivery of the resource to a client. However, the operations that are currently defined are: DESCRIBE, GET_PARAMETER, OPTIONS, PLAY, PLAY_NOTIFY, PAUSE, REDIRECT, SETUP, SET_PARAMETER, and TEARDOWN.

Encoding considerations: IRIs in this scheme are defined and needs to be encoded as RTSP URIs when used within the RTSP protocol. That encoding is done according to RFC 3987.

Applications/protocols that use this URI scheme name: RTSP 1.0 (RFC 2326), RTSP 2.0 (RFC XXXX)

Interoperability considerations: The "rtsps" scheme was never officially defined for RTSP 1.0, however it has seen widespread use in actual deployments of RTSP 1.0. Therefore this section discusses the believed changes between the unspecified RTSP 1.0 "rtsps" scheme and RTSP 2.0 definition. The extensions in the URI syntax performed between RTSP 1.0 and 2.0 can create interoperability issues. The changes are:

Support for IPV6 literal in host part and future IP literals through RFC 3986 defined mechanism.

A new relative format to use in the RTSP protocol elements that is not required to start with "/".

The above changes should have no impact on interoperability as in detail discussed in Section 4.2 of RFCXXXX.

Security considerations: All the security threats identified in Section 7 of RFC 3986 apply also to this scheme. They need to be reviewed and considered in any implementation utilizing this scheme.

Contact: Magnus Westerlund, magnus.westerlund@ericsson.com

Author/Change controller: IETF

References: RFC 2326, RFC 3986, RFC 3987, RFC XXXX

22.14.3. The rtspu URI Scheme

URI scheme name: rtspu

Status: Permanent

URI scheme syntax: See Section 3.2 of RFC 2326.

URI scheme semantics: The rtspu scheme is used to indicate resources accessible through the usage of the Real-time Streaming Protocol (RTSP) over unreliable datagram transport. RTSP allows different operations on the resource identified by the URI, but the primary purpose is the streaming delivery of the resource to a client. However, the operations that are currently defined are: DESCRIBE, GET_PARAMETER, OPTIONS, REDIRECT, PLAY, PLAY_NOTIFY, PAUSE, SETUP, SET_PARAMETER, and TEARDOWN.

Encoding considerations: This scheme is not intended to be used with characters outside the US-ASCII repertoire.

Applications/protocols that use this URI scheme name: RTSP 1.0 (RFC 2326)

Interoperability considerations: The definition of the transport mechanism of RTSP over UDP has interoperability issues. That makes the usage of this scheme problematic.

Security considerations: All the security threats identified in Section 7 of RFC 3986 apply also to this scheme. They needs to be reviewed and considered in any implementation utilizing this scheme.

Contact: Magnus Westerlund, magnus.westerlund@ericsson.com

Author/Change controller: IETF

References: RFC 2326

22.15. SDP attributes

This specification defines three SDP [RFC4566] attributes that it is requested that IANA register.

SDP Attribute ("att-field"):

Attribute name: range
Long form: Media Range Attribute
Type of name: att-field
Type of attribute: Media and session level
Subject to charset: No
Purpose: RFC XXXX
Reference: RFC XXXX, RFC 2326
Values: See ABNF definition.

Attribute name: control
Long form: RTSP control URI
Type of name: att-field
Type of attribute: Media and session level
Subject to charset: No
Purpose: RFC XXXX
Reference: RFC XXXX, RFC 2326
Values: Absolute or Relative URIs.

Attribute name: mtag
Long form: Message Tag
Type of name: att-field
Type of attribute: Media and session level
Subject to charset: No
Purpose: RFC XXXX
Reference: RFC XXXX
Values: See ABNF definition

22.16. Media Type Registration for text/parameters

Type name: text

Subtype name: parameters

Required parameters:

Optional parameters: charset: The charset parameter is applicable to the encoding of the parameter values. The default charset is UTF-8, if the 'charset' parameter is not present.

Encoding considerations: 8bit

Security considerations: This format may carry any type of parameters. Some can have security requirements, like privacy, confidentiality or integrity requirements. The format has no built in security protection. For the usage it was defined the

transport can be protected between server and client using TLS. However, care must be taken to consider if also the proxies are trusted with the parameters in case hop-by-hop security is used. If stored as a file in file system, the necessary precautions need to be taken in relation to the parameters requirements including object security such as S/MIME [RFC5751].

Interoperability considerations: This media type was mentioned as a fictional example in [RFC2326], but was not formally specified. This has resulted in usage of this media type which may not match its formal definition.

Published specification: RFC XXXX, Appendix F.

Applications that use this media type: Applications that use RTSP and have additional parameters they like to read and set using the RTSP GET_PARAMETER and SET_PARAMETER methods.

Additional information:

Magic number(s):

File extension(s):

Macintosh file type code(s):

Person & email address to contact for further information: Magnus Westerlund (magnus.westerlund@ericsson.com)

Intended usage: Common

Restrictions on usage: None

Author: Magnus Westerlund (magnus.westerlund@ericsson.com)

Change controller: IETF

Addition Notes:

23. References

23.1. Normative References

[FIPS-pub-180-2]

National Institute of Standards and Technology (NIST),
"Federal Information Processing Standards Publications
(FIPS PUBS) 180-2: Secure Hash Standard", August 2002.

- [I-D.ietf-avtcore-rtp-circuit-breakers]
Perkins, C. and V. Singh, "Multimedia Congestion Control: Circuit Breakers for Unicast RTP Sessions", draft-ietf-avtcore-rtp-circuit-breakers-04 (work in progress), January 2014.
- [I-D.ietf-mmusic-rtsp-nat]
Goldberg, J., Westerlund, M., and T. Zeng, "A Network Address Translator (NAT) Traversal Mechanism for Media Controlled by Real-Time Streaming Protocol (RTSP)", draft-ietf-mmusic-rtsp-nat-20 (work in progress), February 2014.
- [RFC0768] Postel, J., "User Datagram Protocol", STD 6, RFC 768, August 1980.
- [RFC0793] Postel, J., "Transmission Control Protocol", STD 7, RFC 793, September 1981.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, December 1998.
- [RFC2616] Fielding, R., Gettys, J., Mogul, J., Frystyk, H., Masinter, L., Leach, P., and T. Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1", RFC 2616, June 1999.
- [RFC2617] Franks, J., Hallam-Baker, P., Hostetler, J., Lawrence, S., Leach, P., Luotonen, A., and L. Stewart, "HTTP Authentication: Basic and Digest Access Authentication", RFC 2617, June 1999.
- [RFC2818] Rescorla, E., "HTTP Over TLS", RFC 2818, May 2000.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.
- [RFC3551] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, July 2003.
- [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, RFC 3629, November 2003.

- [RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.
- [RFC3830] Arkko, J., Carrara, E., Lindholm, F., Naslund, M., and K. Norrman, "MIKEY: Multimedia Internet KEYing", RFC 3830, August 2004.
- [RFC3986] Berners-Lee, T., Fielding, R., and L. Masinter, "Uniform Resource Identifier (URI): Generic Syntax", STD 66, RFC 3986, January 2005.
- [RFC3987] Duerst, M. and M. Suignard, "Internationalized Resource Identifiers (IRIs)", RFC 3987, January 2005.
- [RFC4086] Eastlake, D., Schiller, J., and S. Crocker, "Randomness Requirements for Security", BCP 106, RFC 4086, June 2005.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.
- [RFC4395] Hansen, T., Hardie, T., and L. Masinter, "Guidelines and Registration Procedures for New URI Schemes", BCP 35, RFC 4395, February 2006.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC4571] Lazzaro, J., "Framing Real-time Transport Protocol (RTP) and RTP Control Protocol (RTCP) Packets over Connection-Oriented Transport", RFC 4571, July 2006.
- [RFC4585] Ott, J., Wenger, S., Sato, N., Burmeister, C., and J. Rey, "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)", RFC 4585, July 2006.
- [RFC4648] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", RFC 4648, October 2006.
- [RFC4738] Ignjatic, D., Dondeti, L., Audet, F., and P. Lin, "MIKEY-RSA-R: An Additional Mode of Key Distribution in Multimedia Internet KEYing (MIKEY)", RFC 4738, November 2006.
- [RFC5124] Ott, J. and E. Carrara, "Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)", RFC 5124, February 2008.

- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.
- [RFC5246] Dierks, T. and E. Rescorla, "The Transport Layer Security (TLS) Protocol Version 1.2", RFC 5246, August 2008.
- [RFC5280] Cooper, D., Santesson, S., Farrell, S., Boeyen, S., Housley, R., and W. Polk, "Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile", RFC 5280, May 2008.
- [RFC5322] Resnick, P., Ed., "Internet Message Format", RFC 5322, October 2008.
- [RFC5646] Phillips, A. and M. Davis, "Tags for Identifying Languages", BCP 47, RFC 5646, September 2009.
- [RFC5751] Ramsdell, B. and S. Turner, "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Message Specification", RFC 5751, January 2010.
- [RFC5761] Perkins, C. and M. Westerlund, "Multiplexing RTP Data and Control Packets on a Single Port", RFC 5761, April 2010.
- [RFC5888] Camarillo, G. and H. Schulzrinne, "The Session Description Protocol (SDP) Grouping Framework", RFC 5888, June 2010.
- [RFC6838] Freed, N., Klensin, J., and T. Hansen, "Media Type Specifications and Registration Procedures", BCP 13, RFC 6838, January 2013.
- [SMPTE_TC] Society of Motion Picture and Television Engineers, "SMPTE Standard for Television -- Time and Control Code, ST 12M-1-2008", .
- [TS-26234] Third Generation Partnership Project (3GPP), "Transparent end-to-end Packet-switched Streaming Service (PSS); Protocols and codecs; Technical Specification 26.234", December 2002.

23.2. Informative References

- [ISO.13818-6.1995]
International Organization for Standardization,
"Information technology - Generic coding of moving
pictures and associated audio information - part 6:
Extension for digital storage media and control", ISO
Draft Standard 13818-6, November 1995.
- [ISO.8601.2000]
International Organization for Standardization, "Data
elements and interchange formats - Information interchange
- Representation of dates and times", ISO/IEC Standard
8601, December 2000.
- [RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, September
1981.
- [RFC1123] Braden, R., "Requirements for Internet Hosts - Application
and Support", STD 3, RFC 1123, October 1989.
- [RFC2068] Fielding, R., Gettys, J., Mogul, J., Nielsen, H., and T.
Berners-Lee, "Hypertext Transfer Protocol -- HTTP/1.1",
RFC 2068, January 1997.
- [RFC2326] Schulzrinne, H., Rao, A., and R. Lanphier, "Real Time
Streaming Protocol (RTSP)", RFC 2326, April 1998.
- [RFC2663] Srisuresh, P. and M. Holdrege, "IP Network Address
Translator (NAT) Terminology and Considerations", RFC
2663, August 1999.
- [RFC2974] Handley, M., Perkins, C., and E. Whelan, "Session
Announcement Protocol", RFC 2974, October 2000.
- [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston,
A., Peterson, J., Sparks, R., Handley, M., and E.
Schooler, "SIP: Session Initiation Protocol", RFC 3261,
June 2002.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model
with Session Description Protocol (SDP)", RFC 3264, June
2002.
- [RFC3339] Klyne, G., Ed. and C. Newman, "Date and Time on the
Internet: Timestamps", RFC 3339, July 2002.

- [RFC4145] Yon, D. and G. Camarillo, "TCP-Based Media Transport in the Session Description Protocol (SDP)", RFC 4145, September 2005.
- [RFC4567] Arkko, J., Lindholm, F., Naslund, M., Norrman, K., and E. Carrara, "Key Management Extensions for Session Description Protocol (SDP) and Real Time Streaming Protocol (RTSP)", RFC 4567, July 2006.
- [RFC4588] Rey, J., Leon, D., Miyazaki, A., Varsa, V., and R. Hakenberg, "RTP Retransmission Payload Format", RFC 4588, July 2006.
- [RFC4855] Casner, S., "Media Type Registration of RTP Payload Formats", RFC 4855, February 2007.
- [RFC4856] Casner, S., "Media Type Registration of Payload Formats in the RTP Profile for Audio and Video Conferences", RFC 4856, February 2007.
- [RFC5104] Wenger, S., Chandra, U., Westerlund, M., and B. Burman, "Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)", RFC 5104, February 2008.
- [RFC5245] Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols", RFC 5245, April 2010.
- [RFC5389] Rosenberg, J., Mahy, R., Matthews, P., and D. Wing, "Session Traversal Utilities for NAT (STUN)", RFC 5389, October 2008.
- [RFC5583] Schierl, T. and S. Wenger, "Signaling Media Decoding Dependency in the Session Description Protocol (SDP)", RFC 5583, July 2009.
- [RFC5905] Mills, D., Martin, J., Burbank, J., and W. Kasch, "Network Time Protocol Version 4: Protocol and Algorithms Specification", RFC 5905, June 2010.
- [RFC6298] Paxson, V., Allman, M., Chu, J., and M. Sargent, "Computing TCP's Retransmission Timer", RFC 6298, June 2011.
- [Stevens98] Stevens, W., "Unix Networking Programming - Volume 1, second edition", 1998.

Appendix A. Examples

This section contains several different examples trying to illustrate possible ways of using RTSP. The examples can also help with the understanding of how functions of RTSP work. However, remember that these are examples and the normative and syntax description in the other sections take precedence. Please also note that many of the examples contain syntax illegal line breaks to accommodate the formatting restriction that the RFC series impose.

A.1. Media on Demand (Unicast)

This is an example of media on demand streaming of a media stored in a container file. For purposes of this example, a container file is a storage entity in which multiple continuous media types pertaining to the same end-user presentation are present. In effect, the container file represents an RTSP presentation, with each of its components being RTSP controlled media streams. Container files are a widely used means to store such presentations. While the components are transported as independent streams, it is desirable to maintain a common context for those streams at the server end.

This enables the server to keep a single storage handle open easily. It also allows treating all the streams equally in case of any prioritization of streams by the server.

It is also possible that the presentation author may wish to prevent selective retrieval of the streams by the client in order to preserve the artistic effect of the combined media presentation. Similarly, in such a tightly bound presentation, it is desirable to be able to control all the streams via a single control message using an aggregate URI.

The following is an example of using a single RTSP session to control multiple streams. It also illustrates the use of aggregate URIs. In a container file it is also desirable to not write any URI parts which are not kept, when the container is distributed, like the host and most of the path element. Therefore this example also uses the "*" and relative URI in the delivered SDP.

Also this presentation description (SDP) is not cacheble, as the Expires header is set to an equal value with date indicating immediate expiration of its validity.

Client C requests a presentation from media server M. The movie is stored in a container file. The client has obtained an RTSP URI to the container file.

C->M: DESCRIBE rtsp://example.com/twister.3gp RTSP/2.0
CSeq: 1
User-Agent: PhonyClient/1.2

M->C: RTSP/2.0 200 OK
CSeq: 1
Server: PhonyServer/1.0
Date: Fri, 20 Dec 2013 10:20:32 +0000
Content-Type: application/sdp
Content-Length: 271
Content-Base: rtsp://example.com/twister.3gp/
Expires: Fri, 20 Dec 2013 12:20:32 +0000

v=0
o=- 2890844256 2890842807 IN IP4 198.51.100.5
s=RTSP Session
i=An Example of RTSP Session Usage
e=adm@example.com
c=IN IP4 0.0.0.0
a=control: *
a=range:npt=00:00:00-00:10:34.10
t=0 0
m=audio 0 RTP/AVP 0
a=control: trackID=1
m=video 0 RTP/AVP 26
a=control: trackID=4

```
C->M: SETUP rtsp://example.com/twister.3gp/trackID=1 RTSP/2.0
      CSeq: 2
      User-Agent: PhonyClient/1.2
      Require: play.basic
      Transport: RTP/AVP;unicast;dest_addr=":8000"/":8001"
      Accept-Ranges: npt, smpte, clock

M->C: RTSP/2.0 200 OK
      CSeq: 2
      Server: PhonyServer/1.0
      Transport: RTP/AVP;unicast; ssrc=93CB001E;
                dest_addr="192.0.2.53:8000"/"192.0.2.53:8001";
                src_addr="198.51.100.5:9000"/"198.51.100.5:9001"
      Session: 12345678
      Expires: Fri, 20 Dec 2013 12:20:33 +0000
      Date: Fri, 20 Dec 2013 10:20:33 +0000
      Accept-Ranges: npt
      Media-Properties: Random-Access=0.02, Immutable, Unlimited

C->M: SETUP rtsp://example.com/twister.3gp/trackID=4 RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Require: play.basic
      Transport: RTP/AVP;unicast;dest_addr=":8002"/":8003"
      Session: 12345678
      Accept-Ranges: npt, smpte, clock

M->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/1.0
      Transport: RTP/AVP;unicast; ssrc=A813FC13;
                dest_addr="192.0.2.53:8002"/"192.0.2.53:8003";
                src_addr="198.51.100.5:9002"/"198.51.100.5:9003";

      Session: 12345678
      Expires: Fri, 20 Dec 2013 12:20:33 +0000
      Date: Fri, 20 Dec 2013 10:20:33 +0000
      Accept-Range: NPT
      Media-Properties: Random-Access=0.8, Immutable, Unlimited

C->M: PLAY rtsp://example.com/twister.3gp/ RTSP/2.0
      CSeq: 4
      User-Agent: PhonyClient/1.2
      Range: npt=30-
      Seek-Style: RAP
      Session: 12345678
```

```
M->C: RTSP/2.0 200 OK
      CSeq: 4
      Server: PhonyServer/1.0
      Date: Fri, 20 Dec 2013 10:20:34 +0000
      Session: 12345678
      Range: npt=30-634.10
      Seek-Style: RAP
      RTP-Info: url="rtsp://example.com/twister.3gp/trackID=4"
                ssrc=0D12F123:seq=12345;rtptime=3450012,
                url="rtsp://example.com/twister.3gp/trackID=1"
                ssrc=4F312DD8:seq=54321;rtptime=2876889

C->M: PAUSE rtsp://example.com/twister.3gp/ RTSP/2.0
      CSeq: 5
      User-Agent: PhonyClient/1.2
      Session: 12345678

# Pause happens 0.87 seconds after starting to play

M->C: RTSP/2.0 200 OK
      CSeq: 5
      Server: PhonyServer/1.0
      Date: Fri, 20 Dec 2013 10:20:35 +0000
      Session: 12345678
      Range: npt=30.87-634.10

C->M: PLAY rtsp://example.com/twister.3gp/ RTSP/2.0
      CSeq: 6
      User-Agent: PhonyClient/1.2
      Range: npt=30.87-634.10
      Seek-Style: Next
      Session: 12345678

M->C: RTSP/2.0 200 OK
      CSeq: 6
      Server: PhonyServer/1.0
      Date: Fri, 20 Dec 2013 10:22:13 +0000
      Session: 12345678
      Range: npt=30.87-634.10
      Seek-Style: Next
      RTP-Info: url="rtsp://example.com/twister.3gp/trackID=4"
                ssrc=0D12F123:seq=12555;rtptime=6330012,
                url="rtsp://example.com/twister.3gp/trackID=1"
                ssrc=4F312DD8:seq=55021;rtptime=3132889

C->M: TEARDOWN rtsp://example.com/twister.3gp/ RTSP/2.0
      CSeq: 7
```

User-Agent: PhonyClient/1.2
Session: 12345678

M->C: RTSP/2.0 200 OK
CSeq: 7
Server: PhonyServer/1.0
Date: Fri, 20 Dec 2013 10:31:53 +0000

A.2. Media on Demand using Pipelining

This example is basically the example above (Appendix A.1), but now utilizing pipelining to speed up the setup. It requires only two round trip times until the media starts flowing. First of all, the session description is retrieved to determine what media resources need to be setup. In the second step, one sends the necessary SETUP requests and the PLAY request to initiate media delivery.

Client C requests a presentation from media server M. The movie is stored in a container file. The client has obtained an RTSP URI to the container file.

C->M: DESCRIBE rtsp://example.com/twister.3gp RTSP/2.0
CSeq: 1
User-Agent: PhonyClient/1.2

M->C: RTSP/2.0 200 OK
CSeq: 1
Server: PhonyServer/1.0
Date: Fri, 20 Dec 2013 10:20:32 +0000
Content-Type: application/sdp
Content-Length: 271
Content-Base: rtsp://example.com/twister.3gp/
Expires: Fri, 20 Dec 2013 12:20:32 +0000

```
v=0
o=- 2890844256 2890842807 IN IP4 192.0.2.5
s=RTSP Session
i=An Example of RTSP Session Usage
e=adm@example.com
c=IN IP4 0.0.0.0
a=control: *
a=range:npt=00:00:00-00:10:34.10
t=0 0
m=audio 0 RTP/AVP 0
a=control: trackID=1
m=video 0 RTP/AVP 26
a=control: trackID=4
```

```
C->M: SETUP rtsp://example.com/twister.3gp/trackID=1 RTSP/2.0
      CSeq: 2
      User-Agent: PhonyClient/1.2
      Require: play.basic
      Transport: RTP/AVP;unicast;dest_addr=":8000"/":8001"
      Accept-Ranges: npt, smpte, clock
      Pipelined-Requests: 7654

C->M: SETUP rtsp://example.com/twister.3gp/trackID=4 RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Require: play.basic
      Transport: RTP/AVP;unicast;dest_addr=":8002"/":8003"
      Accept-Ranges: npt, smpte, clock
      Pipelined-Requests: 7654

C->M: PLAY rtsp://example.com/twister.3gp/ RTSP/2.0
      CSeq: 4
      User-Agent: PhonyClient/1.2
      Range: npt=0-
      Seek-Style: RAP
      Pipelined-Requests: 7654

M->C: RTSP/2.0 200 OK
      CSeq: 2
      Server: PhonyServer/1.0
      Transport: RTP/AVP;unicast;
                dest_addr="192.0.2.53:8000"/"192.0.2.53:8001";
                src_addr="198.51.100.5:9000"/"198.51.100.5:9001";
                ssrc=93CB001E
      Session: 12345678
      Expires: Fri, 20 Dec 2013 12:20:32 +0000
      Date: Fri, 20 Dec 2013 10:20:32 +0000
      Accept-Ranges: npt
      Pipelined-Requests: 7654
      Media-Properties: Random-Access=0.2, Immutable, Unlimited

M->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/1.0
      Transport: RTP/AVP;unicast;
                dest_addr="192.0.2.53:8002"/"192.0.2.53:8003";
                src_addr="198.51.100.5:9002"/"198.51.100.5:9003";
                ssrc=A813FC13
      Session: 12345678
      Expires: Sat, 21 Dec 2013 10:20:32 +0000
      Date: Fri, 20 Dec 2013 10:20:32 +0000
      Accept-Range: NPT
```

Pipelined-Requests: 7654
Media-Properties: Random-Access=0.8, Immutable, Unlimited

M->C: RTSP/2.0 200 OK
CSeq: 4
Server: PhonyServer/1.0
Date: Fri, 20 Dec 2013 10:20:32 +0000
Session: 12345678
Range: npt=0-623.10
Seek-Style: RAP
RTP-Info: url="rtsp://example.com/twister.3gp/trackID=4"
 ssrc=0D12F123:seq=12345;rtptime=3450012,
 url="rtsp://example.com/twister.3gp/trackID=1"
 ssrc=4F312DD8:seq=54321;rtptime=2876889
Pipelined-Requests: 7654

A.3. Secured Media Session for on Demand Content

This example is basically the above example (Appendix A.2), but now including establishment of SRTP crypto contexts to get a secured media delivery. First of all, the client attempts to fetch this insecurely, but the server redirects to a URI indicating a requirement on using a secure connection for the RTSP messages. The client establishes a TCP/TLS connections and the session description is retrieved to determine what media resources need to be setup. In the this session description secure media (SRTP) is indicated. In the next step, the client sends the necessary SETUP requests including MIKEY messages. This is pipeline with a PLAY request to initiate media delivery.

Client C requests a presentation from media server M. The movie is stored in a container file. The client has obtained an RTSP URI to the container file.

Note: The MIKEY messages below are not valid MIKEY message and are BASE64 encoded random data to represent where the MIKEY messages would go.

C->M: DESCRIBE rtsp://example.com/twister.3gp RTSP/2.0
CSeq: 1
User-Agent: PhonyClient/1.2

M->C: RTSP/2.0 301 Moved Permanently
CSeq: 1
Server: PhonyServer/1.0
Date: Fri, 20 Dec 2013 10:25:32 +0000
Location: rtsp://example.com/twister.3gp

C->M: Establish TCP/TLS connection and verify server's
certificate that is represents example.com.
Used for all below RTSP messages.

C->M: DESCRIBE rtsp://example.com/twister.3gp RTSP/2.0
CSeq: 2
User-Agent: PhonyClient/1.2

M->C: RTSP/2.0 200 OK
CSeq: 2
Server: PhonyServer/1.0
Date: Fri, 20 Dec 2013 10:25:33 +0000
Content-Type: application/sdp
Content-Length: 271
Content-Base: rtsp://example.com/twister.3gp/
Expires: Fri, 20 Dec 2013 12:25:33 +0000

v=0
o=- 2890844256 2890842807 IN IP4 192.0.2.5
s=RTSP Session
i=An Example of RTSP Session Usage
e=adm@example.com
c=IN IP4 0.0.0.0
a=control: *
a=range:npt=00:00:00-00:10:34.10
t=0 0
m=audio 0 RTP/SAVP 0
a=control: trackID=1
m=video 0 RTP/SAVP 26
a=control: trackID=4

C->M: SETUP rtsp://example.com/twister.3gp/trackID=1 RTSP/2.0
CSeq: 3
User-Agent: PhonyClient/1.2
Require: play.basic
Transport: RTP/SAVP;unicast;dest_addr=":8000"/":8001";
MIKEY=VGhpcyBpcyB0aGUgZmlyc3Qgc3RyZWftcyBNSUthFWSBtZXNzYWdl
Accept-Ranges: npt, smpte, clock
Pipelined-Requests: 7654

C->M: SETUP rtsp://example.com/twister.3gp/trackID=4 RTSP/2.0
CSeq: 4
User-Agent: PhonyClient/1.2
Require: play.basic
Transport: RTP/SAVP;unicast;dest_addr=":8002"/":8003";
MIKEY=TUlLRVkgZm9yIHNOcmVhbSB0d2lzdGVyLjNncC90cmFja01EPTQ=
Accept-Ranges: npt, smpte, clock
Pipelined-Requests: 7654

```
C->M: PLAY rtsp://example.com/twister.3gp/ RTSP/2.0
      CSeq: 5
      User-Agent: PhonyClient/1.2
      Range: npt=0-
      Seek-Style: RAP
      Pipelined-Requests: 7654

M->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/1.0
      Transport: RTP/SAVP;unicast;
        dest_addr="192.0.2.53:8000"/"192.0.2.53:8001";
        src_addr="198.51.100.5:9000"/"198.51.100.5:9001";
        ssrc=93CB001E;
        MIKEY=TUllRVkgUmVzcG9uc2UgdHdpc3Rlci4zZ3AvdHJhY2tJRD0x
      Session: 12345678
      Expires: Fri, 20 Dec 2013 12:25:34 +0000
      Date: Fri, 20 Dec 2013 10:25:34 +0000
      Accept-Ranges: npt
      Pipelined-Requests: 7654
      Media-Properties: Random-Access=0.2, Immutable, Unlimited

M->C: RTSP/2.0 200 OK
      CSeq: 4
      Server: PhonyServer/1.0
      Transport: RTP/SAVP;unicast;
        dest_addr="192.0.2.53:8002"/"192.0.2.53:8003";
        src_addr="198.51.100.5:9002"/"198.51.100.5:9003";
        ssrc=A813FC13;
        MIKEY=TUllRVkgUmVzcG9uc2UgdHdpc3Rlci4zZ3AvdHJhY2tJRD00
      Session: 12345678
      Expires: Fri, 20 Dec 2013 12:25:34 +0000
      Date: Fri, 20 Dec 2013 10:25:34 +0000
      Accept-Range: NPT
      Pipelined-Requests: 7654
      Media-Properties: Random-Access=0.8, Immutable, Unlimited

M->C: RTSP/2.0 200 OK
      CSeq: 5
      Server: PhonyServer/1.0
      Date: Fri, 20 Dec 2013 10:25:34 +0000
      Session: 12345678
      Range: npt=0-623.10
      Seek-Style: RAP
      RTP-Info: url="rtsp://example.com/twister.3gp/trackID=4"
        ssrc=0D12F123:seq=12345;rtptime=3450012,
        url="rtsp://example.com/twister.3gp/trackID=1"
        ssrc=4F312DD8:seq=54321;rtptime=2876889;
```


Pipelined-Requests: 7654

A.4. Media on Demand (Unicast)

An alternative example of media on demand with a bit more tweaks is the following. Client C requests a movie distributed from two different media servers A (audio.example.com) and V (video.example.com). The media description is stored on a web server W. The media description contains descriptions of the presentation and all its streams, including the codecs that are available, dynamic RTP payload types, the protocol stack, and content information such as language or copyright restrictions. It may also give an indication about the timeline of the movie.

In this example, the client is only interested in the last part of the movie.

```
C->W: GET /twister.sdp HTTP/1.1
      Host: www.example.com
      Accept: application/sdp

W->C: HTTP/1.1 200 OK
      Date: Wed, 23 Jan 2013 15:35:06 GMT
      Content-Type: application/sdp
      Content-Length: 278
      Expires: Thu, 24 Jan 2013 15:35:06 GMT

      v=0
      o=- 2890844526 2890842807 IN IP4 198.51.100.5
      s=RTSP Session
      e=adm@example.com
      c=IN IP4 0.0.0.0
      a=range:npt=00:00:00-01:49:34
      t=0 0
      m=audio 0 RTP/AVP 0
      a=control:rtsp://audio.example.com/twister/audio.en
      m=video 0 RTP/AVP 31
      a=control:rtsp://video.example.com/twister/video

C->A: SETUP rtsp://audio.example.com/twister/audio.en RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2
      Transport: RTP/AVP/UDP;unicast;dest_addr=":3056"/":3057",
                 RTP/AVP/TCP;unicast;interleaved=0-1
      Accept-Ranges: npt, smpte, clock

A->C: RTSP/2.0 200 OK
      CSeq: 1
      Session: 12345678
      Transport: RTP/AVP/UDP;unicast;
                 dest_addr="192.0.2.53:3056"/"192.0.2.53:3057";
                 src_addr="198.51.100.5:5000"/"198.51.100.5:5001"
      Date: Wed, 23 Jan 2013 15:35:12 +0000
      Server: PhonyServer/1.0
      Expires: Thu, 24 Jan 2013 15:35:12 +0000
      Cache-Control: public
      Accept-Ranges: npt, smpte
      Media-Properties: Random-Access=0.02, Immutable, Unlimited
```

```
C->V: SETUP rtsp://video.example.com/twister/video RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2
      Transport: RTP/AVP/UDP;unicast;
                  dest_addr="192.0.2.53:3058"/"192.0.2.53:3059",
                  RTP/AVP/TCP;unicast;interleaved=0-1
      Accept-Ranges: npt, smpte, clock

V->C: RTSP/2.0 200 OK
      CSeq: 1
      Session: 23456789
      Transport: RTP/AVP/UDP;unicast;
                  dest_addr="192.0.2.53:3058"/"192.0.2.53:3059";
                  src_addr="198.51.100.5:5002"/"198.51.100.5:5003"
      Date: Wed, 23 Jan 2013 15:35:12 +0000
      Server: PhonyServer/1.0
      Cache-Control: public
      Expires: Thu, 24 Jan 2013 15:35:12 +0000
      Accept-Ranges: npt, smpte
      Media-Properties: Random-Access=1.2, Immutable, Unlimited

C->V: PLAY rtsp://video.example.com/twister/video RTSP/2.0
      CSeq: 2
      User-Agent: PhonyClient/1.2
      Session: 23456789
      Range: smpte=0:10:00-

V->C: RTSP/2.0 200 OK
      CSeq: 2
      Session: 23456789
      Range: smpte=0:10:00-1:49:23
      Seek-Style: First-Prior
      RTP-Info: url="rtsp://video.example.com/twister/video"
                  ssrc=A17E189D;seq=12312232;rtptime=78712811
      Server: PhonyServer/2.0
      Date: Wed, 23 Jan 2013 15:35:13 +0000
```

```
C->A: PLAY rtsp://audio.example.com/twister/audio.en RTSP/2.0
      CSeq: 2
      User-Agent: PhonyClient/1.2
      Session: 12345678
      Range: smpte=0:10:00-

A->C: RTSP/2.0 200 OK
      CSeq: 2
      Session: 12345678
      Range: smpte=0:10:00-1:49:23
      Seek-Style: First-Prior
      RTP-Info: url="rtsp://audio.example.com/twister/audio.en"
                 ssrc=3D124F01:seq=876655;rtptime=1032181
      Server: PhonyServer/1.0
      Date: Wed, 23 Jan 2013 15:35:13 +0000


C->A: TEARDOWN rtsp://audio.example.com/twister/audio.en RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Session: 12345678

A->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/1.0
      Date: Wed, 23 Jan 2013 15:36:52 +0000


C->V: TEARDOWN rtsp://video.example.com/twister/video RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Session: 23456789

V->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/2.0
      Date: Wed, 23 Jan 2013 15:36:52 +0000
```

Even though the audio and video track are on two different servers that may start at slightly different times and may drift with respect to each other over time, the client can perform initial synchronization of the two media using RTP-Info and Range received in the PLAY responses. If the two servers are time synchronized the RTCP packets can also be used to maintain synchronization.

A.5. Single Stream Container Files

Some RTSP servers may treat all files as though they are "container files", yet other servers may not support such a concept. Because of this, clients needs to use the rules set forth in the session description for Request-URIs, rather than assuming that a consistent URI may always be used throughout. Below is an example of how a multi-stream server might expect a single-stream file to be served:

```
C->S: DESCRIBE rtsp://foo.example.com/test.wav RTSP/2.0
      Accept: application/x-rtsp-mh, application/sdp
      CSeq: 1
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 1
      Content-base: rtsp://foo.example.com/test.wav/
      Content-type: application/sdp
      Content-length: 163
      Server: PhonyServer/1.0
      Date: Wed, 23 Jan 2013 15:36:52 +0000
      Expires: Thu, 24 Jan 2013 15:36:52 +0000
```

```
v=0
o=- 872653257 872653257 IN IP4 192.0.2.5
s=mu-law wave file
i=audio test
c=IN IP4 0.0.0.0
t=0 0
a=control: *
m=audio 0 RTP/AVP 0
a=control:streamid=0
```

```
C->S: SETUP rtsp://foo.example.com/test.wav/streamid=0 RTSP/2.0
      Transport: RTP/AVP/UDP;unicast;
            dest_addr=":6970"/":6971";mode="PLAY"
      CSeq: 2
      User-Agent: PhonyClient/1.2
      Accept-Ranges: npt, smpte, clock
```

```
S->C: RTSP/2.0 200 OK
      Transport: RTP/AVP/UDP;unicast;
            dest_addr="192.0.2.53:6970"/"192.0.2.53:6971";
            src_addr="198.51.100.5:6970"/"198.51.100.5:6971";
            mode="PLAY";ssrc=EAB98712
      CSeq: 2
      Session: 2034820394
      Expires: Thu, 24 Jan 2013 15:36:52 +0000
      Server: PhonyServer/1.0
      Date: Wed, 23 Jan 2013 15:36:52 +0000
      Accept-Ranges: npt
      Media-Properties: Random-Acces=0.5, Immutable, Unlimited
```

```
C->S: PLAY rtsp://foo.example.com/test.wav/ RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Session: 2034820394
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/1.0
      Date: Wed, 23 Jan 2013 15:36:52 +0000
      Session: 2034820394
      Range: npt=0-600
      Seek-Style: RAP
      RTP-Info: url="rtsp://foo.example.com/test.wav/streamid=0"
            ssrc=0D12F123:seq=981888;rtptime=3781123
```

Note the different URI in the SETUP command, and then the switch back to the aggregate URI in the PLAY command. This makes complete sense when there are multiple streams with aggregate control, but is less than intuitive in the special case where the number of streams is one. However, the server has declared the aggregated control URI in the SDP and therefore this is legal.

In this case, it is also required that servers accept implementations that use the non-aggregated interpretation and use the individual media URI, like this:

```
C->S: PLAY rtsp://example.com/test.wav/streamid=0 RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Session: 2034820394
```

A.6. Live Media Presentation Using Multicast

The media server M chooses the multicast address and port. Here, it is assumed that the web server only contains a pointer to the full description, while the media server M maintains the full description.

```
C->W: GET /sessions.html HTTP/1.1
      Host: www.example.com
```

```
W->C: HTTP/1.1 200 OK
      Content-Type: text/html
```

```
<html>
...
  <a href "rtsp://live.example.com/concert/audio">
    Streamed Live Music performance </a>
...
</html>
```

```
C->M: DESCRIBE rtsp://live.example.com/concert/audio RTSP/2.0
      CSeq: 1
      Supported: play.basic, play.scale
      User-Agent: PhonyClient/1.2
```

```
M->C: RTSP/2.0 200 OK
      CSeq: 1
      Content-Type: application/sdp
      Content-Length: 183
      Server: PhonyServer/1.0
      Date: Wed, 23 Jan 2013 15:36:52 +0000
      Supported: play.basic
```

```
v=0
o=- 2890844526 2890842807 IN IP4 192.0.2.5
s=RTSP Session
t=0 0
m=audio 3456 RTP/AVP 0
c=IN IP4 233.252.0.54/16
a=control: rtsp://live.example.com/concert/audio
a=range:npt=0-
```

```
C->M: SETUP rtsp://live.example.com/concert/audio RTSP/2.0
      CSeq: 2
      Transport: RTP/AVP;multicast;
                dest_addr="233.252.0.54:3456"/"233.252.0.54:3457";ttl=16
      Accept-Ranges: npt, smpte, clock
      User-Agent: PhonyClient/1.2

M->C: RTSP/2.0 200 OK
      CSeq: 2
      Server: PhonyServer/1.0
      Date: Wed, 23 Jan 2013 15:36:52 +0000
      Transport: RTP/AVP;multicast;
                dest_addr="233.252.0.54:3456"/"233.252.0.54:3457";ttl=16
                ;ssrc=4D12AB92/0DF876A3
      Session: 0456804596
      Accept-Ranges: npt, clock
      Media-Properties: No-Seeking, Time-Progressing, Time-Duration=0

C->M: PLAY rtsp://live.example.com/concert/audio RTSP/2.0
      CSeq: 3
      Session: 0456804596
      User-Agent: PhonyClient/1.2

M->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/1.0
      Date: Wed, 23 Jan 2013 15:36:52 +0000
      Session: 0456804596
      Seek-Style: Next
      Range:npt=1256-
      RTP-Info: url="rtsp://live.example.com/concert/audio"
                ssrc=0D12F123;seq=1473; rtptime=80000
```

A.7. Capability Negotiation

This example illustrates how the client and server determine their capability to support a special feature, in this case "play.scale". The server, through the clients request and the included Supported header, learns the client supports RTSP 2.0, and also supports the playback time scaling feature of RTSP. The server's response contains the following feature related information to the client; it supports the basic media delivery functions (play.basic), the extended functionality of time scaling of content (play.scale), and one "example.com" proprietary feature (com.example.flight). The client also learns the methods supported (Public header) by the server for the indicated resource.


```
C->S: OPTIONS rtsp://media.example.com/movie/twister.3gp RTSP/2.0
      CSeq: 1
      Supported: play.basic, play.scale
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 1
      Public: OPTIONS, SETUP, PLAY, PAUSE, TEARDOWN, DESCRIBE, GET_PARAMETER
      Allow: OPTIONS, SETUP, PLAY, PAUSE, TEARDOWN, DESCRIBE
      Server: PhonyServer/2.0
      Supported: play.basic, play.scale, com.example.flight
```

When the client sends its SETUP request it tells the server that it requires support of the play.scale feature for this session by including the Require header.

```
C->S: SETUP rtsp://media.example.com/twister.3gp/trackID=1 RTSP/2.0
      CSeq: 3
      User-Agent: PhonyClient/1.2
      Transport: RTP/AVP/UDP;unicast;
                 dest_addr="192.0.2.53:3056"/"192.0.2.53:3057",
                 RTP/AVP/TCP;unicast;interleaved=0-1
      Require: play.scale
      Accept-Ranges: npt, smpte, clock
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 3
      Session: 12345678
      Transport: RTP/AVP/UDP;unicast;
                 dest_addr="192.0.2.53:3056"/"192.0.2.53:3057";
                 src_addr="198.51.100.5:5000"/"198.51.100.5:5001"
      Server: PhonyServer/2.0
      Accept-Ranges: npt, smpte
      Media-Properties: Random-Access=0.8, Immutable, Unlimited
```

Appendix B. RTSP Protocol State Machine

The RTSP session state machine describes the behavior of the protocol from RTSP session initialization through RTSP session termination. It is probably easiest to think of this as the server's state and then view the the client as needing to track what it believes the server's state will be based on sent or received RTSP messages. Thus in most cases the state tables below can be read as: If the client does X, and assuming it fulfills any pre-requisite(s), the (server) state will move to the new state and the indicated response will be returned. However, there are also server to client notifications or requests, where the action describes what notification or request

will occur, its requisites and what new state will result after the server has received the response, as well as describing the client's response to the action.

The State machine is defined on a per session basis which is uniquely identified by the RTSP session identifier. The session may contain one or more media streams depending on state. If a single media stream is part of the session it is in non-aggregated control. If two or more is part of the session it is in aggregated control.

The below state machine is an informative description of the protocols behavior. In case of ambiguity with the earlier parts of this specification, the description in the earlier parts take precedence.

B.1. States

The state machine contains three states, described below. For each state there exists a table which shows which requests and events are allowed and whether they will result in a state change.

Init: Initial state no session exists.

Ready: Session is ready to start playing.

Play: Session is playing, i.e., sending media stream data in the direction S->C.

B.2. State variables

This representation of the state machine needs more than its state to work. A small number of variables are also needed and they are explained below.

NRM: The number of media streams part of this session.

RP: Resume point, the point in the presentation time line at which a request to continue playing will resume from. A time format for the variable is not mandated.

B.3. Abbreviations

To make the state tables more compact a number of abbreviations are used, which are explained below.

IFI: IF Implemented.

md: Media

PP: Pause Point, the point in the presentation time line at which the presentation was paused.

Prs: Presentation, the complete multimedia presentation.

RedP: Redirect Point, the point in the presentation time line at which a REDIRECT was specified to occur.

SES: Session.

B.4. State Tables

This section contains a table for each state. The table contains all the requests and events that this state is allowed to act on. The events which are method names are, unless noted, requests with the given method in the direction client to server (C->S). In some cases there exist one or more requisite. The response column tells what type of response actions should be performed. Possible actions that are requested for an event include: response codes, e.g., 200, headers that need to be included in the response, setting of state variables, or setting of other session related parameters. The new state column tells which state the state machine changes to.

The response to a valid request meeting the requisites is normally a 2xx (SUCCESS) unless otherwise noted in the response column. The exceptions need to be given a response according to the response column. If the request does not meet the requisite, is erroneous or some other type of error occurs, the appropriate response code is to be sent. If the response code is a 4xx the session state is unchanged. A response code of 3rr will result in that the session is ended and its state is changed to Init. A response code of 304 results in no state change. However, there are restrictions to when a 3rr response may be used. A 5xx response does not result in any change of the session state, except if the error is not possible to recover from. A unrecoverable error results in the ending of the session. As it in the general case can't be determined if it was a unrecoverable error or not the client will be required to test. In the case that the next request after a 5xx is responded with 454 (Session Not Found) the client knows that the session has ended. For any request message that cannot be responded to within the time defined in Section 10.4, a 100 response must be sent.

The server will timeout the session after the period of time specified in the SETUP response, if no activity from the client is detected. Therefore there exists a timeout event for all states except Init.

In the case that $NRM = 1$ the presentation URI is equal to the media URI or a specified presentation URI. For $NRM > 1$ the presentation URI needs to be other than any of the medias that are part of the session. This applies to all states.

Event	Prerequisite	Response
DESCRIBE	Needs REDIRECT	3rr, Redirect
DESCRIBE		200, Session description
OPTIONS	Session ID	200, Reset session timeout timer
OPTIONS		200
SET_PARAMETER	Valid parameter	200, change value of parameter
GET_PARAMETER	Valid parameter	200, return value of parameter

Table 13: None state-machine changing events

The methods in Table 13 do not have any effect on the state machine or the state variables. However, some methods do change other session related parameters, for example SET_PARAMETER which will set the parameter(s) specified in its body. Also all of these methods that allow Session header will also update the keep-alive timer for the session.

Action	Requisite	New State	Response
SETUP		Ready	$NRM=1$, $RP=0.0$
SETUP	Needs Redirect	Init	3rr Redirect
S -> C: REDIRECT	No Session hdr	Init	Terminate all SES

Table 14: State: Init

The initial state of the state machine, see Table 14 can only be left by processing a correct SETUP request. As seen in the table the two state variables are also set by a correct request. This table also shows that a correct SETUP can in some cases be redirected to another URI and/or server by a 3rr response.

Action	Requisite	New State	Response
SETUP	New URI	Ready	NRM +=1
SETUP	URI Setup prior	Ready	Change transport param
TEARDOWN	Prs URI,	Init	No session hdr, NRM = 0
TEARDOWN	md URI,NRM=1	Init	No Session hdr, NRM = 0
TEARDOWN	md URI,NRM>1	Ready	Session hdr, NRM -= 1
PLAY	Prs URI, No range	Play	Play from RP
PLAY	Prs URI, Range	Play	According to range
PLAY	md URI, NRM=1, Range	Play	According to range
PLAY	md URI, NRM=1	Play	Play from RP
PAUSE	Prs URI	Ready	Return PP
SC:REDIRECT	Terminate-Reason	Ready	Set RedP
SC:REDIRECT	No Terminate-Reason time parameter	Init	Session is removed
Timeout		Init	
RedP reached		Init	TEARDOWN of session

Table 15: State: Ready

In the Ready state, see Table 15, some of the actions are depending on the number of media streams (NRM) in the session, i.e., aggregated or non-aggregated control. A SETUP request in the Ready state can either add one more media stream to the session or, if the media stream (same URI) already is part of the session, change the

transport parameters. TEARDOWN is depending on both the Request-URI and the number of media streams within the session. If the Request-URI is the presentations URI the whole session is torn down. If a media URI is used in the TEARDOWN request and more than one media exists in the session, the session will remain and a session header is returned in the response. If only a single media stream remains in the session when performing a TEARDOWN with a media URI the session is removed. The number of media streams remaining after tearing down a media stream determines the new state.

Action	Requisite	New State	Response
PAUSE	Prs URI	Ready	Set RP to present point
End of media	All media	Play	Set RP = End of media
End of range		Play	Set RP = End of range
PLAY	Prs URI, No range	Play	Play from present point
PLAY	Prs URI, Range	Play	According to range
SC:PLAY_NOTIFY		Play	200
SETUP	New URI	Play	455
SETUP	Setuped URI	Play	455
SETUP	Setuped URI, IFI	Play	Change transport param.
TEARDOWN	Prs URI	Init	No session hdr
TEARDOWN	md URI,NRM=1	Init	No Session hdr, NRM=0
TEARDOWN	md URI	Play	455
SC:REDIRECT	Terminate Reason with Time parameter	Play	Set RedP

SC:REDIRECT		Init	Session is removed
RedP reached		Init	TEARDOWN of session
Timeout		Init	Stop Media playout

Table 16: State: Play

The Play state table, see Table 16, contains a number of requests that need a presentation URI (labeled as Prs URI) to work on (i.e., the presentation URI has to be used as the Request-URI). This is due to the exclusion of non-aggregated stream control in sessions with more than one media stream.

To avoid inconsistencies between the client and server, automatic state transitions are avoided. This can be seen at for example "End of media" event when all media has finished playing, the session still remains in Play state. An explicit PAUSE request needs to be sent to change the state to Ready. It may appear that there exist automatic transitions in "RedP reached" and "PP reached". However, they are requested and acknowledged before they take place. The time at which the transition will happen is known by looking at the range header. If the client sends a request close in time to these transitions it needs to be prepared for receiving error messages, as the state may or may not have changed.

Appendix C. Media Transport Alternatives

This section defines how certain combinations of protocols, profiles and lower transports are used. This includes the usage of the Transport header's source and destination address parameters "src_addr" and "dest_addr".

C.1. RTP

This section defines the interaction of RTSP with respect to the RTP protocol [RFC3550]. It also defines any necessary media transport signaling with regards to RTP.

The available RTP profiles and lower layer transports are described below along with rules on signaling the available combinations.

C.1.1.1. AVP

The usage of the "RTP Profile for Audio and Video Conferences with Minimal Control" [RFC3551] when using RTP for media transport over different lower layer transport protocols is defined below in regards to RTSP.

One such case is defined within this document: the use of embedded (interleaved) binary data as defined in Section 14. The usage of this method is indicated by including the "interleaved" parameter.

When using embedded binary data the "src_addr" and "dest_addr" MUST NOT be used. This addressing and multiplexing is used as defined with use of channel numbers and the interleaved parameter.

C.1.1.2. AVP/UDP

This part describes sending of RTP [RFC3550] over lower transport layer UDP [RFC0768] according to the profile "RTP Profile for Audio and Video Conferences with Minimal Control" defined in RFC 3551 [RFC3551]. Implementations of RTP/AVP/UDP MUST implement RTCP (Appendix C.1.6). This profile requires one or two uni- or bi-directional UDP flows per media stream. The first UDP flow is for RTP and the second is for RTCP. Multiplexing of RTP and RTCP (Appendix C.1.6.4) MAY be used, in which case a single UDP flow is used for both parts. Embedding of RTP data with the RTSP messages, in accordance with Section 14, SHOULD NOT be performed when RTSP messages are transported over unreliable transport protocols, like UDP [RFC0768].

The RTP/UDP and RTCP/UDP flows can be established using the Transport header's "src_addr", and "dest_addr" parameters.

In RTSP PLAY mode, the transmission of RTP packets from client to server is unspecified. The behavior in regards to such RTP packets MAY be defined in future.

The "src_addr" and "dest_addr" parameters are used in the following way for media delivery and playback mode, i.e., Mode=PLAY:

- o The "src_addr" and "dest_addr" parameters MUST contain either 1 or 2 address specifications. Note that two address specifications MAY be provided even if RTP and RTCP multiplexing is negotiated.
- o Each address specification for RTP/AVP/UDP or RTP/AVP/TCP MUST contain either:
 - * both an address and a port number, or

* a port number without an address.

- o The first address specification given in either of the parameters applies to the RTP stream. The second specification if present applies to the RTCP stream, unless in case RTP and RTCP multiplexing is negotiated where both RTP and RTCP will use the first specification.
- o The RTP/UDP packets from the server to the client MUST be sent to the address and port given by the first address specification of the "dest_addr" parameter.
- o The RTCP/UDP packets from the server to the client MUST be sent to the address and port given by the second address specification of the "dest_addr" parameter, unless RTP and RTCP multiplexing has been negotiated, in which case RTCP MUST be sent to the first address specification. If no second pair is specified and RTP and RTCP multiplexing has not been negotiated, RTCP MUST NOT be sent.
- o The RTCP/UDP packets from the client to the server MUST be sent to the address and port given by the second address specification of the "src_addr" parameter, unless RTP and RTCP multiplexing has been negotiated, in which case RTCP MUST be sent to the first address specification. If no second pair is specified and RTP and RTCP multiplexing has not been negotiated, RTCP MUST NOT be sent.
- o The RTP/UDP packets from the client to the server MUST be sent to the address and port given by the first address specification of the "src_addr" parameter.
- o RTP and RTCP Packets SHOULD be sent from the corresponding receiver port, i.e., RTCP packets from the server should be sent from the "src_addr" parameters second address port pair, unless RTP and RTCP multiplexing has been negotiated in which case the first address port pair is used.

C.1.3. AVPF/UDP

The RTP profile "Extended RTP Profile for RTCP-based Feedback (RTP/AVPF)" [RFC4585] MAY be used as RTP profiles in sessions using RTP. All that is defined for AVP MUST also apply for AVPF.

The usage of AVPF is indicated by the media initialization protocol used. In the case of SDP it is indicated by media lines (m=) containing the profile RTP/AVPF. That SDP MAY also contain further AVPF related SDP attributes configuring the AVPF session regarding reporting interval and feedback messages to be used [RFC4585]. This configuration MUST be followed.

C.1.4. SAVP/UDP

The RTP profile "The Secure Real-time Transport Protocol (SRTP)" [RFC3711] is an RTP profile (SAVP) that MAY be used in RTSP sessions using RTP. All that is defined for AVP MUST also apply for SAVP.

The usage of SRTP requires that a security context is established. The default key-management unless otherwise signalled SHALL be MIKEY in RSA-R mode as defined in Appendix C.1.4.1, and not according to the procedure defined in "Key Management Extensions for Session Description Protocol (SDP) and Real Time Streaming Protocol (RTSP)" [RFC4567]. The reason is that RFC 4567 sends the initial MIKEY message in SDP, thus both requiring the usage of the DESCRIBE method and forcing the server to keep state for clients performing DESCRIBE in anticipation that they might require key management.

MIKEY is selected as default method for establishing SRTP cryptographic context within an RTSP session as it can be embedded in the RTSP messages, while still ensuring confidentiality of content of the keying material, even when using hop-by-hop TLS security for the RTSP messages. This method does also support pipelining of the RTSP messages.

C.1.4.1. MIKEY Key Establishment

This method for using MIKEY [RFC3830] to establish the SRTP cryptographic context is initiated in the client's SETUP request, and the server's response to the SETUP carries the MIKEY response. This ensures that the crypto context establishment happens simultaneously with the establishment of the media stream being protected. By using MIKEY's RSA-R mode [RFC4738] the client can be the initiator and still allow the server to set the parameters in accordance with the actual media stream.

The SRTP cryptographic context establishment is done according to the following process:

1. The client determines that SAVP or SAVPF shall be used from media description format, e.g., SDP. If no other key management method is explicitly signalled, then MIKEY SHALL be used as defined herein. The use of SRTP with RTSP is only defined with MIKEY with keys established as defined in this Section. Future documents may define how an RTSP implementation treats SDP that indicates some other key mechanism to be used. The need for such specification include [RFC4567] that is not defined for use in RTSP 2.0 within this document.

2. The client SHALL establish a TLS connection for RTSP messages, directly or hop by hop with the server. If hop-by-hop TLS security is used, the User method SHALL be indicated in the Accept-Credentials header. Note that using hop-by-hop does allow the proxy to insert itself as a man in the middle also in the MIKEY exchange by providing one of its certificates, rather than the server's in the Connection-Credentials header. The client SHALL therefore validate the server certificate.
3. The client retrieves the server's certificate from a direct TLS connection, or if hop by hop from Connection-Credentials header. The client then checks that the server certificate is valid and belongs to the server.
4. The client forms the MIKEY Initiator message using RSA-R mode in unicast mode as specified in [RFC4738]. The client SHOULD use the same certificate for TLS and in MIKEY to enable the server to bind the two together. The client's certificate SHALL be included in the MIKEY message. The client SHALL indicate its SRTP capabilities in the message.
5. The MIKEY message from the previous step is base64 [RFC4648] encoded and becomes the value of the MIKEY parameter that is included in the transport specification(s) that specifies a SRTP based profile (SAVP, SAVPF) in the SETUP request.
6. Any proxy encountering the MIKEY parameter SHALL forward it without modification. A proxy requiring to understand transport specification which doesn't support SAVP/SAVPF with MIKEY will discard the whole transport specification. Most types of proxies can easily support SAVP and SAVPF with MIKEY. If possible bypassing the proxy should be tried.
7. The server upon receiving the SETUP request, will need to decide upon the transport specification to use, if multiple are included by the client. In the determination of which transport specifications that are supported and preferred, the server SHOULD decode the MIKEY message to take the embedded SRTP parameters into account. If all transport specs require SRTP but no MIKEY parameter or other supported keying method is included, the server SHALL respond with 403.
8. Upon generating a response the following outcomes can occur:
 - * A transport spec not using SRTP and MIKEY is selected. Thus the response will not contain any MIKEY parameter.

- * A transport spec using SRTP and MIKEY is selected but an error is encountered in the MIKEY processing. In that case an RTSP error response code of 466 "Key Management Error" SHALL be used. A MIKEY message describing the error MAY be included.
 - * A transport spec using SRTP and MIKEY is selected and a MIKEY response message can be created. The server SHOULD use the same certificate for TLS and in MIKEY to enable client to bind the two together. If a different certificate is used it SHALL be included in the MIKEY message. It is RECOMMENDED that the envelope key cache type is set to 'Cache' and that a single envelope key is reused for all MIKEY messages to the client. That message is included in the MIKEY parameter part of the single selected transport specification in the SETUP response. The server will set the SRTP parameters as preferred for this media stream within the supported range by the client.
9. The server transmits the SETUP response back to the client.
 10. The client receives the SETUP response and if the response code indicates a successful request it decodes the MIKEY message and establishes the SRTP cryptographic context from the parameters in the MIKEY response.

In the above method the client's certificate may be self-signed in cases where the client's identity is not necessary to authenticate and the security goal is only to ensure that the RTSP signaling client is the same as the one receiving the SRTP security context.

C.1.5. SAVPF/UDP

The RTP profile "Extended Secure RTP Profile for RTCP-based Feedback (RTP/SAVPF)" [RFC5124] is an RTP profile (SAVPF) that MAY be used in RTSP sessions using RTP. All that is defined for AVPF MUST also apply for SAVPF.

The usage of SRTP requires that a cryptographic context is established. The default mechanism for establishing that security association is to use MIKEY[RFC3830] with RTSP as defined in Appendix C.1.4.1.

C.1.6. RTCP usage with RTSP

RTCP has several usages when RTP is used for media transport as explained below. Due to that RTCP MUST be supported if an RTSP agent handles RTP.

C.1.6.1. Media synchronization

RTCP provides media synchronization and clock drift compensation. The initial media synchronization is available from RTP-Info header. However, to be able to handle any clock drift between the media streams, RTCP is needed.

C.1.6.2. RTSP Session keep-alive

RTCP traffic from the RTSP client to the RTSP server MUST function as keep-alive. This requires an RTSP server supporting RTP to use the received RTCP packets as indications that the client desires the related RTSP session to be kept alive.

C.1.6.3. Bit-rate adaption

RTCP Receiver reports and any additional feedback from the client MUST be used to adapt the bit-rate used over the transport for all cases when RTP is sent over UDP. An RTP sender without reserved resources MUST NOT use more than its fair share of the available resources. This can be determined by comparing on short to medium term (some seconds) the used bit-rate and adapt it so that the RTP sender sends at a bit-rate comparable to what a TCP sender would achieve on average over the same path.

To ensure that the implementation's adaptation mechanism has a well defined outer envelope, all implementations using a non-congestion controlled unicast transport protocol, like UDP, MUST implement Multimedia Congestion Control: Circuit Breakers for Unicast RTP Sessions [I-D.ietf-avtcore-rtp-circuit-breakers].

C.1.6.4. RTP and RTCP Multiplexing

RTSP can be used to negotiate the usage of RTP and RTCP multiplexing as described in [RFC5761]. This allows servers and client to reduce the amount of resources required for the session by only requiring one underlying transport stream per media stream instead of two when using RTP and RTCP. This lessens the server port consumption and also the necessary state and keep-alive work when operating across Network and Address Translators [RFC2663].

Content must be prepared with some consideration for RTP and RTCP multiplexing, mainly ensuring that the RTP payload types used do not collide with the ones used for RTCP packet types. This option likely needs explicit support from the content unless the RTP payload types can be remapped by the server and that is correctly reflected in the session description. Beyond that support of this feature should come at little cost and much gain.

It is recommended that if the content and server support RTP and RTCP multiplexing that this is indicated in the session description, for example using the SDP attribute "a=rtcp-mux". If the SDP message contains the a=rtcp-mux attribute for a media stream, the server MUST support RTP and RTCP multiplexing. If indicated or otherwise desired by the client it can include the Transport parameter "RTCP-mux" in any transport specification where it desires to use RTCP-mux. The server will indicate if it supports RTCP-mux. Servers and Clients SHOULD support RTP and RTCP multiplexing.

For capability exchange, an RTSP feature tag for RTP and RTCP multiplexing is defined: "setup.rtp.rtcp.mux".

To minimize the risk of negotiation failure while using RTP and RTCP multiplexing some recommendations are here provided. If the session description includes explicit indication of support (a=rtcp-mux in SDP), then a RTSP agent can safely create a SETUP request with a transport specification with only a single dest_addr parameter address specification. If no such explicit indication is provided, then even if the feature tag "setup.rtp.rtcp.mux" is provided in a Supported header by the RTSP server or the feature tag included in the Required header in the SETUP request, the media resource may not support RTP and RTCP multiplexing. Thus, to maximize the probability of successful negotiation the RTSP agent is recommended to include two dest_addr parameter address specifications in the first or first set (if pipelining is used) of SETUP request(s) for any media resource aggregate. That way the RTSP server can either accept RTP and RTCP multiplexing and only use the first address specification, and if not use both specifications. The RTSP agent after having received the response for a successful negotiation of the usage of RTP and RTCP multiplexing, can then release the resources associated with the second address specification.

C.2. RTP over TCP

Transport of RTP over TCP can be done in two ways: over independent TCP connections using RFC 4571 [RFC4571] or interleaved in the RTSP connection. In both cases the protocol MUST be "rtp" and the lower layer MUST be TCP. The profile may be any of the above specified ones; AVP, AVPF, SAVP or SAVPF.

C.2.1. Interleaved RTP over TCP

The use of embedded (interleaved) binary data transported on the RTSP connection is possible as specified in Section 14. When using this declared combination of interleaved binary data the RTSP messages MUST be transported over TCP. TLS may or may not be used. If TLS is used both RTSP messages and the binary data will be protected by TLS.

One should, however, consider that this will result in all media streams go through any proxy. Using independent TCP connections can avoid that issue.

C.2.2. RTP over independent TCP

In this Appendix, it is described the sending of RTP [RFC3550] over lower transport layer TCP [RFC0793] according to "Framing Real-time Transport Protocol (RTP) and RTP Control Protocol (RTCP) Packets over Connection-Oriented Transport" [RFC4571]. This Appendix adapts the guidelines for using RTP over TCP within SIP/SDP [RFC4145] to work with RTSP.

A client codes the support of RTP over independent TCP by specifying an RTP/AVP/TCP transport option without an interleaved parameter in the Transport line of a SETUP request. This transport option **MUST** include the "unicast" parameter.

If the client wishes to use RTP with RTCP, two address specifications needs to be included in the dest_addr parameter. If the client wishes to use RTP without RTCP, one address specification is included in the dest_addr parameter. If the client wishes to multiplex RTP and RTCP on a single transport flow (see Appendix C.1.6.4), one or two address specifications are included in the dest_addr parameter in addition to the RTCP-mux transport parameter. Two address specifications are allowed to allow successful negotiation when server or content can't support RTP and RTCP multiplexing. Ordering rules of dest_addr ports follow the rules for RTP/AVP/UDP.

If the client wishes to play the active role in initiating the TCP connection, it **MAY** set the "setup" parameter (See Section 18.54) on the Transport line to be "active", or it **MAY** omit the setup parameter, as active is the default. If the client signals the active role, the ports in the address specifications in the dest_addr parameter **MUST** be set to 9 (the discard port).

If the client wishes to play the passive role in TCP connection initiation, it **MUST** set the "setup" parameter on the Transport line to be "passive". If the client is able to assume the active or the passive role, it **MUST** set the "setup" parameter on the Transport line to be "actpass". In either case, the dest_addr parameter's address specification port value for RTP **MUST** be set to the TCP port number on which the client is expecting to receive the TCP connection for RTP, and the dest_addr's address specification port value for RTCP **MUST** be set to the TCP port number on which the client is expecting to receive the TCP connection for RTCP. In the case that the client wishes to multiplex RTP and RTCP on a single transport flow, the RTCP-mux parameter is included and one or two dest_addr parameter

address specifications are included, as mentioned earlier in this section.

If upon receipt of a non-interleaved RTP/AVP/TCP SETUP request, a server decides to accept this requested option, the 2xx reply MUST contain a Transport option that specifies RTP/AVP/TCP (without using the interleaved parameter, and with using the unicast parameter). The `dest_addr` parameter value MUST be echoed from the parameter value in the client request unless the destination address (only port) was not provided in which case the server MAY include the source address of the RTSP TCP connection with the port number unchanged.

In addition, the server reply MUST set the `setup` parameter on the Transport line, to indicate the role the server will play in the connection setup. Permissible values are "active" (if a client set "setup" to "passive" or "actpass") and "passive" (if a client set "setup" to "active" or "actpass").

If a server sets "setup" to "passive", the "src_addr" in the reply MUST indicate the ports the server is willing to receive an TCP connection for RTP and (if the client requested an TCP connection for RTCP by specifying two `dest_addr` address specifications) an TCP/RTCP connection. If a server sets "setup" to "active", the ports specified in "src_addr" address specifications MUST be set to 9. The server MAY use the "ssrc" parameter, following the guidance in Section 18.54. The server sets only one address specification in the case that the client has indicated only a single address specification or in case RTP and RTCP multiplexing was requested and accepted by server. Port ordering for `src_addr` follows the rules for RTP/AVP/UDP.

Servers MUST support taking the passive role and MAY support taking the active role. Servers with a public IP address take the passive role, thus enabling clients behind NATs and Firewalls a better chance of successful connect to the server by actively connecting outwards. Therefore the clients are RECOMMENDED to take the active role.

After sending (receiving) a 2xx reply for a SETUP method for a non-interleaved RTP/AVP/TCP media stream, the active party SHOULD initiate the TCP connection as soon as possible. The client MUST NOT send a PLAY request prior to the establishment of all the TCP connections negotiated using SETUP for the session. In case the server receives a PLAY request in a session that has not yet established all the TCP connections, it MUST respond using the 464 "Data Transport Not Ready Yet" (Section 17.4.29) error code.

Once the PLAY request for a media resource transported over non-interleaved RTP/AVP/TCP occurs, media begins to flow from server to

client over the RTP TCP connection, and RTCP packets flow bidirectionally over the RTCP TCP connection. Unless RTP and RTCP multiplexing has been negotiated in which case RTP and RTCP will flow over a common TCP connection. As in the RTP/UDP case, client to server traffic on a RTP only TCP session is unspecified by this memo. The packets that travel on these connections MUST be framed using the protocol defined in [RFC4571], not by the framing defined for interleaving RTP over the RTSP connection defined in Section 14.

A successful PAUSE request for a media being transported over RTP/AVP/TCP pauses the flow of packets over the connections, without closing the connections. A successful TEARDOWN request signals that the TCP connections for RTP and RTCP are to be closed by the RTSP client as soon as possible.

Subsequent SETUP requests on an already-SETUP RTP/AVP/TCP URI may be ambiguous in the following way: does the client wish to open up new TCP connection for RTP or RTCP for the URI, or does the client wish to continue using the existing TCP connections? The client SHOULD use the "connection" parameter (defined in Section 18.54) on the Transport line to make its intention clear (by setting "connection" to "new" if new connections are needed, and by setting "connection" to "existing" if the existing connections are to be used). After a 2xx reply for a SETUP request for a new connection, parties should close the pre-existing connections, after waiting a suitable period for any stray RTP or RTCP packets to arrive.

The usage of SRTP, i.e., either SAVP or SAVPF profiles, requires that a security association is established. The default mechanism for establishing that security association is to use MIKEY[RFC3830] with RTSP as defined Appendix C.1.4.1.

Below, a rewritten version of the example "media on demand" (Appendix A.1) shows the use of RTP/AVP/TCP non-interleaved:

```
C->M: DESCRIBE rtsp://example.com/twister.3gp RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2

M->C: RTSP/2.0 200 OK
      CSeq: 1
      Server: PhonyServer/1.0
      Date: Wed, 23 Jan 2013 15:36:52 +0000
      Content-Type: application/sdp
      Content-Length: 227
      Content-Base: rtsp://example.com/twister.3gp/
      Expires: Thu, 24 Jan 2013 15:36:52 +0000

      v=0
      o=- 2890844256 2890842807 IN IP4 198.51.100.34
      s=RTSP Session
      i=An Example of RTSP Session Usage
      e=adm@example.com
      c=IN IP4 0.0.0.0
      a=control: *
      a=range:npt=00:00:00-00:10:34.10
      t=0 0
      m=audio 0 RTP/AVP 0
      a=control: trackID=1

C->M: SETUP rtsp://example.com/twister.3gp/trackID=1 RTSP/2.0
      CSeq: 2
      User-Agent: PhonyClient/1.2
      Require: play.basic
      Transport: RTP/AVP/TCP;unicast;dest_addr=":9"/":9";
                setup=active;connection=new
      Accept-Ranges: npt, smpte, clock
```

```
M->C: RTSP/2.0 200 OK
      CSeq: 2
      Server: PhonyServer/1.0
      Transport: RTP/AVP/TCP;unicast;
                dest_addr=":9"/":9";
                src_addr="198.51.100.5:53478"/"198.51.100:54091";
                setup=passive;connection=new;ssrc=93CB001E
      Session: 12345678
      Expires: Thu, 24 Jan 2013 15:36:52 +0000
      Date: Wed, 23 Jan 2013 15:36:52 +0000
      Accept-Ranges: npt
      Media-Properties: Random-Access=0.8, Immutable, Unlimited

C->M: TCP Connection Establishment x2

C->M: PLAY rtsp://example.com/twister.3gp/ RTSP/2.0
      CSeq: 4
      User-Agent: PhonyClient/1.2
      Range: npt=30-
      Session: 12345678

M->C: RTSP/2.0 200 OK
      CSeq: 4
      Server: PhonyServer/1.0
      Date: Wed, 23 Jan 2013 15:36:54 +0000
      Session: 12345678
      Range: npt=30-623.10
      Seek-Style: First-Prior
      RTP-Info: url="rtsp://example.com/twister.3gp/trackID=1"
                ssrc=4F312DD8:seq=54321;rtptime=2876889
```

C.3. Handling Media Clock Time Jumps in the RTP Media Layer

RTSP allows media clients to control selected, non-contiguous sections of media presentations, rendering those streams with an RTP media layer [RFC3550]. Two cases occur, the first is when a new PLAY request replaces an old ongoing request and the new request results in a jump in the media. This should produce in the RTP layer a continuous media stream. A client may also directly following a completed PLAY request perform a new PLAY request. This will result in some gap in the media layer. The below text will look into both cases.

A PLAY request that replaces an ongoing request allows the media layer rendering the RTP stream without being affected by jumps in media clock time. The RTP timestamps for the new media range is set so that they become continuous with the previous media range in the previous request. The RTP sequence number for the first packet in

the new range will be the next following the last packet in the previous range, i.e., monotonically increasing. The goal is to allow the media rendering layer to work without interruption or reconfiguration across the jumps in media clock. This should be possible in all cases of replaced PLAY requests for media that has random-access properties. In this case care is needed to align frames or similar media dependent structures.

In cases where jumps in media clock time are a result of RTSP signaling operations arriving after a completed PLAY operation, the request timing will result in that media becomes non-continuous. The server becomes unable to send the media so that it arrives timely and still carry timestamps to make the media stream continuous. In these cases the server will produce RTP streams where there are gaps in the RTP timeline for the media. In such cases, if the media has frame structure, aligning the timestamp for the next frame with the previous structure reduces the burden to render this media. The gap should represent the time the server hasn't been serving media, e.g., the time between the end of the media stream or a PAUSE request and the new PLAY request. In these cases the RTP sequence number would normally be monotonically increasing across the gap.

For RTSP sessions with media that lacks random access properties, such as live streams, any media clock jump is commonly the result of a correspondingly long pause of delivery. The RTP timestamp will have increased in direct proportion to the duration of the paused delivery. Note also that in this case the RTP sequence number should be the next packet number. If not, the RTCP packet loss reporting will indicate as loss all packets not received between the point of pausing and later resuming. This may trigger congestion avoidance mechanisms. An allowed exception from the above recommendation on monotonically increasing RTP sequence number is live media streams, likely being relayed. In this case, when the client resumes delivery, it will get the media that is currently being delivered to the server itself. For this type of basic delivery of live streams to multiple users over unicast, individual rewriting of RTP sequence numbers becomes quite a burden. For solutions that anyway caches media, timeshifts, etc, the rewriting should be a minor issue.

The goal when handling jumps in media clock time is that the provided stream is continuous without gaps in RTP timestamp or sequence number. However, when delivery has been halted for some reason the RTP timestamp when resuming MUST represent the duration the delivery was halted. RTP sequence number MUST generally be the next number, i.e., monotonically increasing modulo 65536. For media resources with the properties Time-Progressing and Time-Duration=0.0 the server MAY create RTP media streams with RTP sequence number jumps in them due to the client first halting delivery and later resuming it (PAUSE

and then later PLAY). However, servers utilizing this exception must take into consideration the resulting RTCP receiver reports that likely contain loss reports for all the packets part of the discontinuity. A client cannot rely on that a server will align when resuming playing even if it is RECOMMENDED. The RTP-Info header will provide information on how the server acts in each case.

One cannot assume that the RTSP client can communicate with the RTP media agent, as the two may be independent processes. If the RTP timestamp shows the same gap as the NPT, the media agent will assume that there is a pause in the presentation. If the jump in NPT is large enough, the RTP timestamp may roll over and the media agent may believe later packets to be duplicates of packets just played out. Having the RTP timestamp jump will also affect the RTCP measurements based on this.

As an example, assume an RTP timestamp frequency of 8000 Hz, a packetization interval of 100 ms and an initial sequence number and timestamp of zero.

```
C->S: PLAY rtsp://example.com/fizzle RTSP/2.0
      CSeq: 4
      Session: abcdefgh
      Range: npt=10-15
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 4
      Session: abcdefgh
      Range: npt=10-15
      RTP-Info: url="rtsp://example.com/fizzle/audiotrack"
                  ssrc=0D12F123:seq=0;rtptime=0
```

The ensuing RTP data stream is depicted below:

```
S -> C: RTP packet - seq = 0,  rtptime = 0,      NPT time = 10s
S -> C: RTP packet - seq = 1,  rtptime = 800,    NPT time = 10.1s
. . .
S -> C: RTP packet - seq = 49, rtptime = 39200, NPT time = 14.9s
```

Upon the completion of the requested delivery the server sends a
PLAY_NOTIFY

```
S->C: PLAY_NOTIFY rtsp://example.com/fizzle RTSP/2.0
      CSeq: 5
      Notify-Reason: end-of-stream
      Request-Status: cseq=4 status=200 reason="OK"
      Range: npt=-15
      RTP-Info:url="rtsp://example.com/fizzle/audiotrack"
              ssrc=0D12F123:seq=49;rtptime=39200
      Session: abcdefgh
```

```
C->S: RTSP/2.0 200 OK
      CSeq: 5
      User-Agent: PhonyClient/1.2
```

Upon the completion of the play range, the client follows up with a request to PLAY from a new NPT.

```
C->S: PLAY rtsp://example.com/fizzle RTSP/2.0
      CSeq: 6
      Session: abcdefg
      Range: npt=18-20
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 6
      Session: abcdefg
      Range: npt=18-20
      RTP-Info: url="rtsp://example.com/fizzle/audiotrack"
              ssrc=0D12F123:seq=50;rtptime=40100
```

The ensuing RTP data stream is depicted below:

```
S->C: RTP packet - seq = 50, rtptime = 40100, NPT time = 18s
S->C: RTP packet - seq = 51, rtptime = 40900, NPT time = 18.1s
. . .
S->C: RTP packet - seq = 69, rtptime = 55300, NPT time = 19.9s
```

In this example, first, NPT 10 through 15 is played, then the client requests the server to skip ahead and play NPT 18 through 20. The first segment is presented as RTP packets with sequence numbers 0 through 49 and timestamp 0 through 39,200. The second segment consists of RTP packets with sequence number 50 through 69, with timestamps 40,100 through 55,200. While there is a gap in the NPT, there is no gap in the sequence number space of the RTP data stream.

The RTP timestamp gap is present in the above example due to the time it takes to perform the second play request, in this case 12.5 ms (100/8000).

C.4. Handling RTP Timestamps after PAUSE

During a PAUSE / PLAY interaction in an RTSP session, the duration of time for which the RTP transmission was halted MUST be reflected in the RTP timestamp of each RTP stream. The duration can be calculated for each RTP stream as the time elapsed from when the last RTP packet was sent before the PAUSE request was received and when the first RTP packet was sent after the subsequent PLAY request was received. The duration includes all latency incurred and processing time required to complete the request.

The RTP RFC [RFC3550] states that: The RTP timestamp for each unit [packet] would be related to the wallclock time at which the unit becomes current on the virtual presentation timeline.

In order to satisfy the requirements of [RFC3550], the RTP timestamp space needs to increase continuously with real time. While this is not optimal for stored media, it is required for RTP and RTCP to function as intended. Using a continuous RTP timestamp space allows the same timestamp model for both stored and live media and allows better opportunity to integrate both types of media under a single control.

As an example, assume a clock frequency of 8000 Hz, a packetization interval of 100 ms and an initial sequence number and timestamp of zero.

```
C->S: PLAY rtsp://example.com/fizzle RTSP/2.0
      CSeq: 4
      Session: abcdefg
      Range: npt=10-15

      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 4
      Session: abcdefg
      Range: npt=10-15
      RTP-Info: url="rtsp://example.com/fizzle/audiotrack"
                ssrc=0D12F123:seq=0;rtptime=0
```

The ensuing RTP data stream is depicted below:

```
S -> C: RTP packet - seq = 0, rtptime = 0,    NPT time = 10s
S -> C: RTP packet - seq = 1, rtptime = 800,   NPT time = 10.1s
S -> C: RTP packet - seq = 2, rtptime = 1600,   NPT time = 10.2s
S -> C: RTP packet - seq = 3, rtptime = 2400,   NPT time = 10.3s
```

The client then sends a PAUSE request:

```
C->S: PAUSE rtsp://example.com/fizzle RTSP/2.0
      CSeq: 5
      Session: abcdefg
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 5
      Session: abcdefg
      Range: npt=10.4-15
```

20 seconds elapse and then the client sends a PLAY request. In addition the server requires 15 ms to process the request:

```
C->S: PLAY rtsp://example.com/fizzle RTSP/2.0
      CSeq: 6
      Session: abcdefg
      User-Agent: PhonyClient/1.2
```

```
S->C: RTSP/2.0 200 OK
      CSeq: 6
      Session: abcdefg
      Range: npt=10.4-15
      RTP-Info: url="rtsp://example.com/fizzle/audiotrack"
                 ssrc=0D12F123:seq=4;rtptime=164400
```

The ensuing RTP data stream is depicted below:

```
S -> C: RTP packet - seq = 4, rtptime = 164400, NPT time = 10.4s
S -> C: RTP packet - seq = 5, rtptime = 165200, NPT time = 10.5s
S -> C: RTP packet - seq = 6, rtptime = 166000, NPT time = 10.6s
```

First, NPT 10 through 10.3 is played, then a PAUSE is received by the server. After 20 seconds a PLAY is received by the server which takes 15 ms to process. The duration of time for which the session was paused is reflected in the RTP timestamp of the RTP packets sent after this PLAY request.

A client can use the RTSP range header and RTP-Info header to map NPT time of a presentation with the RTP timestamp.

Note: In RFC 2326 [RFC2326], this matter was not clearly defined and was misunderstood commonly. However, for RTSP 2.0 it is expected that this will be handled correctly and no exception handling will be required.

Note further: It may be required to reset some of the state to ensure the correct media decoding and the usual jitter-buffer handling when issuing a PLAY request.

C.5. RTSP / RTP Integration

For certain data types, tight integration between the RTSP layer and the RTP layer will be necessary. This by no means precludes the above restrictions. Combined RTSP/RTP media clients should use the RTP-Info field to determine whether incoming RTP packets were sent before or after a seek or before or after a PAUSE.

C.6. Scaling with RTP

For scaling (see Section 18.46), RTP timestamps should correspond to the rendering timing. For example, when playing video recorded at 30 frames/second at a scale of two and speed (Section 18.50) of one, the server would drop every second frame to maintain and deliver video packets with the normal timestamp spacing of 3,000 per frame, but NPT would increase by 1/15 second for each video frame.

Note: The above scaling puts requirements on the media codec or a media stream to support it. For example motion JPEG or other non-predictive video coding can easier handle the above example.

C.7. Maintaining NPT synchronization with RTP timestamps

The client can maintain a correct display of NPT (Normal Play Time) by noting the RTP timestamp value of the first packet arriving after repositioning. The sequence parameter of the RTP-Info (Section 18.45) header provides the first sequence number of the next segment.

C.8. Continuous Audio

For continuous audio, the server SHOULD set the RTP marker bit at the beginning of serving a new PLAY request or at jumps in timeline. This allows the client to perform playout delay adaptation.

C.9. Multiple Sources in an RTP Session

Note that more than one SSRC MAY be sent in the media stream. If it happens all sources are expected to be rendered simultaneously.

C.10. Usage of SSRCs and the RTCP BYE Message During an RTSP Session

The RTCP BYE message indicates the end of use of a given SSRC. If all sources leave an RTP session, it can, in most cases, be assumed to have ended. Therefore, a client or server **MUST NOT** send an RTCP BYE message until it has finished using a SSRC. A server **SHOULD** keep using a SSRC until the RTP session is terminated. Prolonging the use of a SSRC allows the established synchronization context associated with that SSRC to be used to synchronize subsequent PLAY requests even if the PLAY response is late.

An SSRC collision with the SSRC that transmits media does also have consequences, as it will normally force the media sender to change its SSRC in accordance with the RTP specification [RFC3550]. However, an RTSP server may wait and see if the client changes and thus resolve the conflict to minimize the impact. As media sender SSRC change will result in a loss of synchronization context, and require any receiver to wait for RTCP sender reports for all media requiring synchronization before being able to play out synchronized. Due to these reasons a client joining a session should take care to not select the same SSRC(s) as the server indicates in the `ssrc` Transport header parameter. Any SSRC signalled in the Transport header **MUST** be avoided. A client detecting a collision prior to sending any RTP or RTCP messages **SHALL** also select a new SSRC.

C.11. Future Additions

It is the intention that any future protocol or profile regarding media delivery and lower transport should be easy to add to RTSP. This section provides the necessary steps that needs to be meet.

The following things needs to be considered when adding a new protocol or profile for use with RTSP:

- o The protocol or profile needs to define a name tag representing it. This tag is required to be an ABNF "token" to be possible to use in the Transport header specification.
- o The useful combinations of protocol, profiles and lower layer transport for this extension needs to be defined. For each combination declare the necessary parameters to use in the Transport header.
- o For new media protocols the interaction with RTSP needs to be addressed. One important factor will be the media synchronization. It may be necessary to have new headers similar to RTP info to carry this information.

- o Discuss congestion control for media, especially if transport without built in congestion control is used.

See the IANA section (Section 22) for information how to register new attributes.

Appendix D. Use of SDP for RTSP Session Descriptions

The Session Description Protocol (SDP, [RFC4566]) may be used to describe streams or presentations in RTSP. This description is typically returned in reply to a DESCRIBE request on a URI from a server to a client, or received via HTTP from a server to a client.

This appendix describes how an SDP file determines the operation of an RTSP session. Thus, it is worth pointing out that the interpretation of the SDP is done in the context of the SDP receiver, which is the one being configured. This is the same as in SAP [RFC2974]; this differs from SDP Offer/Answer [RFC3264] where each SDP is interpreted in the context of the agent providing it.

SDP as is provides no mechanism by which a client can distinguish, without human guidance, between several media streams to be rendered simultaneously and a set of alternatives (e.g., two audio streams spoken in different languages). The SDP extension "Grouping of Media Lines in the Session Description Protocol (SDP)" [RFC5888] provides such functionality to some degree. Appendix D.4 describes the usage of SDP media line grouping for RTSP.

D.1. Definitions

The terms "session-level", "media-level" and other key/attribute names and values used in this appendix are to be used as defined in SDP[RFC4566]:

D.1.1. Control URI

The "a=control:" attribute is used to convey the control URI. This attribute is used both for the session and media descriptions. If used for individual media, it indicates the URI to be used for controlling that particular media stream. If found at the session level, the attribute indicates the URI for aggregate control (presentation URI). The session level URI MUST be different from any media level URI. The presence of a session level control attribute MUST be interpreted as support for aggregated control. The control attribute MUST be present on media level unless the presentation only contains a single media stream, in which case the attribute MAY be present on the session level only and then also apply to that single media stream.

ABNF for the attribute is defined in Section 20.3.

Example:

```
a=control:rtsp://example.com/foo
```

This attribute MAY contain either relative or absolute URIs, following the rules and conventions set out in RFC 3986 [RFC3986]. Implementations MUST look for a base URI in the following order:

1. the RTSP Content-Base field;
2. the RTSP Content-Location field;
3. the RTSP Request-URI.

If this attribute contains only an asterisk (*), then the URI MUST be treated as if it were an empty embedded URI, and thus inherit the entire base URI.

Note, RFC 2326 was very unclear on the processing of relative URI and several RTSP 1.0 implementations at the point of publishing this document did not perform RFC 3986 processing to determine the resulting URI, instead simple concatenation is common. To avoid this issue completely it is recommended to use absolute URI in the SDP.

The URI handling for SDPs from container files need special consideration. For example let's assume that a container file has the URI: "rtsp://example.com/container.mp4". Let's further assume this URI is the base URI, and that there is an absolute media level URI: "rtsp://example.com/container.mp4/trackID=2". A relative media level URI that resolves in accordance with RFC 3986 [RFC3986] to the above given media URI is: "container.mp4/trackID=2". It is usually not desirable to need to include in or modify the SDP stored within the container file with the server local name of the container file. To avoid this, one can modify the base URI used to include a trailing slash, e.g., "rtsp://example.com/container.mp4/". In this case the relative URI for the media will only need to be: "trackID=2". However, this will also mean that using "*" in the SDP will result in control URI including the trailing slash, i.e., "rtsp://example.com/container.mp4/".

Note: The usage of TrackID in the above is not a standardized form, but one example out of several similar strings such as TrackID, Track_ID, StreamID that is used by different server vendors to indicate a particular piece of media inside a container file.

D.1.2. Media Streams

The "m=" field is used to enumerate the streams. It is expected that all the specified streams will be rendered with appropriate synchronization. If the session is over multicast, the port number indicated SHOULD be used for reception. The client MAY try to override the destination port, through the Transport header. The servers MAY allow this, the response will indicate if allowed or not. If the session is unicast, the port numbers are the ones RECOMMENDED by the server to the client, about which receiver ports to use; the client MUST still include its receiver ports in its SETUP request. The client MAY ignore this recommendation. If the server has no preference, it SHOULD set the port number value to zero.

The "m=" lines contain information about which transport protocol, profile, and possibly lower-layer is to be used for the media stream. The combination of transport, profile and lower layer, like RTP/AVP/UDP needs to be defined for how to be used with RTSP. The currently defined combinations are defined in Appendix C, further combinations MAY be specified.

Example:

```
m=audio 0 RTP/AVP 31
```

D.1.3. Payload Type(s)

The payload type(s) are specified in the "m=" line. In case the payload type is a static payload type from RFC 3551 [RFC3551], no other information may be required. In case it is a dynamic payload type, the media attribute "rtpmap" is used to specify what the media is. The "encoding name" within the "rtpmap" attribute may be one of those specified in [RFC4856], or a media type registered with IANA according to [RFC4855], or an experimental encoding as specified in SDP [RFC4566]). Codec-specific parameters are not specified in this field, but rather in the "fmtp" attribute described below.

The selection of the RTP payload type numbers used may be required to consider RTP and RTCP Multiplexing [RFC5761] if that is to be supported by the server.

D.1.4. Format-Specific Parameters

Format-specific parameters are conveyed using the "fmtp" media attribute. The syntax of the "fmtp" attribute is specific to the encoding(s) that the attribute refers to. Note that some of the format specific parameters may be specified outside of the fmtp

parameters, like for example the "ptime" attribute for most audio encodings.

D.1.5. Directionality of media stream

The SDP attributes "a=sendrecv", "a=recvonly" and "a=sendonly" provide instructions about the direction the media streams flow within a session. When using RTSP the SDP can be delivered to a client using either RTSP DESCRIBE or a number of RTSP external methods, like HTTP, FTP, and email. Based on this the SDP applies to how the RTSP client will see the complete session. Thus media streams delivered from the RTSP server to the client, would be given the "a=recvonly" attribute.

"a=recvonly" in a SDP provided to the RTSP client indicates that media delivery will only occur in the direction from the RTSP server to the client. SDP provided to the RTSP client that lacks any of the directionality attributes (a=recvonly, a=sendonly, a=sendrecv) would be interpreted as having a=sendrecv. At the time of writing there exist no RTSP mode suitable for media traffic in the direction from the RTSP client to the server. Thus all RTSP SDP SHOULD have a=recvonly attribute when using the PLAY mode defined in this document. If future modes are defined for media in client to server direction, then usage of a=sendonly, or a=sendrecv may become suitable to indicate intended media directions.

D.1.6. Range of Presentation

The "a=range" attribute defines the total time range of the stored session or an individual media. Non-seekable live sessions can be indicated as specified below, while the length of live sessions can be deduced from the "t=" and "r=" SDP parameters.

The attribute is both a session and a media level attribute. For presentations that contain media streams of the same duration, the range attribute SHOULD only be used at session-level. In case of different lengths the range attribute MUST be given at media level for all media, and SHOULD NOT be given at session level. If the attribute is present at both media level and session level the media level values MUST be used.

Note: Usually one will specify the same length for all media, even if there isn't media available for the full duration on all media. However, that requires that the server accepts PLAY requests within that range.

Servers MUST take care to provide RTSP Range (see Section 18.40) values that are consistent with what is presented in the SDP for the

content. There is no reason for non dynamic content, like media clips provided on demand to have inconsistent values. Inconsistent values between the SDP and the actual values for the content handled by the server is likely to generate some failure, like 457 "Invalid Range", in case the client uses PLAY requests with a Range header. In case the content is dynamic in length and it is infeasible to provide a correct value in the SDP the server is recommended to describe this as non-seekable content (see below). The server MAY override that property in the response to a PLAY request using the correct values in the Range header.

The unit is specified first, followed by the value range. The units and their values are as defined in Section 4.4.1, Section 4.4.2 and Section 4.4.3 and MAY be extended with further formats. Any open ended range (start-), i.e., without stop range, is of unspecified duration and MUST be considered as non-seekable content unless this property is overridden. Multiple instances carrying different clock formats MAY be included at either session or media level.

ABNF for the attribute is defined in Section 20.3.

Examples:

```
a=range:npt=0-34.4368
a=range:clock=19971113T211503Z-19971113T220300Z
Non seekable stream of unknown duration:
a=range:npt=0-
```

D.1.7. Time of Availability

The "t=" field defines when the SDP is valid. For on-demand content the server SHOULD indicate a stop time value for which it guarantees the description to be valid, and a start time that is equal to or before the time at which the DESCRIBE request was received. It MAY also indicate start and stop times of 0, meaning that the session is always available.

For sessions that are of live type, i.e., specific start time, unknown stop time, likely unseekable, the "t=" and "r=" field SHOULD be used to indicate the start time of the event. The stop time SHOULD be given so that the live event will have ended at that time, while still not be unnecessary long into the future.

D.1.8. Connection Information

In SDP used with RTSP, the "c=" field contains the destination address for the media stream. If a multicast address is specified the client SHOULD use this address in any SETUP request as

destination address, including any additional parameters, such as TTL. For on-demand unicast streams and some multicast streams, the destination address MAY be specified by the client via the SETUP request, thus overriding any specified address. To identify streams without a fixed destination address, where the client is required to specify a destination address, the "c=" field SHOULD be set to a null value. For addresses of type "IP4", this value MUST be "0.0.0.0", and for type "IP6", this value MUST be "0:0:0:0:0:0:0:0" (can also be written as "::"), i.e., the unspecified address according to RFC 4291 [RFC4291].

D.1.9. Message Body Tag

The optional "a=mtag" attribute identifies a version of the session description. It is opaque to the client. SETUP requests may include this identifier in the If-Match field (see Section 18.24) to only allow session establishment if this attribute value still corresponds to that of the current description. The attribute value is opaque and may contain any character allowed within SDP attribute values.

ABNF for the attribute is defined in Section 20.3.

Example:

```
a=mtag:"158bb3e7c7fd62ce67f12b533f06b83a"
```

One could argue that the "o=" field provides identical functionality. However, it does so in a manner that would put constraints on servers that need to support multiple session description types other than SDP for the same piece of media content.

D.2. Aggregate Control Not Available

If a presentation does not support aggregate control no session level "a=control:" attribute is specified. For a SDP with multiple media sections specified, each section will have its own control URI specified via the "a=control:" attribute.

Example:


```
v=0
o=- 2890844256 2890842807 IN IP4 192.0.2.56
s=I came from a web page
e=adm@example.com
c=IN IP4 0.0.0.0
t=0 0
m=video 8002 RTP/AVP 31
a=control:rtsp://audio.example.com/movie.aud
m=audio 8004 RTP/AVP 3
a=control:rtsp://video.example.com/movie.vid
```

Note that the position of the control URI in the description implies that the client establishes separate RTSP control sessions to the servers audio.example.com and video.example.com.

It is recommended that an SDP file contains the complete media initialization information even if it is delivered to the media client through non-RTSP means. This is necessary as there is no mechanism to indicate that the client should request more detailed media stream information via DESCRIBE.

D.3. Aggregate Control Available

In this scenario, the server has multiple streams that can be controlled as a whole. In this case, there are both a media-level "a=control:" attributes, which are used to specify the stream URIs, and a session-level "a=control:" attribute which is used as the Request-URI for aggregate control. If the media-level URI is relative, it is resolved to absolute URIs according to Appendix D.1.1 above.

Example:

```
C->M: DESCRIBE rtsp://example.com/movie RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2
```

```
M->C: RTSP/2.0 200 OK
      CSeq: 1
      Date: Wed, 23 Jan 2013 15:36:52 +0000
      Expires: Wed, 23 Jan 2013 16:36:52 +0000
      Content-Type: application/sdp
      Content-Base: rtsp://example.com/movie/
      Content-Length: 227
```

```
v=0
o=- 2890844256 2890842807 IN IP4 192.0.2.211
s=I contain
i=<more info>
e=adm@example.com
c=IN IP4 0.0.0.0
a=control:*
t=0 0
m=video 8002 RTP/AVP 31
a=control:trackID=1
m=audio 8004 RTP/AVP 3
a=control:trackID=2
```

In this example, the client is recommended to establish a single RTSP session to the server, and uses the URIs `rtsp://example.com/movie/trackID=1` and `rtsp://example.com/movie/trackID=2` to set up the video and audio streams, respectively. The URI `rtsp://example.com/movie/`, which is resolved from the "*", controls the whole presentation (movie).

A client is not required to issue SETUP requests for all streams within an aggregate object. Servers should allow the client to ask for only a subset of the streams.

D.4. Grouping of Media Lines in SDP

For some types of media it is desirable to express a relationship between various media components, for instance, for lip synchronization or Scalable Video Codec (SVC) [RFC5583]. This relationship is expressed on the SDP level by grouping of media lines, as described in [RFC5888] and can be exposed to RTSP.

For RTSP it is mainly important to know how to handle grouped medias received by means of SDP, i.e., if the media are under aggregate control (see Appendix D.3) or if aggregate control is not available (see Appendix D.2).

It is RECOMMENDED that grouped medias are handled by aggregate control, to give the client the ability to control either the whole presentation or single medias.

D.5. RTSP external SDP delivery

There are some considerations that need to be made when the session description is delivered to the client outside of RTSP, for example via HTTP or email.

First of all, the SDP needs to contain absolute URIs, since relative will in most cases not work as the delivery will not correctly forward the base URI.

The writing of the SDP session availability information, i.e., "t=" and "r=", needs to be carefully considered. When the SDP is fetched by the DESCRIBE method, the probability that it is valid is very high. However, the same is much less certain for SDPs distributed using other methods. Therefore the publisher of the SDP should take care to follow the recommendations about availability in the SDP specification [RFC4566] in Section 4.2.

Appendix E. RTSP Use Cases

This Appendix describes the most important and considered use cases for RTSP. They are listed in descending order of importance in regards to ensuring that all necessary functionality is present. This specification only fully supports usage of the two first. Also in these first two cases, there are special cases or exceptions that are not supported without extensions, e.g., the redirection of media delivery to another address than the controlling agent's (client's).

E.1. On-demand Playback of Stored Content

An RTSP capable server stores content suitable for being streamed to a client. A client desiring playback of any of the stored content uses RTSP to set up the media transport required to deliver the desired content. RTSP is then used to initiate, halt and manipulate the actual transmission (playout) of the content. RTSP is also required to provide necessary description and synchronization information for the content.

The above high level description can be broken down into a number of functions that RTSP needs to be capable of.

Presentation Description: Provide initialization information about the presentation (content); for example, which media codecs are needed for the content. Other information that is important

includes the number of media streams the presentation contains, the transport protocols used for the media streams, and identifiers for these media streams. This information is required before setup of the content is possible and to determine if the client is even capable of using the content.

This information need not be sent using RTSP; other external protocols can be used to transmit the transport presentation descriptions. Two good examples are the use of HTTP [RFC2616] or email to fetch or receive presentation descriptions like SDP [RFC4566]

Setup: Set up some or all of the media streams in a presentation. The setup itself consists of selecting the protocol for media transport and the necessary parameters for the protocol, like addresses and ports.

Control of Transmission: After the necessary media streams have been established the client can request the server to start transmitting the content. The client must be allowed to start or stop the transmission of the content at arbitrary times. The client must also be able to start the transmission at any point in the timeline of the presentation.

Synchronization: For media transport protocols like RTP [RFC3550] it might be beneficial to carry synchronization information within RTSP. This may be due to either the lack of inter-media synchronization within the protocol itself, or the potential delay before the synchronization is established (which is the case for RTP when using RTCP).

Termination: Terminate the established contexts.

For this use case there are a number of assumptions about how it works. These are:

On-Demand content: The content is stored at the server and can be accessed at any time during a time period when it is intended to be available.

Independent sessions: A server is capable of serving a number of clients simultaneously, including from the same piece of content at different points in that presentations time-line.

Unicast Transport: Content for each individual client is transmitted to them using unicast traffic.

It is also possible to redirect the media traffic to a different destination than that of the agent controlling the traffic. However, allowing this without appropriate mechanisms for checking that the destination approves of this allows for distributed denial of service attacks (DDoS).

E.2. Unicast Distribution of Live Content

This use case is similar to the above on-demand content case (see Appendix E.1) the difference is the nature of the content itself. Live content is continuously distributed as it becomes available from a source; i.e., the main difference from on-demand is that one starts distributing content before the end of it has become available to the server.

In many cases the consumer of live content is only interested in consuming what actually happens "now"; i.e., very similar to broadcast TV. However, in this case it is assumed that there exists no broadcast or multicast channel to the users, and instead the server functions as a distribution node, sending the same content to multiple receivers, using unicast traffic between server and client. This unicast traffic and the transport parameters are individually negotiated for each receiving client.

Another aspect of live content is that it often has a very limited time of availability, as it is only available for the duration of the event the content covers. An example of such a live content could be a music concert which lasts 2 hour and starts at a predetermined time. Thus there is a need to announce when and for how long the live content is available.

In some cases, the server providing live content may be saving some or all of the content to allow clients to pause the stream and resume it from the paused point, or to "rewind" and play continuously from a point earlier than the live point. Hence, this use case does not necessarily exclude playing from other than the live point of the stream, playing with scales other than 1.0, etc.

E.3. On-demand Playback using Multicast

It is possible to use RTSP to request that media be delivered to a multicast group. The entity setting up the session (the controller) will then control when and what media is delivered to the group. This use case has some potential for denial of service attacks by flooding a multicast group. Therefore, a mechanism is needed to indicate that the group actually accepts the traffic from the RTSP server.

An open issue in this use case is how one ensures that all receivers listening to the multicast or broadcast receives the session presentation configuring the receivers. This specification has to rely on an external solution to solve this issue.

E.4. Inviting an RTSP server into a conference

If one has an established conference or group session, it is possible to have an RTSP server distribute media to the whole group. Transmission to the group is simplest when controlled by a single participant or leader of the conference. Shared control might be possible, but would require further investigation and possibly extensions.

This use case assumes that there exists either multicast or a conference focus that redistribute media to all participants.

This use case is intended to be able to handle the following scenario: A conference leader or participant (hereafter called the controller) has some pre-stored content on an RTSP server that he wants to share with the group. The controller sets up an RTSP session at the streaming server for this content and retrieves the session description for the content. The destination for the media content is set to the shared multicast group or conference focus. When desired by the controller, he/she can start and stop the transmission of the media to the conference group.

There are several issues with this use case that are not solved by this core specification for RTSP:

Denial of service: To avoid an RTSP server from being an unknowing participant in a denial of service attack the server needs to be able to verify the destination's acceptance of the media. Such a mechanism to verify the approval of received media does not yet exist; instead, only policies can be used, which can be made to work in controlled environments.

Distributing the presentation description to all participants in the group:

To enable a media receiver to correctly decode the content the media configuration information needs to be distributed reliably to all participants. This will most likely require support from an external protocol.

Passing control of the session: If it is desired to pass control of the RTSP session between the participants, some support will be required by an external protocol to exchange state

information and possibly floor control of who is controlling the RTSP session.

E.5. Live Content using Multicast

This use case in its simplest form does not require any use of RTSP at all; this is what multicast conferences being announced with SAP [RFC2974] and SDP are intended to handle. However, in use cases where more advanced features like access control to the multicast session are desired, RTSP could be used for session establishment.

A client desiring to join a live multicasted media session with cryptographic (encryption) access control could use RTSP in the following way. The source of the session announces the session and gives all interested an RTSP URI. The client connects to the server and requests the presentation description, allowing configuration for reception of the media. In this step it is possible for the client to use secured transport and any desired level of authentication; for example, for billing or access control. An RTSP link also allows for load balancing between multiple servers.

If these were the only goals, they could be achieved by simply using HTTP. However, for cases where the sender likes to keep track of each individual receiver of a session, and possibly use the session as a side channel for distributing key-updates or other information on a per-receiver basis, and the full set of receivers is not known prior to the session start, the state establishment that RTSP provides can be beneficial. In this case a client would establish an RTSP session for this multicast group with the RTSP server. The RTSP server will not transmit any media, but instead will point to the multicast group. The client and server will be able to keep the session alive for as long as the receiver participates in the session thus enabling, for example, the server to push updates to the client.

This use case will most likely not be able to be implemented without some extensions to the server-to-client push mechanism. Here the PLAY_NOTIFY method (see Section 13.5) with a suitable extension could provide clear benefits.

Appendix F. Text format for Parameters

A resource of type "text/parameters" consists of either 1) a list of parameters (for a query) or 2) a list of parameters and associated values (for an response or setting of the parameter). Each entry of the list is a single line of text. Parameters are separated from values by a colon. The parameter name MUST only use US-ASCII visible characters while the values are UTF-8 text strings. The media type registration form is in Section 22.16.

There is a potential interoperability issue for this format. It was named in RFC 2326 but never defined, even if used in examples that hint at the syntax. This format matches the purpose and its syntax supports the examples provided. However, it goes further by allowing UTF-8 in the value part, thus usage of UTF-8 strings may not be supported. However, as individual parameters are not defined, the using application anyway needs to have out-of-band agreement or using feature-tag to determine if the end-point supports the parameters.

The ABNF [RFC5234] grammar for "text/parameters" content is:

```

file           = *((parameter / parameter-value) CRLF)
parameter      = 1*visible-except-colon
parameter-value = parameter *WSP ":" value
visible-except-colon = %x21-39 / %x3B-7E      ; VCHAR - ":"
value           = *(TEXT-UTF8char / WSP)
TEXT-UTF8char   = <as defined in Section 20.1>
WSP             = <See RFC 5234> ; Space or HTAB
VCHAR          = <See RFC 5234>
CRLF           = <See RFC 5234>

```

Appendix G. Requirements for Unreliable Transport of RTSP

This appendix provides guidance for those who want to implement RTSP messages over unreliable transports as has been defined in RTSP 1.0 [RFC2326]. RFC 2326 defined the "rtspu" URI scheme and provided some basic information for the transport of RTSP messages over UDP. The information is being provided here as there has been at least one commercial implementation and compatibility with that should be maintained.

The following points should be considered for an interoperable implementation:

- o Requests shall be acknowledged by the receiver. If there is no acknowledgement, the sender may resend the same message after a timeout of one round-trip time (RTT). Any retransmissions due to lack of acknowledgement must carry the same sequence number as the original request.
- o The round-trip time can be estimated as in TCP (RFC 6298) [RFC6298], with an initial round-trip value of 500 ms. An implementation may cache the last RTT measurement as the initial value for future connections.
- o The Timestamp header (Section 18.53) is used to avoid the retransmission ambiguity problem [Stevens98].

- o The registered default port for RTSP over UDP for the server is 554.
- o RTSP messages can be carried over any lower-layer transport protocol that is 8-bit clean.
- o RTSP messages are vulnerable to bit errors and should not be subjected to them.
- o Source authentication, or at least validation that RTSP messages comes from the same entity becomes extremely important, as session hijacking may be substantially easier for RTSP message transport using an unreliable protocol like UDP than for TCP.

There are two RTSP headers that are primarily intended for being used by the unreliable handling of RTSP messages and which will be maintained:

- o CSeq: See Section 18.20. It should be noted that the CSeq header is also required to match requests and responses independent whether a reliable or unreliable transport is used.
- o Timestamp: See Section 18.53

Appendix H. Backwards Compatibility Considerations

This section contains notes on issues about backwards compatibility with clients or servers being implemented according to RFC 2326 [RFC2326]. Note that there exists no requirement to implement RTSP 1.0; in fact this document recommend against it as it is difficult to do in an interoperable way.

A server implementing RTSP/2.0 MUST include an RTSP-Version of RTSP/2.0 in all responses to requests containing RTSP-Version RTSP/2.0. If a server receives an RTSP/1.0 request, it MAY respond with an RTSP/1.0 response if it chooses to support RFC 2326. If the server chooses not to support RFC 2326, it MUST respond with a 505 (RTSP Version not supported) status code. A server MUST NOT respond to an RTSP-Version RTSP/1.0 request with an RTSP-Version RTSP/2.0 response.

Clients implementing RTSP/2.0 MAY use an OPTIONS request with a RTSP-Version of 2.0 to determine whether a server supports RTSP/2.0. If the server responds with either an RTSP-Version of 1.0 or a status code of 505 (RTSP Version not supported), the client will have to use RTSP/1.0 requests if it chooses to support RFC 2326.

H.1. Play Request in Play State

The behavior in the server when a Play is received in Play state has changed (Section 13.4). In RFC 2326, the new PLAY request would be queued until the current Play completed. Any new PLAY request now takes effect immediately replacing the previous request.

H.2. Using Persistent Connections

Some server implementations of RFC 2326 maintain a one-to-one relationship between a connection and an RTSP session. Such implementations require clients to use a persistent connection to communicate with the server and when a client closes its connection, the server may remove the RTSP session. This is worth noting if a RTSP 2.0 client also supporting 1.0 connects to a 1.0 server.

Appendix I. Changes

This appendix briefly lists the differences between RTSP 1.0 [RFC2326] and RTSP 2.0 for an informational purpose. For implementers of RTSP 2.0 it is recommended to read carefully through this memo and not to rely on the list of changes below to adapt from RTSP 1.0 to RTSP 2.0, as RTSP 2.0 is not intended to be backwards compatible with RTSP 1.0 [RFC2326] other than the version negotiation mechanism.

I.1. Brief Overview

The following protocol elements were removed in RTSP 2.0 compared to RTSP 1.0:

- o there is no section on minimal implementation anymore, but more the definition of RTSP 2.0 core;
- o the RECORD and ANNOUNCE methods and all related functionality (including 201 (Created) and 250 (Low On Storage Space) status codes);
- o the use of UDP for RTSP message transport was removed due to missing interest and to broken specification;
- o the use of PLAY method for keep-alive in Play state.

The following protocol elements were added or changed in RTSP 2.0 compared to RTSP 1.0:

- o RTSP session TEARDOWN from the server to the client;

- o IPv6 support;
- o extended IANA registries (e.g., transport headers parameters, transport-protocol, profile, lower-transport, and mode);
- o request pipelining for quick session start-up;
- o fully reworked state-machine;
- o RTSP messages now use URIs rather than URLs;
- o incorporated much of related HTTP text ([RFC2616]) in this memo, compared to just referencing the sections in HTTP, to avoid ambiguities;
- o the REDIRECT method was expanded and diversified for different situations;
- o Includes a new section about how to setup different media transport alternatives and their profiles, and lower layer protocols. This caused the appendix on RTP interaction to be moved there instead of being in the part which describes RTP. The section also includes guidelines what to consider when writing usage guidelines for new protocols and profiles;
- o Added an asynchronous notification method `PLAY_NOTIFY`. This method is used by the RTSP server to asynchronously notify clients about session changes while in Play state. To a limited extent this is comparable with some implementations of `ANNOUNCE` in RTSP 1.0 not intended for Recording.

I.2. Detailed List of Changes

Compared to RTSP 1.0 (RFC 2326), the below changes has been made when defining RTSP 2.0. Note that this list does not reflect minor changes in wording or correction of typographical errors.

- o The section on minimal implementation was deleted without substitution.
- o The Transport header has been changed in the following way:
 - * The ABNF has been changed to define that extensions are possible, and that unknown parameters result in that servers ignore the transport specification.

- * To prevent backwards compatibility issues, any extension or new parameter requires the usage of a feature-tag combined with the Require header.
 - * Syntax unclarities with the Mode parameter have been resolved.
 - * Syntax error with ";" for multicast and unicast has been resolved.
 - * Two new addressing parameters have been defined, src_addr and dest_addr. These replace the parameters "port", "client_port", "server_port", "destination", "source".
 - * Support for IPv6 explicit addresses in all address fields has been included.
 - * To handle URI definitions that contain ";" or "," a quoted URI format has been introduced and is required.
 - * Defined IANA registries for the transport headers parameters, transport-protocol, profile, lower-transport, and mode.
 - * The transport headers interleaved parameter's text was made more strict and uses formal requirements levels. It was also clarified that the interleaved channels are symmetric and that it is the server that sets the channel numbers.
 - * It has been clarified that the client can't request of the server to use a certain RTP SSRC, using a request with the transport parameter SSRC.
 - * Syntax definition for SSRC has been clarified to require 8HEX. It has also been extended to allow multiple values for clients supporting this version.
 - * Clarified the text on the transport headers "dest_addr" parameters regarding what security precautions the server is required to perform.
- o The Range formats has been changed in the following way:
- * The NPT format has been given an initial NPT identifier that must now be used.
 - * All formats now support initial open ended formats of type "npt=-10" and also format only "Range: smpte" ranges for usage with GET_PARAMETER requests.

- * The npt-hhmmss notation now follows ISO 8601 more stricter.
- o RTSP message handling has been changed in the following way:
 - * RTSP messages now use URIs rather than URLs.
 - * It has been clarified that a 4xx message due to missing CSeq header shall be returned without a CSeq header.
 - * The 300 (Multiple Choices) response code has been removed.
 - * Rules for how to handle timing out RTSP messages has been added.
 - * Extended Pipelining rules allowing for quick session startup.
 - * Sequence numbering and proxy handling of sequence number defined, including case when response arrive out of order.
- o The HTTP references have been updated to RFC 2616 and RFC 2617. Most of the text has been copied and then altered to fit RTSP into this specification. Public, and the Content-Base header has also been imported from RFC 2068 so that they are defined in the RTSP specification. Known effects on RTSP due to HTTP clarifications:
 - * Content-Encoding header can include encoding of type "identity".
- o The state machine section has been completely rewritten. It now includes more details and is also more clear about the model used.
- o An IANA section has been included which contains a number of registries and their rules. This will allow us to use IANA to keep track of RTSP extensions.
- o The transport of RTSP messages has seen the following changes:
 - * The use of UDP for RTSP message transport has been deprecated due to missing interest and to broken specification.
 - * The rules for how TCP connections are to be handled has been clarified. Now it is made clear that servers should not close the TCP connection unless they have been unused for significant time.
 - * Strong recommendations why server and clients should use persistent connections have also been added.

- * There is now a requirement on the servers to handle non-persistent connections as this provides fault tolerance.
- * Added wording on the usage of Connection:Close for RTSP.
- * Specified usage of TLS for RTSP messages, including a scheme to approve a proxy's TLS connection to the next hop.
- o The following header related changes have been made:
 - * Accept-Ranges response-header is added. This header clarifies which range formats that can be used for a resource.
 - * Fixed the missing definitions for the Cache-Control header. Also added to the syntax definition the missing delta-seconds for max-stale and min-fresh parameters.
 - * Put requirement on CSeq header that the value is increased by one for each new RTSP request. A Recommendation to start at 0 has also been added.
 - * Added requirement that the Date header must be used for all messages with message body and the Server should always include it.
 - * Removed possibility of using Range header with Scale header to indicate when it is to be activated, since it can't work as defined. Also added rule that lack of Scale header in response indicates lack of support for the header. Feature-tags for scaled playback has been defined.
 - * The Speed header must now be responded to indicate support and the actual speed going to be used. A feature-tag is defined. Notes on congestion control were also added.
 - * The Supported header was borrowed from SIP [RFC3261] to help with the feature negotiation in RTSP.
 - * Clarified that the Timestamp header can be used to resolve retransmission ambiguities.
 - * The Session header text has been expanded with an explanation on keep-alive and which methods to use. SET_PARAMETER is now recommended to use if only keep-alive within RTSP is desired.
 - * It has been clarified how the Range header formats are used to indicate pause points in the PAUSE response.

- * Clarified that RTP-Info URIs that are relative, use the Request-URI as base URI. Also clarified that the used URI must be the one that was used in the SETUP request. The URIs are now also required to be quoted. The header also expresses the SSRC for the provided RTP timestamp and sequence number values.
 - * Added text that requires the Range to always be present in PLAY responses. Clarified what should be sent in case of live streams.
 - * The headers table has been updated using a structure borrowed from SIP. Those tables convey much more information and should provide a good overview of the available headers.
 - * It has been clarified that any message with a message body is required to have a Content-Length header. This was the case in RFC 2326, but could be misinterpreted.
 - * ETag has changed name to MTag.
 - * To resolve functionality around MTag. The MTag and If-None-Match header have been added from HTTP with necessary clarification in regards to RTSP operation.
 - * Imported the Public header from HTTP RFC 2068 [RFC2068] since it has been removed from HTTP due to lack of use. Public is used quite frequently in RTSP.
 - * Clarified rules for populating the Public header so that it is an intersection of the capabilities of all the RTSP agents in a chain.
 - * Added the Media-Range header for listing the current availability of the media range.
 - * Added the Notify-Reason header for giving the reason when sending PLAY_NOTIFY requests.
 - * A new header Seek-Style has been defined to direct and inform how any seek operation should/have been performed.
- o The Protocol Syntax has been changed in the following way:
 - * All ABNF definitions are updated according to the rules defined in RFC 5234 [RFC5234] and have been gathered in a separate Section 20.

- * The ABNF for the User-Agent and Server headers have been corrected.
- * Some definitions in the introduction regarding the RTSP session have been changed.
- * The protocol has been made fully IPv6 capable.
- * The CHAR rule has been changed to exclude NULL.
- o The Status codes have been changed in the following way:
 - * The use of status code 303 "See Other" has been deprecated as it does not make sense to use in RTSP.
 - * When sending response 451 and 458 the response body should contain the offending parameters.
 - * Clarification on when a 3rr redirect status code can be received has been added. This includes receiving 3rr as a result of a request within a established session. This provides clarification to a previous unspecified behavior.
 - * Removed the 201 (Created) and 250 (Low On Storage Space) status codes as they are only relevant to recording, which is deprecated.
 - * Several new Status codes have been defined: 464 "Data Transport Not Ready Yet", 465 "Notification Reason Unknown", 470 "Connection Authorization Required", 471 "Connection Credentials not accepted", 472 "Failure to establish secure connection".
- o The following functionality has been deprecated from the protocol:
 - * The use of Queued Play.
 - * The use of PLAY method for keep-alive in Play state.
 - * The RECORD and ANNOUNCE methods and all related functionality. Some of the syntax has been removed.
 - * The possibility to use timed execution of methods with the time parameter in the Range header.
 - * The description on how rtspu works is not part of the core specification and will require external description. Only that

it exists is defined here and some requirements for the transport is provided.

- o The following changes have been made in relation to methods:
 - * The OPTIONS method has been clarified with regards to the use of the Public and Allow headers.
 - * Added text clarifying the usage of SET_PARAMETER for keep-alive and usage without any body.
 - * PLAY method is now allowed to be pipelined with the pipelining of one or more SETUP requests following the initial that generates the session for aggregated control.
 - * REDIRECT has been expanded and diversified for different situations.
 - * Added a new method PLAY_NOTIFY. This method is used by the RTSP server to asynchronously notify clients about session changes.
- o Wrote a new section about how to setup different media transport alternatives and their profiles, and lower layer protocols. This caused the appendix on RTP interaction to be moved there instead of being in the part which describes RTP. The section also includes guidelines what to consider when writing usage guidelines for new protocols and profiles.
- o Setup and usage of independent TCP connections for transport of RTP has been specified.
- o Added a new section describing the available mechanisms to determine if functionality is supported, called "Capability Handling". Renamed option-tags to feature-tags.
- o Added a contributors section with people who have contributed actual text to the specification.
- o Added a section Use Cases that describes the major use cases for RTSP.
- o Clarified the usage of a=range and how to indicate live content that are not seekable with this header.
- o Text specifying the special behavior of PLAY for live content.
- o Security features of RTSP have been clarified:

- * HTTP based authorization has been clarified requiring both Basic and DIGEST support
- * TLS support mandated
- * IF one implements RTP then SRTP and defined MIKEY based key-exchange must be supported
- * Various minor mitigations discussed or resulted in protocol changes.

Appendix J. Acknowledgements

This memorandum defines RTSP version 2.0 which is a revision of the Proposed Standard RTSP version 1.0 which is defined in [RFC2326]. The authors of RFC 2326 are Henning Schulzrinne, Anup Rao, and Robert Lanphier.

Both RTSP version 1.0 and RTSP version 2.0 borrow format and descriptions from HTTP/1.1.

Robert Sparks and especially Elwyn Davies provided very valuable and detailed reviews in the IETF last call that greatly improved the document and resolved many issues, especially regarding consistency.

This document has benefited greatly from the comments of all those participating in the MMUSIC-WG. In addition to those already mentioned, the following individuals have contributed to this specification:

Rahul Agarwal, Claudio Allocchio, Jeff Ayars, Milko Boic, Torsten Braun, Brent Browning, Bruce Butterfield, Steve Casner, Maureen Chesire, Jinhang Choi, Francisco Cortes, Elwyn Davies, Spencer Dawkins, Kelly Djahandari, Martin Dunsmuir, Adrian Farrel, Stephen Farrell, Ross Finlayson, Eric Fleischman, Jay Geagan, Andy Grignon, Christian Groves, V. Guruprasad, Peter Haight, Mark Handley, Brad Hefta-Gaub, Volker Hilt, John K. Ho, Patrick Hoffman, Go Hori, Philipp Hoschka, Anne Jones, Ingemar Johansson, Jae-Hwan Kim, Anders Klemets, Ruth Lang, Barry Leiba, Stephanie Leif, Jonathan Lennox, Eduardo F. Llach, Chris Lonvick, Xavier Marjou, Thomas Marshall, Rob McCool, Martti Mela, David Oran, Joerg Ott, Joe Pallas, Maria Papadopouli, Sujal Patel, Ema Patki, Alagu Periyannan, Colin Perkins, Pekka Pessi, Igor Plotnikov, Pete Resnick, Peter Saint-Andre, Holger Schmidt, Jonathan Sergeant, Pinaki Shah, David Singer, Lior Sion, Jeff Smith, Alexander Sokolsky, Dale Stamm, John Francis Stracke, Geetha Srikantan, Scott Taylor, David Walker, Stephan Wenger, Dale R. Worley, and Byungjo Yoon, and especially to Flemming Andreassen.

J.1. Contributors

The following people have made written contributions that were included in the specification:

- o Tom Marshall contributed text on the usage of 3rr status codes.
- o Thomas Zheng contributed text on the usage of the Range in PLAY responses and proposed an earlier version of the PLAY_NOTIFY method.
- o Sean Sheedy contributed text on the timeout behavior of RTSP messages and connections, the 463 status code, and proposed an earlier version of the PLAY_NOTIFY method.
- o Greg Sherwood proposed an earlier version of the PLAY_NOTIFY method.
- o Fredrik Lindholm contributed text about the RTSP security framework.
- o John Lazzaro contributed the text for RTP over Independent TCP.
- o Aravind Narasimhan contributed by rewriting Media Transport Alternatives (Appendix C) and editorial improvements on a number of places in the specification.
- o Torbjorn Einarsson has done some editorial improvements of the text.

Appendix K. RFC Editor Consideration

Please replace RFC XXXX with the RFC number this specification receives.

Authors' Addresses

Henning Schulzrinne
Columbia University
1214 Amsterdam Avenue
New York, NY 10027
USA

Email: schulzrinne@cs.columbia.edu

Anup Rao
Cisco
USA

Email: anrao@cisco.com

Rob Lanphier
Seattle, WA
USA

Email: robla@robla.net

Magnus Westerlund
Ericsson AB
Faeroegatan 6
STOCKHOLM SE-164 80
SWEDEN

Email: magnus.westerlund@ericsson.com

Martin Stiernerling
NEC Laboratories Europe, NEC Europe Ltd.
Kurfuersten-Anlage 36
Heidelberg 69115
Germany

Phone: +49 (0) 6221 4342 113
Email: mls.ietf@gmail.com
URI: <http://www.stiernerling.org>

MMUSIC WG
Internet-Draft
Intended status: Standards Track
Expires: May 3, 2012

M. Garcia-Martin
Ericsson
S. Veikkolainen
Nokia
October 31, 2011

Session Description Protocol (SDP) Extension For Setting Up Audio and
Video Media Streams Over Circuit-Switched Bearers In The Public Switched
Telephone Network (PSTN)
draft-ietf-mmusic-sdp-cs-09

Abstract

This memo describes use cases, requirements, and protocol extensions for using the Session Description Protocol (SDP) Offer/Answer model for establishing audio and video media streams over circuit-switched bearers in the Public Switched Telephone Network (PSTN).

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 3, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	4
2. Conventions Used in This Document	5
3. Requirements	5
4. Overview of Operation	6
4.1. Example Call Flow	6
5. Protocol Description	8
5.1. Level of Compliance	8
5.2. Extensions to SDP	8
5.2.1. Connection Data	8
5.2.2. Media Descriptions	9
5.2.3. Correlating the PSTN Circuit-Switched Bearer with SDP	10
5.2.3.1. The "cs-correlation" attribute	11
5.2.3.2. Caller-ID Correlation Mechanism	12
5.2.3.3. User-User Information Element Correlation Mechanism	13
5.2.3.4. DTMF Correlation Mechanism	14
5.3. Negotiating the correlation mechanisms	15
5.3.1. Determining the Direction of the Circuit-Switched Bearer Setup	15
5.3.2. Populating the cs-correlation attribute	16
5.3.3. Considerations on successful correlation	16
5.4. Considerations for Usage of Existing SDP	17
5.4.1. Originator of the Session	17
5.4.2. Contact information	17
5.5. Offer/Answer mode extensions	18
5.5.1. Generating the Initial Offer	18
5.5.2. Generating the Answer	20
5.5.3. Offerer processing the Answer	22
5.5.4. Modifying the session	23
5.6. Formal Syntax	24
6. Example	25
7. Security Considerations	26
8. IANA Considerations	27
8.1. Registration of new cs-correlation SDP attribute	27
8.2. Registration of a new "nettype" value	28
8.3. Registration of new "addrtype" values	28
8.4. Registration of a new "proto" value	28
9. Acknowledgments	28
10. References	29
10.1. Normative References	29
10.2. Informative References	29
Authors' Addresses	30

1. Introduction

The Session Description Protocol (SDP) [RFC4566] is intended for describing multimedia sessions for the purposes of session announcement, session invitation, and other forms of multimedia session initiation. SDP is most commonly used for describing media streams that are transported over the Real-Time Transport Protocol (RTP) [RFC3550], using the profiles for audio and video media defined in RTP Profile for Audio and Video Conferences with Minimal Control [RFC3551].

However, SDP can be used to describe other transport protocols than RTP. Previous work includes SDP conventions for describing ATM bearer connections [RFC3108] and the Message Session Relay Protocol [RFC4975].

SDP is commonly carried in Session Initiation Protocol (SIP) [RFC3261] messages in order to agree on a common media description among the endpoints. An Offer/Answer Model with Session Description Protocol (SDP) [RFC3264] defines a framework by which two endpoints can exchange SDP media descriptions and come to an agreement as to which media streams should be used, along with the media related parameters.

In some scenarios it might be desirable to establish the media stream over a circuit-switched bearer connection even if the signaling for the session is carried over an IP bearer. An example of such a scenario is illustrated with two mobile devices capable of both circuit-switched and packet-switched communication over a low-bandwidth radio bearer. The radio bearer may not be suitable for carrying real-time audio or video media, and using a circuit-switched bearer would offer a better perceived quality of service. So, according to this scenario, SDP and its higher layer session control protocol (e.g., the Session Initiation Protocol (SIP) [RFC3261]) are used over regular IP connectivity, while the audio or video is received through the classical circuit-switched bearer.

Setting up a signaling relationship in the IP domain instead of just setting up a circuit-switched call offers also the possibility of negotiating in the same session other IP based media that is not sensitive to jitter and delay, for example, text messaging or presence information.

At a later point in time the mobile device might move to an area where a high-bandwidth packet-switched bearer, for example a Wireless Local Area Network (WLAN) connection, is available. At this point the mobile device may perform a handover and move the audio or video media streams over to the high-speed bearer. This implies a new

exchange of SDP Offer/Answer that lead to a re-negotiation of the media streams.

Other use cases exist. For example, an endpoint might have at its disposal circuit-switched and packet-switched connectivity, but the same audio or video codecs are not feasible for both access networks. For example, the circuit-switched audio or video stream supports narrow-bandwidth codecs, while the packet-switched access allows any other audio or video codec implemented in the endpoint. In this case, it might be beneficial for the endpoint to describe different codecs for each access type and get an agreement on the bearer together with the remote endpoint.

There are additional use cases related to third party call control where the session setup time is improved when the circuit-switched bearer in the PSTN is described together with one or more codecs.

The rest of the document is structured as follows: Section 2 provides the document conventions, Section 3 introduces the requirements, Section 4 presents an overview of the proposed solutions, and Section 5 contains the protocol description. Section 6 provides an example of descriptions of circuit-switched audio or video streams in SDP. Section 8 and Section 7 contain the IANA and Security considerations, respectively.

2. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [RFC2119] and indicate requirement levels for compliant implementations.

3. Requirements

This section presents the general requirements that are specific for the audio or video media streams over circuit-switched bearers.

REQ-1: A mechanism for endpoints to negotiate and agree on an audio or video media stream established over a circuit-switched bearer MUST be available.

REQ-2: The mechanism MUST allow the endpoints to combine circuit-switched audio or video media streams with other complementary media streams, for example, text messaging.

- REQ-3: The mechanism MUST allow the endpoint to negotiate the direction of the circuit-switched bearer, i.e., which endpoint is active when initiating the circuit-switched bearer.
- REQ-4: The mechanism MUST be independent of the type of the circuit-switched access (e.g., Integrated Services Digital Network (ISDN), Global System for Mobile Communication (GSM), etc.)
- REQ-5: There MUST be a mechanism that helps an endpoint to correlate an incoming circuit-switched bearer with the one negotiated in SDP, as opposed to another incoming call that is not related to that.
- REQ-6: It MUST be possible for endpoints to advertise different list of audio or video codecs in the circuit-switched audio or video stream from those used in a packet-switched audio or video stream.
- REQ-7: It MUST be possible for endpoints to not advertise the list of available codecs for circuit-switched audio or video streams.

4. Overview of Operation

The mechanism defined in this memo extends SDP and allows describing an audio or video media stream established over a circuit-switched bearer. New tokens are registered in the "c=" and "m=" lines to be able to describe a media stream over a circuit-switched bearer. These SDP extensions are described in Section 5.2. Since circuit-switched bearers are connection-oriented media streams, the mechanism re-uses the connection-oriented extensions defined in RFC 4145 [RFC4145] to negotiate the active and passive sides of a connection setup. This is further described in Section 5.3.1.

4.1. Example Call Flow

Consider the example presented in Figure 1. In this example, Alice is located in an environment where she has access to both IP and circuit-switched bearers for communicating with other endpoints. Alice decides that the circuit-switched bearer offers a better perceived quality of service for voice, and issues an SDP Offer containing the description of an audio media stream over circuit-switched bearer.

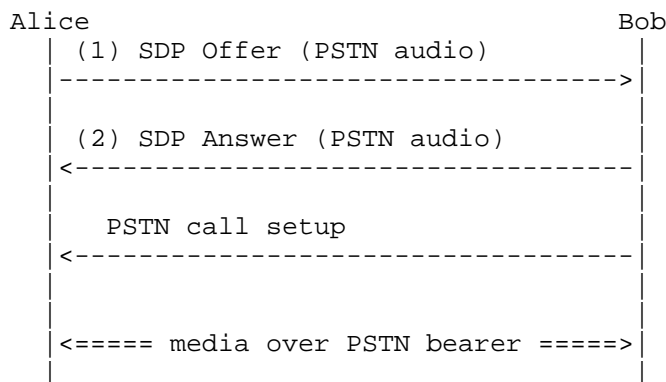


Figure 1: Example Flow

Bob receives the SDP offer and determines that he is located in an environment where the IP based bearer is not suitable for real-time audio media. However he also has PSTN circuit-switched bearer available for audio. Bob generates an SDP answer containing a description of the audio media stream over a circuit-switched bearer.

During the offer-answer exchange Alice and Bob also agree the direction in which the circuit-switched bearer should be established. In this example, Bob becomes the active party, in other words, he establishes the circuit-switched call to the other endpoint. The Offer/Answer exchange contains identifiers or references that can be used on the circuit-switched network for addressing the other endpoint, as well as information that is used to determine that the incoming circuit-switched bearer establishment is related to the ongoing session between Alice and Bob.

Bob establishes a circuit-switched bearer towards Alice using whatever mechanisms are defined for the network type in question. When receiving the incoming circuit-switched connection attempt, Alice is able to determine that the attempt is related to the session she is just establishing with Bob.

Alice accepts the circuit-switched connection; the circuit-switched bearer setup is completed. Bob and Alice can now use the circuit-switched connection for two-way audio media.

If, for some reason, Bob would like to reject the offered stream, he would set the port number of the specific stream to zero, as specified in RFC3264 [RFC3264]. Also, if Bob does not understand some of the SDP attributes specified in this document, he would ignore them, as specified in RFC4566 [RFC4566].

5. Protocol Description

5.1. Level of Compliance

Implementations according to this specification MUST implement the SDP extensions described in Section 5.2, and MUST implement the considerations discussed in Section 5.3, Section 5.4 and Section 5.5.

5.2. Extensions to SDP

This section provides the syntax and semantics of the extensions required for providing a description of audio or video media streams over circuit-switched bearers in SDP.

5.2.1. Connection Data

According to SDP [RFC4566], the connection data line in SDP has the following syntax:

```
c=<nettype> <addrtype> <connection-address>
```

where <nettype> indicates the network type, <addrtype> indicates the address type, and the <connection-address> is the connection address, which is dependent on the address type.

At the moment, the only network type defined is "IN", which indicates Internet network type. The address types "IP4" and "IP6" indicate the type of IP addresses.

This memo defines a new network type for describing a circuit-switched bearer network type in the PSTN. The mnemonic "PSTN" is used for this network type.

For the address type, we initially consider the possibility of describing E.164 telephone numbers. We define a new "E164" address type. When used, the "E164" address type indicates that the connection address contains an international E.164 number represented according to the ITU-T E.164 [ITU.E164.1991] recommendation.

It is a common convention that an international E.164 number contains a leading '+' sign. For consistency's sake, we also require the E.164 telephone is prepended with a '+', even if that is not necessary for routing of the call in the PSTN network.

There are cases, though, when the endpoint is merely aware of a circuit-switched bearer, without having further information about the address type or the E.164 number allocated to it. In these cases a dash "-" is used to indicate an unknown address type or connection

address. This makes the connection data line be according to the SDP syntax.

Note that <addrtype> and/or <connection-address> should not be omitted without being set to a "-" since this would violate basic syntax of SDP [RFC4566].

The following are examples of the extension to the connection data line:

```
c=PSTN E164 +35891234567
```

```
c=PSTN - -
```

When the <addrtype> is PSTN, the connection address is defined as follows:

- o an international E.164 number

When the <addrtype> is "-", the connection address is defined as follows:

- o the value "-", signifying that the address is unknown
- o any syntactically valid value, which is to be ignored

5.2.2. Media Descriptions

According to SDP [RFC4566], the media descriptions line in SDP has the following syntax:

```
m=<media> <port> <proto> <fmt> ...
```

The <media> sub-field carries the media type. For establishing an audio bearer, the existing "audio" media type is used. For establishing a video bearer, the existing "video" media type is used.

The <port> sub-field is the transport port to which the media stream is sent. Circuit-switched access lacks the concept of a port number, and therefore the <port> sub-field is set to the discard port "9".

According to RFC 3264 [RFC3264], a port number of zero in the offer of a unicast stream indicates that the stream is offered but must not be used. If a port number of zero is present in the answer of a unicast stream, it indicates that the stream is rejected. These rules are still valid when the media line in SDP represents a circuit-switched bearer.

The <proto> sub-field is the transport protocol. The circuit-switched bearer uses whatever transport protocol it has available. This subfield SHOULD be set to the mnemonic "PSTN" to be syntactically correct with SDP [RFC4566] and to indicate the usage of circuit-switched protocols in the PSTN.

The <fmt> sub-field is the media format description. In the classical usage of SDP to describe RTP-based media streams, when the <proto> sub-field is set to "RTP/AVP" or "RTP/SAVP", the <fmt> sub-field contains the payload types as defined in the RTP audio profile [RFC3551].

When "RTP/AVP" is used in the <proto> field, the <fmt> sub-field contains the RTP payload type numbers. We use the <fmt> sub-field to indicate the list of available codecs over the circuit-switched bearer, by re-using the conventions and payload type numbers defined for RTP/AVP. The RTP audio and video media types, which, when applied to PSTN circuit-switched bearers, represent merely an audio or video codec.

In some cases, the endpoint is not able to determine the list of available codecs for circuit-switched media streams. In this case, in order to be syntactically compliant with SDP [RFC4566], the endpoint MUST include a single dash "-" in the <fmt> sub-field.

As per RFC 4566 [RFC4566], the media format descriptions are listed in priority order.

Examples of media descriptions for circuit-switched audio streams are:

```
m=audio 9 PSTN 3 0 8
```

```
m=audio 9 PSTN -
```

Similarly, an example of a media description for circuit-switched video stream is:

```
m=video 9 PSTN 34
```

```
m=video 9 PSTN -
```

5.2.3. Correlating the PSTN Circuit-Switched Bearer with SDP

The endpoints should be able to correlate the circuit-switched bearer with the session negotiated with SDP in order to avoid ringing for an incoming circuit-switched bearer that is related to the session controlled with SDP (and SIP).

Several alternatives exist for performing this correlation. This memo provides three mutually non-exclusive correlation mechanisms. Other correlation mechanisms may exist, and their usage will be specified when need arises. All mechanisms share the same principle: some unique information is sent in the SDP and in the circuit-switched signaling protocol. If these pieces of information match, then the circuit-switched bearer is part of the session described in the SDP exchange. Otherwise, there is no guarantee that the circuit-switched bearer is related to such session.

The first mechanism is based on the exchange of PSTN caller-ID between the endpoints. The caller-ID is also available as the Calling Party ID in the circuit-switched signaling.

The second mechanism is based on the inclusion in SDP of a value that is also sent in the User-to-User Information Element that is part of the bearer setup signaling in the PSTN.

The third mechanism is based on sending in SDP a string that represents Dual Tone MultiFrequency (DTMF) digits that will be later sent right after the circuit-switched bearer is established. Implementations MAY use any of these mechanisms and MAY use two or more mechanisms simultaneously.

5.2.3.1. The "cs-correlation" attribute

In order to provide support for the correlation mechanisms, we define a new SDP attribute called "cs-correlation". This "cs-correlation" attribute can include any of the "callerid", "uuie", or "dtmf" sub-fields, which specify additional information required by the Caller-ID, User to User Information, or DTMF correlation mechanisms, respectively. The list of correlation mechanisms may be extended by other specifications.

The following sections provide more detailed information of these subfields. The "cs-correlation" attribute has the following format:

```
a=cs-correlation: <correlation-mechanisms>
correlation-mechanisms = corr-mech *(SP corr-mech)
corr-mech              = caller-id-mech / uuie-mech /
                        dfmt-mech / ext-mech
caller-id-mech         = "callerid" [":" caller-id-value]
uuie-mech              = "uuie" [":" uuie-value]
dtmf-mech              = "dtmf" [":" dtmf-value]
ext-mech               = <ext-mech-name>[":"<ext-mech-value>]
```

The values "callerid", "uuie" and "dtmf" refer to the correlation

mechanisms defined in Section 5.2.3.2, Section 5.2.3.3, and Section 5.2.3.4, respectively. The formal Augmented Backus-Naur Format (ABNF) syntax of the "cs-correlation" attribute is presented in Section 5.6.

5.2.3.2. Caller-ID Correlation Mechanism

The Caller-ID correlation mechanisms consists of an exchange of the calling party number as an international E.164 number in SDP, followed by the availability of the Calling Party Number information element in the call setup signaling of the circuit switched connection. If both pieces of information match, the circuit-switched bearer is correlated to the session described in SDP.

Example of inclusion of an international E.164 number in the "cs-correlation" attribute is:

```
a=cs-correlation:callerid:+35891234567
```

The presence of the "callerid" sub-field indicates that the endpoint supports use of the calling party number as a means of correlating a PSTN call with the session being negotiated. The "callerid" sub-field MAY be accompanied by the international E.164 number of the party inserting the parameter.

Note that there are no warranties that this correlation mechanism works or is even available, due a number of problems:

- o The endpoint might not be aware of its own E.164 number, in which case it cannot populate the SDP appropriately.
- o The Calling Party Number information element in the circuit-switched signaling might not be available, e.g., due to policy restrictions of the network operator or caller restriction due to privacy.
- o The Calling Party Number information element in the circuit-switched signaling might be available, but the digit representation of the E.164 number might differ from the one expressed in the SDP. To mitigate this problem implementations should consider only some of the rightmost digits from the E.164 number for correlation. For example, the numbers +358-9-123-4567 and 09-123-4567 could be considered as the same number. This is also the behavior of some cellular phones, which correlate the incoming calling party with a number stored in the phone book, for the purpose of displaying the caller's name.

5.2.3.3. User-User Information Element Correlation Mechanism

A second correlation mechanism is based on including in SDP a string that represents the User-User Information Element that is part of the call setup signaling of the circuit-switched bearer. The User-User Information Element is specified in ITU-T Q.931 [ITU.Q931.1998] and 3GPP TS 24.008 [TS.24.008], among others. The User-User Information Element has a maximum size of 35 or 131 octets, depending on the actual message of the PSTN protocol where it is included.

The mechanism works as follows: An endpoint creates a User-User Information Element, according to the requirements of the call setup signaling protocol. The same value is included in the SDP offer or SDP answer, in a "cs-correlation:uuie" attribute. When the SDP Offer/Answer exchange is completed, each endpoint has become aware of the value that will be used in the User-User Information Element of the call setup message of the PSTN protocol. The endpoint that initiates the call setup attempt includes this value in the User-User Information Element. The recipient of the call setup attempt can extract the User-User Information Element and correlate it with the value previously received in the SDP. If both values match, then the call setup attempt corresponds to that indicated in the SDP.

The first three octets of the User-User Information Element specified in ITU-T Q.931 [ITU.Q931.1998] are the UUIE identifier, length of the user-user contents, and a protocol discriminator, followed by the actual User information. The first three octets of the UUIE MUST NOT be used for correlation, only the octets carrying the User information value are compared the value of the "cs-correlation:uuie" attribute.

OPEN ISSUE: Need to confirm whether this is acceptable.

Note that, for correlation purposes, the value of the User-User Information Element is considered as a opaque string and only used for correlation purposes. Typically call signaling protocols impose requirements on the creation of User-User Information Element for end-user protocol exchange. The details regarding the generation of the User-User Information Element are outside the scope of this specification.

Please note that there are no warranties that this correlation mechanism works. On one side, policy restrictions might not make the User-User information available end to end in the PSTN. On the other hand, the generation of the User-User Information Element is controlled by the PSTN circuit-switched call protocol, which might not offer enough freedom for generating different values from one endpoint to another one, or from one call to another in the same

endpoint. This might result in the same value of the User-User Information Element for all calls.

5.2.3.4. DTMF Correlation Mechanism

We introduce a third mechanism for correlating the circuit-switched bearer with the session described with SDP. This is based on agreeing on a sequence of digits that are negotiated in the SDP Offer/Answer exchange and sent as Dual Tone Multifrequency (DTMF) tones over the circuit-switched bearer once this bearer is established. If the DTMF digit sequence received through the circuit-switched bearer matches the digit string negotiated in the SDP, the circuit-switched bearer is correlated with the session described in the SDP. The mechanism is similar to many voice conferencing systems which require the user to enter a PIN code using DTMF tones in order to be accepted in a voice conference.

The mechanism works as follows: An endpoint selects a DTMF digit sequence. The same sequence is included in the SDP offer or SDP answer, in a "cs-correlation:dtmf" attribute. When the SDP Offer/Answer exchange is completed, each endpoint has become aware of the DTMF sequence that will be sent right after the circuit-switched bearer is set up. The endpoint that initiates the call setup attempt sends the DTMF digits according to the procedures defined for the circuit-switched bearer technology used. The recipient (passive side of the bearer setup) of the call setup attempt collects the digits and compares them with the value previously received in the SDP. If the digits match, then the call setup attempt corresponds to that indicated in the SDP.

Implementations are advised to select a number of DTMF digits that provide enough assurance that the call is related, but on the other hand do not prolong the bearer setup time unnecessarily.

As an example, an endpoint willing to send DTMF tone sequence "14D*3" would include a "cs-correlation" attribute line as follows:

```
a=cs-correlation:dtmf:14D*3
```

If the endpoints successfully agree on the usage of the DTMF digit correlation mechanism, but the passive side does not receive any DTMF digits after successful circuit-switched bearer setup, or receives a set of DTMF digits that do not match the value of the "dtmf" attribute (including receiving too many digits), the passive side SHOULD treat the circuit-switched bearer as not correlated to the ongoing session.

DTMF digits can only be sent once the circuit-switched bearer is set up. In order to suppress alerting for an incoming circuit-switched call, implementations may choose various mechanisms. For example, alerting may be suppressed for a certain time period for incoming call attempts that originate from the number that was observed during the offer/answer negotiation.

5.3. Negotiating the correlation mechanisms

The three correlation mechanisms presented above (based on called party number, User-User Information Element and DTMF digit sending) are non-exclusive, and can be used independently of each other. In order to know how to populate the "a=cs-correlation" attribute, the endpoints need to agree which endpoint will become the active party, i.e. the one that will set up the circuit-switched bearer.

5.3.1. Determining the Direction of the Circuit-Switched Bearer Setup

In order to avoid a situation where both endpoints attempt to initiate a connection simultaneously, the direction in which the circuit-switched bearer is set up should be negotiated during the Offer/Answer exchange.

The framework defined in RFC 4145 [RFC4145] allows the endpoints to agree which endpoint acts as the active endpoint when initiating a TCP connection. While RFC 4145 [RFC4145] was originally designed for establishing TCP connections, it can be easily extrapolated to the connection establishment of circuit-switched bearers. This specification uses the concepts specified in RFC 4145 [RFC4145] for agreeing on the direction of establishment of a circuit-switched bearer.

RFC 4145 [RFC4145] defines two new attributes in SDP: "setup" and "connection". The "setup" attribute indicates which of the endpoints should initiate the connection establishment of the PSTN circuit-switched bearer. Four values are defined in Section 4 of RFC 4145 [RFC4145]: "active", "passive", "actpass", "holdconn". Please refer to Section 4 of RFC 4145 [RFC4145] for a detailed description of this attribute.

The "connection" attribute indicates whether a new connection is needed or an existing connection is reused. The attribute can take the values "new" or "existing". Please refer to Section 5 of RFC 4145 [RFC4145] for a detailed description of this attribute.

Implementations according to this specification MUST support the "setup" and "connection" attributes specified in RFC 4145 [RFC4145], but applied to circuit-switched bearers in the PSTN.

We define the active party as the one that initiates the circuit-switched bearer after the Offer/Answer process. The passive party is the one receiving the circuit-switched bearer. Either party may indicate its desire to become the active or passive party during the Offer/Answer exchange using the procedures described in Section 5.5.

5.3.2. Populating the cs-correlation attribute

By defining values for the sub-fields in the "a=cs-correlation" attribute, the endpoint indicates that it is willing to become the active party, and that it can use those values in the Calling party number, User-User Information Element, or as DTMF tones during the circuit-switched bearer setup.

Thus, the following rules apply:

An endpoint that can only become the active party in the circuit-switched bearer setup MUST include all correlation mechanisms it supports in the "a=cs-correlation" attribute, and MUST also specify values for the sub-fields.

An endpoint that can only become the passive party in the circuit-switched bearer setup MUST include all correlation mechanisms it supports in the "a=cs-correlation" attribute, but MUST NOT specify values for the sub-fields.

An endpoint that is willing to become either the active or passive party (by including the "a=setup:actpass" attribute in the Offer), MUST include all correlation mechanisms it supports in the "a=cs-correlation" attribute, and MUST also specify values for the sub-fields.

5.3.3. Considerations on successful correlation

Note that, as stated above, it cannot be guaranteed that any given correlation mechanism will succeed even if the usage of those was agreed beforehand. This is due to the fact that the correlation mechanisms require support from the circuit-switched bearer technology used.

Therefore, even a single positive indication using any of these mechanisms SHOULD be interpreted by the passive endpoint so that the circuit-switched bearer establishment is related to the ongoing session, even if the other correlation mechanisms fail.

If, after negotiating one or more correlation mechanisms in the SDP offer/answer exchange, an endpoint receives a circuit-switched bearer with no correlation information present, the endpoint has two

choices: it can either treat the call as unrelated, or treat the call as related to the ongoing session in the IP domain.

An endpoint may for example specify a time window after SDP offer/answer exchange during which received calls are treated as correlated even if the signaling in the circuit-switched domain does not carry any correlation information. In this case, there is a chance that the call is erroneously treated as related to the ongoing session.

An endpoint may also choose to always treat an incoming call as unrelated if the signaling in the circuit-switched domain does not carry any correlation information. In this case, there is a chance that the call is erroneously treated as unrelated.

Since, in these cases, no correlation information can be deduced from the signaling, it is up to the implementation to decide how to behave. One option is also to let the user decide whether to accept the call as related, or to treat the call as unrelated.

5.4. Considerations for Usage of Existing SDP

5.4.1. Originator of the Session

According to SDP [RFC4566], the origin line in SDP has the following syntax:

```
o=<username> <sess-id> <sess-version> <nettype> <addrtype>  
<unicast-address>
```

Of interest here are the <nettype> and <addrtype> fields, which indicate the type of network and type of address, respectively. Typically, this field carries the IP address of the originator of the session. Even if the SDP was used to negotiate an audio or video media stream transported over a circuit-switched bearer, the originator is using SDP over an IP bearer. Therefore, <nettype> and <addrtype> fields in the "o=" line should be populated with the IP address identifying the source of the signaling.

5.4.2. Contact information

SDP [RFC4566] defines the "p=" line which may include the phone number of the person responsible for the conference. Even though this line can carry a phone number, it is not suited for the purpose of defining a connection address for the media. Therefore, we have selected to define the PSTN specific connection addresses in the "c=" line.

5.5. Offer/Answer mode extensions

In this section, we define extensions to the Offer/Answer model defined in The Offer/Answer Model in SDP [RFC3264] and extended in the SDP Capability Negotiation [RFC5939] to allow for PSTN addresses to be used with the Offer/Answer model.

5.5.1. Generating the Initial Offer

The Offerer, wishing to use PSTN audio or video stream, MUST populate the "c=" and "m=" lines as follows.

The endpoint MUST set the <nettype> in the "c=" line to "PSTN", and the <addrtype> to "E164". Furthermore, the endpoint SHOULD set the <connection-address> field to its own international E.164 number (with a leading "+"). If the endpoint is not aware of its own E.164 number, it MUST set the <connection-address> to "-".

In the "m=" line, the endpoint MUST set the <media> subfield to "audio" or "video", depending on the media type, the <port> to "9" (the discard port), and the <proto> sub-field to "PSTN".

The <fmt> sub-field carries the payload type number(s) the endpoint is wishing to use. Payload type numbers in this case refer to the codecs that the endpoint wishes to use. For example, if the endpoint wishes to use the GSM codec, it would add payload type number 3 in the list of codecs.

For dynamic payload types, the endpoint MUST define the set of valid encoding names and related parameters using the "a=rtpmap" attribute line. See Section 6 of SDP [RFC4566] for details.

When generating the Offer, if the Offerer supports any of the correlation mechanisms defined in this memo, it MUST include an attribute line "a=cs-correlation" in the SDP offer. The "a=cs-correlation" line contains an enumeration of the correlation mechanisms supported by the Offerer, in the format of sub-fields.

The current list of sub-fields include "callerid", "uui" and "dtmf" and they refer to the correlation mechanisms defined in Section 5.2.3.2, Section 5.2.3.3, and Section 5.2.3.4, respectively.

If the Offerer supports any of the correlation mechanisms defined in this memo, and is willing to become the active party, the Offerer MUST add the "callerid", "uui", and/or "dtmf" sub-fields and MUST specify values for those sub-fields.

- o the international E.164 number as the value in the "callerid" sub-field,
- o the contents of the User-User information element as the value of the "uuie" sub-field, and/or
- o the DTMF tone string as the value of the "dtmf" sub-field

If the Offerer is only able to become the passive party in the circuit-switched bearer setup, it MUST add the "callerid", "uuie" and/or "dtmf" sub-fields, but MUST NOT specify values for those sub-fields.

For example, if the Offerer is willing to use the User-User Information element and DTMF digit sending mechanisms, but can only become the passive party, it includes the following lines to the SDP:

```
a=cs-correlation:uuie dtmf
```

```
a=setup:passive
```

If, on the other hand, the Offerer is willing to use the User-User Information element and the DTMF correlation mechanisms, and is able to become the active or passive side, it includes the following lines to the SDP:

```
a=cs-correlation:uuie:2890W284hAT452612908awudfjang908 dtmf:14D*3
```

```
a=setup:actpass
```

The negotiation of the value of the 'setup' attribute takes place as defined in Section 4.1 of TCP-Based Media Transport in the SDP [RFC4145].

The Offerer states which role or roles it is willing to perform; and the Answerer, taking the Offerer's willingness into consideration, chooses which roles both endpoints will actually perform during the circuit-switched bearer setup.

By 'active' endpoint, we refer to an endpoint that will establish the circuit-switched bearer; and by 'passive' endpoint, we refer to an endpoint that will receive a circuit-switched bearer.

If an Offerer does not know its international E.164 number, it MUST set the 'a=setup' attribute to the value 'active'. If the Offerer knows its international E.164 number, it MUST set the value to either 'actpass' or 'passive'.

Also 'holdconn' is a permissible value in the 'a=setup' attribute. It indicates that the connection is not established for the time being.

The Offerer uses the "a=connection" attribute to decide whether a new circuit-switched bearer is to be established or not. For the initial Offer, the Offerer MUST use value 'new'.

5.5.2. Generating the Answer

If the Offer contained a circuit-switched audio or video stream, the Answerer first determines whether it is able to accept and use such streams. If the Answerer is not willing to use circuit-switched media for the session, it MUST construct an Answer where the port number for such media stream(s) is set to zero, according to Section 6 of An Offer/Answer Model with the Session Description Protocol (SDP) [RFC3264].

If the Offer included a "-" as the payload type number, it indicates that the Offerer is not willing or able to define any specific payload type. Most often, a "-" is expected to be used instead of the payload type when the endpoint is not aware of or not willing to define the codecs which will eventually be used on the circuit-switched bearer. The circuit-switched signaling protocols have their own means of negotiating or indicating the codecs, therefore an Answerer SHOULD accept such Offers, and SHOULD set the payload type to "-" also in the Answer.

If the Answerer explicitly wants to specify a codec for the circuit-switched media, it MAY set the respective payload numbers in the <fmt> sub-field in the answer. This behavior, however, is NOT RECOMMENDED.

When receiving the Offer, the Answerer MUST determine whether it becomes the active or passive party.

If the SDP in the Offer indicates that the Offerer is only able to become the active party, the Answerer needs to determine whether it is able to become the passive party. If this is not possible e.g. due to the Answerer not knowing its international E.164 number, the Answerer MUST reject the circuit-switched media by setting the port number to zero on the Answer. If the Answerer is aware of its international E.164 number, it MUST include the "a=setup" attribute in the Answer and set it to value "passive" or "holdconn".

If the SDP in the Offer indicates that the Offerer is only able to become the passive party, the Answerer MUST verify that the Offerer's E.164 number is included in the "c=" line of the Offer. If the

number is included, the Answerer MUST include the "a=setup" attribute in the Answer and set it to value "active" or "holdconn". If the number is not included, call establishment is not possible, and the Answerer MUST reject the circuit-switched media by setting the port number to zero in the Answer.

If the SDP in the Offer indicates that the Offerer is able to become either the active or passive party, the Answerer needs to determine which role it would like to take. If the Offer includes an international E.164 number in the "c=" line, the Answerer SHOULD become the active party. If the Offer does not include an E.164 number, and if the Answerer is aware of its international E.164 number, it MUST become the passive party. If the Offer does not include an E.164 number in the "c=" line and the Answerer is not aware of its E.164 number, it MUST reject the circuit-switched media by setting the port number to zero in the Answer.

The Answerer MUST select those correlation mechanisms from the Offer it supports, and include an "a=cs-correlation" attribute line in the Answer containing those mechanisms it supports. The Answerer MUST NOT add any mechanisms which were not included in the offer.

If the Answerer becomes the active party, it MUST add parameter values to the "callerid", "uuie" or "dtmf" sub-fields.

If the Answerer becomes the passive party, it MUST NOT add values to the "callerid", "uuie" and/or "dtmf" sub-fields.

After generating and sending the Answer, if the Answerer became the active party, it

- o MUST extract the E.164 number from the "c=" line of the Offer and MUST establish a circuit-switched bearer to that address.
- o if the SDP Answer contained a value for the "callerid" sub-field, must set the Calling Party Number Information Element to that number,
- o if the SDP Answer contained a value for the "uuie" sub-field, MUST send the User-User Information element according to the rules defined for the circuit-switched technology used, and set the value of the Information Element to that received in the SDP Offer,
- o if the SDP Answer contained a value for the "dtmf" sub-field, MUST send those DTMF digits according to the circuit-switched technology used.

If, on the other hand, the Answerer became the passive party, it

- o MUST be prepared to receive a circuit-switched bearer,
- o if the Offer contained a value for the "callerid" sub-field, MUST compare that value to the Calling Party Number Information Element of the circuit-switched bearer,
- o if the Offer contained a value for the "dtmf" sub-field, MUST be prepared to receive and collect DTMF digits once the circuit-switched bearer is set up. The Answerer MUST compare the received DTMF digits to the value of the "dtmf" sub-field. If the received DTMF digits match the value of the "dtmf" sub-field in the "cs-correlation" attribute, the call SHOULD be treated as correlated to the ongoing session.
- o if the Offer contained a value for the "uuie" sub-field, MUST be prepared to receive a User-User Information element once the circuit-switched bearer is set up. The Answerer MUST compare the received UUI to the value of the "uuie" sub-field. If the value of the received UUI matches the value of the "uuie" sub-field, the call SHOULD be treated as correlated to the ongoing session.

5.5.3. Offerer processing the Answer

When receiving the Answer, if the SDP does not contain "a=cs-correlation" attribute line, the Offerer should take that as an indication that the other party does not support or is not willing to use the procedures defined in the document for this session, and MUST revert to normal processing of SDP.

When receiving the Answer, the Offerer MUST first determine whether it becomes the active or passive party, as described in Section 5.3.1.

If the Offerer becomes the active party, it

- o MUST extract the E.164 number from the "c=" line and MUST establish a circuit-switched bearer to that address.
- o if the SDP Answer contained a value for the "uuie" sub-field, MUST send the User-User Information element according to the rules defined for the circuit-switched technology used, and set the value of the Information Element to that received in the SDP Answer,
- o if the SDP Answer contained a value for the "dtmf" sub-field, MUST send those DTMF digits according to the circuit-switched

technology used.

If the Offerer becomes the passive party, it

- o MUST be prepared to receive a circuit-switched bearer,
- o if the Answer contained a value for the "dtmf" sub-field, MUST be prepared to receive and collect DTMF digits once the circuit-switched bearer is set up. The Offerer SHOULD compare the received DTMF digits to the value of the "dtmf" sub-field. If the received DTMF digits match the value of the "dtmf" sub-field in the "cs-correlation" attribute, the call SHOULD be treated as correlated to the ongoing session.
- o if the Answer contained a value for the "uuie" sub-field, MUST be prepared to receive a User-User Information element once the circuit-switched bearer is set up. The Offerer SHOULD compare the received UUI to the value of the "uuie" sub-field. If the value of the received UUI matches the value of the "uuie" sub-field, the call SHOULD be treated as correlated to the ongoing session.

5.5.4. Modifying the session

If, at a later time, one of the parties wishes to modify the session, e.g., by adding new media stream, or by changing properties used on an existing stream, it may do so via the mechanisms defined for An Offer/Answer Model with SDP [RFC3264].

If there is an existing circuit-switched bearer between the endpoints, and the Offerer wants to reuse that the Offerer MUST set the value of the "a=connection" attribute to 'existing'.

If either party removes the circuit-switched media from the session (by setting the port number to zero), it MUST terminate the circuit-switched bearer using whatever mechanism is appropriate for the technology in question.

If either party wishes to drop and reestablish an existing call, that party MUST first remove the circuit-switched media from the session by setting the port number to zero, and then use another Offer/Answer exchange where it MUST set the "a=connection" attribute to 'new'. If the media types are different (for example, a different codec will be used for the circuit-switched bearer), the media descriptions for terminating the existing bearer and the new bearer can be in the same Offer.

5.6. Formal Syntax

The following is the formal Augmented Backus-Naur Form (ABNF) [RFC5234] syntax that supports the extensions defined in this specification. The syntax is built above the SDP [RFC4566] grammar. Implementations according to this specification MUST be compliant with this syntax.

Figure 2 shows the formal syntax of the extensions defined in this memo.

```
; extension to the connection field originally specified
; in RFC 4566

connection-field    = [%x63 "=" nettype SP addrtype SP
connection-address CRLF]
;nettype and addrtype are defined in RFC 4566

connection-address /= e164-address / "-"
e164-address       = "+" 1*15DIGIT
; DIGIT is specified in RFC 5234

;subrules for correlation attribute
attribute          /= cs-correlation-attr
; attribute defined in RFC 4566
cs-correlation-attr= "cs-correlation:" corr-mechanisms
corr-mechanisms    = corr-mech *(SP corr-mech)
corr-mech          = caller-id-mech / uuie-mech / dtmf-mech / ext-mech
caller-id-mech     = "callerid" [":" caller-id-value]
caller-id-value    = "+" 1*15DIGIT
uuie-mech          = "uuie" [":" uuie-value]
uuie-value         = 1*32(ALPHA/DIGIT)
dtmf-mech         = "dtmf" [":" dtmf-value]
dtmf-value        = 1*32(DIGIT / %x41-44 / %x23 / %x2A )
;0-9, A-D, '#' and '*'
ext-mech          = ext-mech-name[":" ext-mech-value]
ext-mech-name      = token
ext-mech-value     = token
; token is specified in RFC4566
```

Figure 2: Syntax of the SDP extensions

6. Example

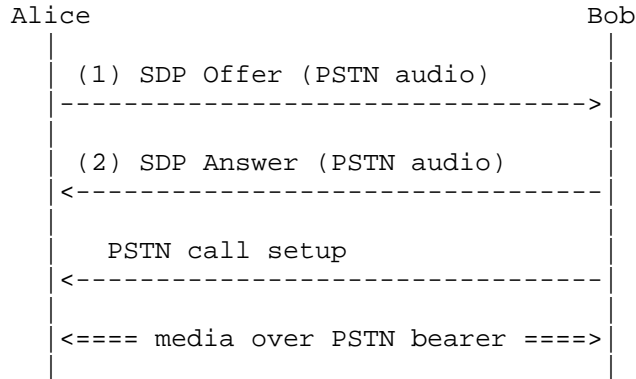


Figure 3: Basic flow

Figure 3 shows a basic example that describes a single audio media stream over a circuit-switched bearer. Alice generates a SDP Offer which is show in Figure 4. The Offer describes a PSTN circuit-switched bearer in the "m=" and "c=" line where it also indicates its international E.164 number format. Additionally, Alice expresses that she can initiate the circuit-switched bearer or be the recipient of it in the "a=setup" attribute line. The SDP Offer also includes a correlation identifiers that this endpoint will be inserting the Calling Party Number and/or User-User Information Element of the PSTN call setup if eventually this endpoint initiates the PSTN call.

```

v=0
o=jdoe 2890844526 2890842807 IN IP4 192.0.2.5
s=
t=0 0
m=audio 9 PSTN -
c=PSTN E164 +35891234567
a=setup:actpass
a=connection:new
a=cs-correlation:callerid:+15551234 uuie:2890W284hAT452612908awudfjang908
  
```

Figure 4: SDP offer (1)

Bob generates a SDP Answer (Figure 5), describing a PSTN audio media on port 9 without information on the media sub-type on the "m=" line. The "c=" line contains Bob's international E.164 number. In the "a=setup" line Bob indicates that he is willing to become the active endpoint when establishing the PSTN call, and he also includes the "a=cs-correlation" attribute line containing the values he is going to include in the Calling Party Number and User-User IE of the PSTN

call establishment.

```
v=0
o=- 2890973824 2890987289 IN IP4 192.0.2.7
s=
t=0 0
m=audio 9 PSTN -
c=PSTN E164 +35897654321
a=setup:active
a=connection:new
a=cs-correlation:callerid:+15554321 uuie:2127W49uThi455916adjfhtow9619361
```

Figure 5: SDP Answer with circuit-switched media

When Alice receives the Answer, she examines that Bob is willing to become the active endpoint when setting up the PSTN call. Alice temporarily stores Bob's E.164 number and the User-User IE value of the "cs-correlation" attribute, and waits for a circuit-switched bearer establishment.

Bob initiates a circuit-switched bearer using whatever circuit-switched technology is available for him. The called party number is set to Alice's number, and calling party number is set to Bob's own number. Bob also sets the User-User Information Element value to the one contained in the SDP Answer.

When Alice receives the circuit-switched bearer establishment, she examines the UUIE and the calling party number, and by comparing those received during O/A exchange determines that the call is related to the SDP session.

It may also be that neither the UUIE nor the calling party number is received by the called party, or the format of the calling party number is changed by the PSTN. Implementations may still accept such call establishment attempts as being related to the session that was established in the IP network. As it cannot be guaranteed that the values used for correlation are always passed intact through the network, they should be treated as additional hints that the circuit-switched bearer is actually related to the session.

7. Security Considerations

This document provides an extension on top of RFC 4566 [RFC4566], and RFC 3264 [RFC3264]. As such, the security considerations of those documents apply.

This memo provides mechanisms to agree on a correlation identifier or

identifiers that are used to evaluate whether an incoming circuit-switched bearer is related to an ongoing session in the IP domain. If an attacker replicates the correlation identifier and establishes a call within the time window the receiving endpoint is expecting a call, the attacker may be able to hijack the circuit-switched bearer. These types of attacks are not specific to the mechanisms presented in this memo. For example, caller ID spoofing is a well known attack in the PSTN. Users are advised to use the same caution before revealing sensitive information as they would on any other phone call. Furthermore, users are advised that mechanisms that may be in use in the IP domain for securing the media, like Secure RTP (SRTP) [RFC3711], are not available in the CS domain.

8. IANA Considerations

This document instructs IANA to register a number of SDP tokens according to the following data.

8.1. Registration of new cs-correlation SDP attribute

Contact: Miguel Garcia <miguel.a.garcia@ericsson.com>

Attribute name: cs-correlation

Long-form attribute name: PSTN Correlation Identifier

Type of attribute: media level only

This attribute is subject to the charset attribute

Description: This attribute provides the Correlation Identifier used in PSTN signaling

Specification: RFC XXXX

The IANA is requested to create a subregistry for 'cs-correlation' attribute under the Session Description Protocol (SDP) Parameters registry. The initial values for the subregistry are presented in the following, and IANA is requested to add them into its database:

Value of 'cs-correlation' attribute	Reference	Description
-----		callerid
RFC XXXX Caller ID	uuie	RFC XXXX User-User Information Element
RFC XXXX Dual-tone Multifrequency	dtmf	

Note for the RFC Editor: 'RFC XXXX' above should be replaced by a reference to the RFC number of this draft.

As per the terminology in [RFC2434], the registration policy for new values of 'cs-correlation' parameter is 'Specification Required'.

8.2. Registration of a new "nettype" value

This memo provides instructions to IANA to register a new "nettype" in the Session Description Protocol Parameters registry [1]. The registration data, according to RFC 4566 [RFC4566] follows.

Type	SDP Name	Reference
----	-----	-----
nettype	PSTN	[RFCxxxx]

8.3. Registration of new "addrtype" values

This memo provides instructions to IANA to register a new "addrtype" in the Session Description Protocol Parameters registry [1]. The registration data, according to RFC 4566 [RFC4566] follows.

Type	SDP Name	Reference
----	-----	-----
addrtype	E164	[RFCxxxx]
	-	[RFCxxxx]

8.4. Registration of a new "proto" value

This memo provides instructions to IANA to register a new "proto" in the Session Description Protocol Parameters registry [1]. The registration data, according to RFC 4566 [RFC4566] follows.

Type	SDP Name	Reference
-----	-----	-----
proto	PSTN	[RFCxxxx]

9. Acknowledgments

The authors want to thank Paul Kyzivat, Flemming Andreassen, Thomas Belling, John Elwell, Jari Mutikainen, Miikka Poikselka, Jonathan Rosenberg, Ingemar Johansson, Christer Holmberg, and Alf Heidermark for providing their insight and comments on this document.

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.
- [RFC3108] Kumar, R. and M. Mostafa, "Conventions for the use of the Session Description Protocol (SDP) for ATM Bearer Connections", RFC 3108, May 2001.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.
- [RFC4145] Yon, D. and G. Camarillo, "TCP-Based Media Transport in the Session Description Protocol (SDP)", RFC 4145, September 2005.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.
- [RFC5939] Andreassen, F., "Session Description Protocol (SDP) Capability Negotiation", RFC 5939, September 2010.

10.2. Informative References

- [ITU.E164.1991] International Telecommunications Union, "The International Public Telecommunication Numbering Plan", ITU-T Recommendation E.164, 1991.
- [ITU.Q931.1998] "Digital Subscriber Signalling System No. 1 (DSS 1) - ISDN User - Network Interface Layer 3 Specification for Basic Call Control", ISO Standard 9594-1, May 1998.
- [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, June 2002.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V.

Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.

[RFC3551] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, July 2003.

[RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.

[RFC4975] Campbell, B., Mahy, R., and C. Jennings, "The Message Session Relay Protocol (MSRP)", RFC 4975, September 2007.

[TS.24.008]
3GPP, "Mobile radio interface Layer 3 specification; Core network protocols; Stage 3", 3GPP TS 24.008 3.20.0, December 2005.

URIs

[1] <<http://www.iana.org/assignments/sdp-parameters>>

Authors' Addresses

Miguel A. Garcia-Martin
Ericsson
Calle Via de los Poblados 13
Madrid, ES 28033
Spain

Email: miguel.a.garcia@ericsson.com

Simo Veikkolainen
Nokia
P.O. Box 407
NOKIA GROUP, FI 00045
Finland

Phone: +358 50 486 4463
Email: simo.veikkolainen@nokia.com

MMUSIC WG
Internet-Draft
Intended status: Standards Track
Expires: August 15, 2014

M. Garcia-Martin
Ericsson
S. Veikkolainen
Nokia
February 11, 2014

Session Description Protocol (SDP) Extension For Setting Audio and Video
Media Streams Over Circuit-Switched Bearers In The Public Switched
Telephone Network (PSTN)
draft-ietf-mmusic-sdp-cs-23

Abstract

This memo describes use cases, requirements, and protocol extensions for using the Session Description Protocol (SDP) Offer/Answer model for establishing audio and video media streams over circuit-switched bearers in the Public Switched Telephone Network (PSTN).

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on August 15, 2014.

Copyright Notice

Copyright (c) 2014 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of

the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	3
2. Conventions Used in This Document	4
3. Requirements	5
4. Overview of Operation	5
4.1. Example Call Flow	6
5. Protocol Description	7
5.1. Level of Compliance	7
5.2. Extensions to SDP	7
5.2.1. Connection Data	7
5.2.2. Media Descriptions	9
5.2.3. Correlating the PSTN Circuit-Switched Bearer with SDP	10
5.2.3.1. The "cs-correlation" attribute	11
5.2.3.2. Caller-ID Correlation Mechanism	12
5.2.3.3. User-User Information Element Correlation Mechanism	13
5.2.3.4. DTMF Correlation Mechanism	14
5.2.3.5. The external correlation mechanism	15
5.2.3.6. Extensions to correlation mechanisms	16
5.3. Negotiating the correlation mechanisms	16
5.3.1. Determining the Direction of the Circuit-Switched Bearer Setup	17
5.3.2. Populating the cs-correlation attribute	17
5.3.3. Considerations on correlations	18
5.4. Considerations for Usage of Existing SDP	19
5.4.1. Originator of the Session	19
5.4.2. Contact information	19
5.5. Considerations for Usage of Third Party Call Control (3PCC)	20
5.6. Offer/Answer mode extensions	20
5.6.1. Generating the Initial Offer	20
5.6.2. Generating the Answer	23
5.6.3. Offerer processing the Answer	25
5.6.4. Modifying the session	27
5.7. Formal Syntax	28
6. Examples	29
6.1. Single PSTN audio stream	30
6.2. Advanced SDP example: Circuit-Switched Audio and Video Streams	32
7. Security Considerations	33
8. IANA Considerations	34
8.1. Registration of new cs-correlation SDP attribute	34
8.2. Registration of a new "nettype" value	35
8.3. Registration of new "addrtype" value	35

8.4. Registration of a new "proto" value	35
9. Acknowledgments	36
10. References	36
10.1. Normative References	36
10.2. Informative References	37
10.3. URIs	38
Authors' Addresses	38

1. Introduction

The Session Description Protocol (SDP) [RFC4566] is intended for describing multimedia sessions for the purposes of session announcement, session invitation, and other forms of multimedia session initiation. SDP is most commonly used for describing media streams that are transported over the Real-Time Transport Protocol (RTP) [RFC3550], using the profiles for audio and video media defined in RTP Profile for Audio and Video Conferences with Minimal Control [RFC3551].

However, SDP can be used to describe other media transport protocols than RTP. Previous work includes SDP conventions for describing ATM bearer connections [RFC3108] and the Message Session Relay Protocol [RFC4975].

SDP is commonly carried in Session Initiation Protocol (SIP) [RFC3261] messages in order to agree on a common media description among the endpoints. An Offer/Answer Model with Session Description Protocol (SDP) [RFC3264] defines a framework by which two endpoints can exchange SDP media descriptions and come to an agreement as to which media streams should be used, along with the media related parameters.

In some scenarios it might be desirable to establish the media stream over a circuit-switched bearer connection even if the signaling for the session is carried over an IP bearer. An example of such a scenario is illustrated with two mobile devices capable of both circuit-switched and packet-switched communication over a low-bandwidth radio bearer. The radio bearer may not be suitable for carrying real-time audio or video media, and using a circuit-switched bearer would offer a better perceived quality of service. So, according to this scenario, SDP and its higher layer session control protocol (e.g., the Session Initiation Protocol (SIP) [RFC3261]) are used over regular IP connectivity, while the audio or video is received through the classical circuit-switched bearer.

This document addresses only the use of circuit-switched bearers in the PSTN, not a generic circuit-switched network. The mechanisms presented below require a call signaling protocol of the PSTN to be

used (such as ITU-T Q.931 [ITU.Q931.1998] or 3GPP TS 24.008 [TS.24.008]).

Setting up a signaling relationship in the IP domain instead of just setting up a circuit-switched call offers also the possibility of negotiating in the same session other IP based media that is not sensitive to jitter and delay, for example, text messaging or presence information.

At a later point in time the mobile device might move to an area where a high-bandwidth packet-switched bearer, for example a Wireless Local Area Network (WLAN) connection, is available. At this point the mobile device may perform a handover and move the audio or video media streams over to the high-speed bearer. This implies a new exchange of SDP Offer/Answer that leads to a re-negotiation of the media streams.

Other use cases exist. For example, an endpoint might have at its disposal circuit-switched and packet-switched connectivity, but the same audio or video codecs are not feasible for both access networks. For example, the circuit-switched audio or video stream supports narrow-bandwidth codecs, while the packet-switched access allows any other audio or video codec implemented in the endpoint. In this case, it might be beneficial for the endpoint to describe different codecs for each access type and get an agreement on the bearer together with the remote endpoint.

There are additional use cases related to third party call control where the session setup time is improved when the circuit-switched bearer in the PSTN is described together with one or more codecs.

The rest of the document is structured as follows: Section 2 provides the document conventions, Section 3 introduces the requirements, Section 4 presents an overview of the proposed solutions, and Section 5 contains the protocol description. Section 6 provides an example of descriptions of circuit-switched audio or video streams in SDP. Section 7 and Section 8 contain the Security and IANA considerations, respectively.

2. Conventions Used in This Document

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14, RFC 2119 [RFC2119] and indicate requirement levels for compliant implementations.

3. Requirements

This section presents the general requirements that are specific for the audio or video media streams over circuit-switched bearers.

- REQ-1: A mechanism for endpoints to negotiate and agree on an audio or video media stream established over a circuit-switched bearer MUST be available.
- REQ-2: The mechanism MUST allow the endpoints to combine circuit-switched audio or video media streams with other complementary media streams, for example, text messaging.
- REQ-3: The mechanism MUST allow the endpoint to negotiate the direction of the circuit-switched bearer, i.e., which endpoint is active when initiating the circuit-switched bearer.
- REQ-4: The mechanism MUST be independent of the type of the circuit-switched access (e.g., Integrated Services Digital Network (ISDN), Global System for Mobile Communication (GSM), etc.)
- REQ-5: There MUST be a mechanism that helps an endpoint to correlate an incoming circuit-switched bearer with the one negotiated in SDP, as opposed to another incoming call that is not related to that. In case correlation by programmatic means is not possible, correlation may also be performed by the human user.
- REQ-6: It MUST be possible for endpoints to advertise different lists of audio or video codecs in the circuit-switched audio or video stream from those used in a packet-switched audio or video stream.
- REQ-7: It MUST be possible for endpoints to not advertise the list of available codecs for circuit-switched audio or video streams.

4. Overview of Operation

The mechanism defined in this memo extends SDP and allows describing an audio or video media stream established over a circuit-switched bearer. A new network type ("PSTN") and a new protocol type ("PSTN") are defined for the "c=" and "m=" lines to be able to describe a media stream over a circuit-switched bearer. These SDP extensions are described in Section 5.2. Since circuit-switched bearers are connection-oriented media streams, the mechanism re-uses the connection-oriented extensions defined in RFC 4145 [RFC4145] to

negotiate the active and passive sides of a connection setup. This is further described in Section 5.3.1.

4.1. Example Call Flow

Consider the example presented in Figure 1. In this example, Endpoint A is located in an environment where it has access to both IP and circuit-switched bearers for communicating with other endpoints. Endpoint A decides that the circuit-switched bearer offers a better perceived quality of service for voice, and issues an SDP Offer containing the description of an audio media stream over circuit-switched bearer.

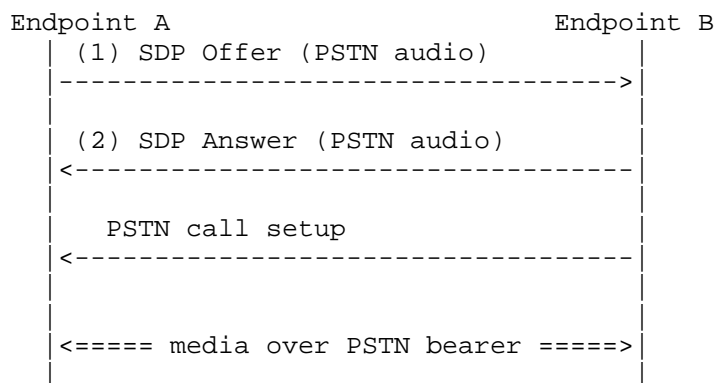


Figure 1: Example Flow

Endpoint B receives the SDP offer and determines that it is located in an environment where the IP based bearer is not suitable for real-time audio media. However, Endpoint B also has PSTN circuit-switched bearer available for audio. Endpoint B generates an SDP answer containing a description of the audio media stream over a circuit-switched bearer.

During the offer-answer exchange Endpoints A and B also agree the direction in which the circuit-switched bearer should be established. In this example, Endpoint B becomes the active party, in other words, it establishes the circuit-switched call to the other endpoint. The Offer/Answer exchange contains identifiers or references that can be used on the circuit-switched network for addressing the other endpoint, as well as information that is used to determine that the incoming circuit-switched bearer establishment is related to the ongoing session between the two endpoints.

Endpoint B establishes a circuit-switched bearer towards Endpoint A using whatever mechanisms are defined for the network type in

question. When receiving the incoming circuit-switched connection attempt, Endpoint A is able to determine that the attempt is related to the session it is just establishing with B.

Endpoint A accepts the circuit-switched connection; the circuit-switched bearer setup is completed. The two endpoints can now use the circuit-switched connection for two-way audio media.

If, for some reason, Endpoint B would like to reject the offered stream, it would set the port number of the specific stream to zero, as specified in RFC3264 [RFC3264]. Also, if B does not understand some of the SDP attributes specified in this document, it would ignore them, as specified in RFC4566 [RFC4566].

5. Protocol Description

5.1. Level of Compliance

Implementations according to this specification MUST implement the SDP extensions described in Section 5.2, and MUST implement the considerations discussed in Section 5.3, Section 5.4 and Section 5.6.

5.2. Extensions to SDP

This section provides the syntax and semantics of the extensions required for providing a description of audio or video media streams over circuit-switched bearers in SDP.

5.2.1. Connection Data

According to SDP [RFC4566], the connection data line in SDP has the following syntax:

```
c=<nettype> <addrtype> <connection-address>
```

where <nettype> indicates the network type, <addrtype> indicates the address type, and the <connection-address> is the connection address, which is dependent on the address type.

At the moment, the only network type defined is "IN", which indicates Internet network type. The address types "IP4" and "IP6" indicate the type of IP addresses.

This memo defines a new network type for describing a circuit-switched bearer network type in the PSTN. The mnemonic "PSTN" is used for this network type.

For the address type, we initially consider the possibility of describing E.164 telephone numbers. We define a new "E164" address type to be used within the context of a "PSTN" network type. The "E164" address type indicates that the connection address contains an E.164 number represented according to the ITU-T E.164 [ITU.E164.1991] recommendation.

It is a common convention that an international E.164 number contains a leading '+' sign. For consistency's sake, we also require the E.164 telephone is prepended with a '+', even if that is not necessary for routing of the call in the PSTN network.

There are cases, though, when the endpoint is merely aware of a circuit-switched bearer, without having further information about the E.164 number allocated to it. In these cases a dash ("-") is used to indicate an unknown connection address. This makes the connection data line be according to the SDP syntax.

Please note that the "E164" address type defined in this memo is exclusively defined to be used in conjunction with the "PSTN" network type in accordance with regular Offer/Answer procedures [RFC4566].

Note: RFC 3108 [RFC3108] also defines address type "E.164". This definition is distinct from the one defined by this memo and shall not be used with <nettype> "PSTN".

This memo exclusively uses the international representation of E.164 numbers, i.e., those including a country code and, as described above prepended with a '+' sign. Implementations conforming to this specification and using the "E164" address type together with the "PSTN" network type MUST use the 'global-number-digits' construction specified in RFC 3966 [RFC3966] for representing international E.164 numbers. This representation requires the presence of the '+' sign, and additionally allows for the presence of one or more 'visual-separator' constructions for easier human readability (see Section 5.7).

Note that <connection-address> MUST NOT be omitted when unknown since this would violate basic syntax of SDP [RFC4566]. In such cases, it MUST be set to a "-".

The following are examples of the extension to the connection data line:

```
c=PSTN E164 +441134960123
```

```
c=PSTN E164 -
```

When the <addrtype> is E164, the connection address is defined as follows:

- o an international E.164 number (prepended with a '+' sign)
- o the value "-", signifying that the address is unknown
- o any other value resulting from the production rule of connection-address in RFC4566 [RFC4566], but in all cases any value encountered will be ignored.

5.2.2. Media Descriptions

According to SDP [RFC4566], the media description line in SDP has the following syntax:

```
m=<media> <port> <proto> <fmt> ...
```

The <media> subfield carries the media type. For establishing an audio bearer, the existing "audio" media type is used. For establishing a video bearer, the existing "video" media type is used.

The <port> subfield is the transport port to which the media stream is sent. Circuit-switched access lacks the concept of a port number, and therefore the <port> subfield does not carry any meaningful value. In order to be compliant with SDP syntax, implementations SHOULD set the <port> subfield to the discard port value "9" and MUST ignore it on reception.

According to RFC 3264 [RFC3264], a port number of zero in the offer of a unicast stream indicates that the stream is offered but must not be used. If a port number of zero is present in the answer of a unicast stream, it indicates that the stream is rejected. These rules are still valid when the media line in SDP represents a circuit-switched bearer.

The <proto> subfield is the transport protocol. The circuit-switched bearer uses whatever transport protocol it has available. This subfield SHOULD be set to the mnemonic "PSTN" to be syntactically correct with SDP [RFC4566] and to indicate the usage of circuit-switched protocols in the PSTN.

The <fmt> subfield is the media format description. In the classical usage of SDP to describe RTP-based media streams, when the <proto> subfield is set to "RTP/AVP" or "RTP/SAVP", the <fmt> subfield contains the payload types as defined in the RTP audio profile [RFC3551].

When "RTP/AVP" is used in the <proto> field, the <fmt> subfield contains the RTP payload type numbers. We use the <fmt> subfield to indicate the list of available codecs over the circuit-switched bearer, by re-using the conventions and payload type numbers defined for RTP/AVP. The RTP audio and video media types, when applied to PSTN circuit-switched bearers, represent merely an audio or video codec. If the endpoint is able to determine the list of available codecs for circuit-switched media streams, it MUST use the corresponding payload type numbers in the <fmt> subfield.

In some cases, the endpoint is not able to determine the list of available codecs for circuit-switched media streams. In this case, in order to be syntactically compliant with SDP [RFC4566], the endpoint MUST include a single dash ("-") in the <fmt> subfield.

As per RFC 4566 [RFC4566], the media format descriptions are listed in priority order.

Examples of media descriptions for circuit-switched audio streams are:

```
m=audio 9 PSTN 3 0 8
```

```
m=audio 9 PSTN -
```

Similarly, an example of a media description for circuit-switched video stream is:

```
m=video 9 PSTN 34
```

```
m=video 9 PSTN -
```

5.2.3. Correlating the PSTN Circuit-Switched Bearer with SDP

The endpoints should be able to correlate the circuit-switched bearer with the session negotiated with SDP in order to avoid ringing for an incoming circuit-switched bearer that is related to the session controlled with SDP (and SIP).

Several alternatives exist for performing this correlation. This memo provides three mutually non-exclusive correlation mechanisms. Additionally, we define a fourth mechanism where correlation may be performed by external means, typically by the human user, in case using other correlation mechanisms is not possible or does not succeed. Other correlation mechanisms may exist, and their usage will be specified when need arises.

All mechanisms share the same principle: some unique information is sent in the SDP and in the circuit-switched signaling protocol. If these pieces of information match, then the circuit-switched bearer is part of the session described in the SDP exchange. Otherwise, there is no guarantee that the circuit-switched bearer is related to such session.

The first mechanism is based on the exchange of PSTN caller-ID between the endpoints. The caller-ID is also available as the Calling Party ID in the circuit-switched signaling.

The second mechanism is based on the inclusion in SDP of a value that is also sent in the User-to-User Information Element that is part of the bearer setup signaling in the PSTN.

The third mechanism is based on sending in SDP a string that represents Dual-Tone Multi-Frequency (DTMF) digits that will be later sent right after the circuit-switched bearer is established.

The fourth correlation mechanism declares support for cases where correlation is done by external means. Typically this means that the decision is left for the human user. This is the way how some current conferencing systems operate: after logging on to the conference, the system calls back to the user's phone number to establish audio communications, and it is up to the human user to accept or reject the incoming call. By declaring explicit support for this mechanism endpoints can use it only when such possibility exist.

Endpoints may opt to implement any combination of the correlation mechanisms specified in Section 5.2.3.2, Section 5.2.3.3, Section 5.2.3.4, and Section 5.2.3.5, including an option of implementing none at all.

5.2.3.1. The "cs-correlation" attribute

In order to provide support for the correlation mechanisms, we define a new media-level SDP attribute called "cs-correlation". There MUST be at most one "cs-correlation" attribute per media description.

This "cs-correlation" attribute MAY contain zero or more subfields, either "callerid", "uuie", "dtmf", or "external" to specify additional information required by the Caller-ID, User to User Information, DTMF, or external correlation mechanisms, respectively. The list of correlation mechanisms may be extended by other specifications, see Section 5.2.3.6 for more details.

The following sections provide more detailed information of these subfields.

The values "callerid", "uuie", "dtmf" and "external" refer to the correlation mechanisms defined in Section 5.2.3.2, Section 5.2.3.3, Section 5.2.3.4 and, Section 5.2.3.5 respectively. The formal Augmented Backus-Naur Format (ABNF) syntax of the "cs-correlation" attribute is presented in Section 5.7.

5.2.3.2. Caller-ID Correlation Mechanism

The Caller-ID correlation mechanisms consists of an exchange of the calling party number as an international E.164 number in SDP, followed by the availability of the Calling Party Number information element in the call setup signaling of the circuit switched connection. If both pieces of information match, the circuit-switched bearer is correlated to the session described in SDP.

Example of inclusion of an international E.164 number in the "cs-correlation" attribute is:

```
a=cs-correlation:callerid:+441134960123
```

The presence of the "callerid" subfield indicates that the endpoint supports use of the calling party number as a means of correlating a PSTN call with the session being negotiated. The "callerid" subfield MAY be accompanied by the international E.164 number of the party inserting the parameter.

Note that there are no guarantees that this correlation mechanism works or is even available, due a number of problems:

- o The endpoint might not be aware of its own E.164 number, in which case it cannot populate the SDP appropriately.
- o The Calling Party Number information element in the circuit-switched signaling might not be available, e.g., due to policy restrictions of the network operator or caller restriction due to privacy.
- o The Calling Party Number information element in the circuit-switched signaling might be available, but the digit representation of the E.164 number might differ from the one expressed in the SDP, due to, e.g., lack of country code. To mitigate this problem implementations should consider only some of the rightmost digits from the E.164 number for correlation. For example, the numbers +44-113-496-0123 and 0113-496-0123 could be considered as the same number. This is also the behavior of some

cellular phones, which correlate the incoming calling party with a number stored in the phone book, for the purpose of displaying the caller's name. Please refer to ITU-T E.164 recommendation [ITU.E164.1991] for consideration of the relevant number of digits to consider.

5.2.3.3. User-User Information Element Correlation Mechanism

A second correlation mechanism is based on including in SDP a string that represents the User-User Information Element that is part of the call setup signaling of the circuit-switched bearer. The User-User Information Element is specified in ITU-T Q.931 [ITU.Q931.1998] and 3GPP TS 24.008 [TS.24.008], among others. The User-User Information Element has a maximum size of 35 or 131 octets, depending on the actual message of the PSTN protocol where it is included and the network settings.

The mechanism works as follows: An endpoint creates a User-User Information Element, according to the requirements of the call setup signaling protocol. The same value is included in the SDP offer or SDP answer, in the "uuie" subfield of the "cs-correlation" attribute. When the SDP Offer/Answer exchange is completed, each endpoint has become aware of the value that will be used in the User-User Information Element of the call setup message of the PSTN protocol. The endpoint that initiates the call setup attempt includes this value in the User-User Information Element. The recipient of the call setup attempt can extract the User-User Information Element and correlate it with the value previously received in the SDP. If both values match, then the call setup attempt corresponds to that indicated in the SDP.

According to ITU-T Q.931 [ITU.Q931.1998], the User-User Information Element (UUIE) identifier is composed of a first octet identifying this as a User-User Information Element, a second octet containing the length of the user-user contents, a third octet containing a Protocol Discriminator, and a value of up to 32 or 128 octets (depending on network settings) containing the actual User Information (see Figure 4-36 in ITU-T Q.931). The first two octets of the UUIE MUST NOT be used for correlation, only the octets carrying the Protocol Discriminator and the User Information value are input to the creation of the value of the "uuie" subfield in the "cs-correlation" attribute. Therefore, the value of the "uuie" subfield in the "cs-correlation" attribute MUST start with the Protocol Discriminator octet, followed by the User Information octets. The value of the Protocol Discriminator octet is not specified in this document; it is expected that organizations using this technology will allocate a suitable value for the Protocol Discriminator.

Once the binary value of the "uuie" subfield in the "cs-correlation" attribute is created, it MUST be base 16 (also known as "hex") encoded before it is inserted in SDP. Please refer to RFC 4648 [RFC4648] for a detailed description of base 16 encoding. The resulting encoded value needs to have an even number of hexadecimal digits, and MUST be considered invalid if it has an odd number.

Note that the encoding of the "uuie" subfield of the "cs-correlation" attribute is largely inspired by the encoding of the same value in the User-to-User header field in SIP, according to the document "A Mechanism for Transporting User to User Call Control Information in SIP" [I-D.ietf-cuss-sip-uui].

As an example, an endpoint willing to send a UUIE containing a protocol discriminator with the hexadecimal value of %x56 and an hexadecimal User Information value of %xA390F3D2B7310023 would include a "cs-correlation" attribute line as follows:

```
a=cs-correlation:uuie:56A390F3D2B7310023
```

Note that the value of the User-User Information Element is considered as an opaque string and only used for correlation purposes. Typically call signaling protocols impose requirements on the creation of User-User Information Element for end-user protocol exchange. The details regarding the generation of the User-User Information Element are outside the scope of this specification.

Please note that there are no guarantees that this correlation mechanism works. On one side, policy restrictions might not make the User-User information available end to end in the PSTN. On the other hand, the generation of the User-User Information Element is controlled by the PSTN circuit-switched call protocol, which might not offer enough freedom for generating different values from one endpoint to another one, or from one call to another in the same endpoint. This might result in the same value of the User-User Information Element for all calls.

5.2.3.4. DTMF Correlation Mechanism

We introduce a third mechanism for correlating the circuit-switched bearer with the session described with SDP. This is based on agreeing on a sequence of digits that are negotiated in the SDP Offer/Answer exchange and sent as Dual-Tone Multi-Frequency (DTMF) ITU-T Recommendation Q.23 [ITU.Q23.1988] tones over the circuit-switched bearer once this bearer is established. If the DTMF digit sequence received through the circuit-switched bearer matches the digit string negotiated in the SDP, the circuit-switched bearer is correlated with the session described in the SDP. The mechanism is similar to many

voice conferencing systems which require the user to enter a PIN code using DTMF tones in order to be accepted in a voice conference.

The mechanism works as follows: An endpoint selects a DTMF digit sequence. The same sequence is included in the SDP offer or SDP answer, in a "dtmf" subfield of the "cs-correlation" attribute. When the SDP Offer/Answer exchange is completed, each endpoint has become aware of the DTMF sequence that will be sent right after the circuit-switched bearer is set up. The endpoint that initiates the call setup attempt sends the DTMF digits according to the procedures defined for the circuit-switched bearer technology used. The recipient (passive side of the bearer setup) of the call setup attempt collects the digits and compares them with the value previously received in the SDP. If the digits match, then the call setup attempt corresponds to that indicated in the SDP.

Implementations are advised to select a number of DTMF digits that provide enough assurance that the call is related, but on the other hand do not prolong the bearer setup time unnecessarily. A number of 5 to 10 digits is a good compromise.

As an example, an endpoint willing to send DTMF tone sequence "14D*3" would include a "cs-correlation" attribute line as follows:

```
a=cs-correlation:dtmf:14D*3
```

If the endpoints successfully agree on the usage of the DTMF digit correlation mechanism, but the passive side does not receive any DTMF digits after successful circuit-switched bearer setup, or receives a set of DTMF digits that do not match the value of the "dtmf" attribute (including receiving too many digits), the passive side SHOULD consider that this DTMF mechanism has failed to correlate the incoming call.

5.2.3.5. The external correlation mechanism

The fourth correlation mechanism relies on external means for correlating the incoming call to the session. Since endpoints can select which correlation mechanisms they support, it may happen that no other common correlation mechanism is found, or that the selected correlation mechanism does not succeed due to the required feature not being supported by the underlying PSTN network. In these cases, the human user can do the decision of accepting or rejecting the incoming call, thus "correlating" the call with the session. Since not all endpoints are operated by a human user, or if there is no other external means implemented by the endpoint for the correlation function, we explicitly define support for such external correlation mechanism.

Endpoints wishing to use this external correlation mechanism would use a subfield "external" in the "a=cs-correlation" attribute. Unlike the three other correlation mechanism, the "external" subfield does not accept a value. An example of a "a=cs-correlation" attribute line would look like this:

```
a=cs-correlation:external
```

Endpoints which are willing to only use the three explicit correlation mechanisms defined in this document ("callerid", "uuie", and/or "dtmf") would not include the "external" mechanism in the Offer/Answer exchange.

The external correlation mechanism typically relies on the human user to do the decision on whether the call is related to the ongoing session or not. After the user accepts the call, that bearer is considered as related to the session. There is a small chance that the user receives at the same time another circuit-switched call which is not related to the ongoing session. The user may reject this call if he is able to determine (e.g. based on the calling line identification) that the call is not related to the session, and continue waiting for another call attempt. If the user accepts the incoming circuit-switched call, but it turns out to be not related to the session, the endpoints need to rely on the human user to take appropriate action (typically, they would hang up).

5.2.3.6. Extensions to correlation mechanisms

New values for the "cs-correlation" attribute may be specified. The registration policy for new values is "Specification Required", see Section 8. Any such specification MUST include a description of how SDP Offer/Answer mechanism is used to negotiate the use of the new values, taking into account how endpoints determine which side will become active or passive (see Section 5.3 for more details).

If, during the Offer/Answer negotiation, either endpoint encounters an unknown value in the "cs-correlation" attribute, it MUST consider that mechanism as unsupported, and MUST NOT include that value in subsequent Offer/Answer negotiation.

5.3. Negotiating the correlation mechanisms

The four correlation mechanisms presented above (based on called party number, User-User Information Element, DTMF digit sending, and external) are non-exclusive, and can be used independently of each other. In order to know how to populate the "cs-correlation" attribute, the endpoints need to agree which endpoint will become the

active party, i.e., the one that will set up the circuit-switched bearer.

5.3.1. Determining the Direction of the Circuit-Switched Bearer Setup

In order to avoid a situation where both endpoints attempt to initiate a connection simultaneously, the direction in which the circuit-switched bearer is set up **MUST** be negotiated during the Offer/Answer exchange.

The framework defined in RFC 4145 [RFC4145] allows the endpoints to agree which endpoint acts as the active endpoint when initiating a TCP connection. While RFC 4145 [RFC4145] was originally designed for establishing TCP connections, it can be easily extrapolated to the connection establishment of circuit-switched bearers. This specification uses the concepts specified in RFC 4145 [RFC4145] for agreeing on the direction of establishment of a circuit-switched bearer.

RFC 4145 [RFC4145] defines two new attributes in SDP: "setup" and "connection". The "setup" attribute indicates which of the endpoints should initiate the connection establishment of the PSTN circuit-switched bearer. Four values are defined in Section 4 of RFC 4145 [RFC4145]: "active", "passive", "actpass", "holdconn". Please refer to Section 4 of RFC 4145 [RFC4145] for a detailed description of this attribute.

The "connection" attribute indicates whether a new connection is needed or an existing connection is reused. The attribute can take the values "new" or "existing". Please refer to Section 5 of RFC 4145 [RFC4145] for a detailed description of this attribute.

Implementations according to this specification **MUST** support the "setup" and "connection" attributes specified in RFC 4145 [RFC4145], but applied to circuit-switched bearers in the PSTN.

We define the active party as the one that initiates the circuit-switched bearer after the Offer/Answer exchange. The passive party is the one receiving the circuit-switched bearer. Either party may indicate its desire to become the active or passive party during the Offer/Answer exchange using the procedures described in Section 5.6.

5.3.2. Populating the cs-correlation attribute

By defining values for the subfields in the "a=cs-correlation" attribute, the endpoint indicates that it is willing to become the active party, and that it can use those values in the Calling party

number, User-User Information Element, or as DTMF tones during the circuit-switched bearer setup.

Thus, the following rules apply:

An endpoint that can only become the active party in the circuit-switched bearer setup MUST include all correlation mechanisms it supports in the "a=cs-correlation" attribute, and MUST also specify values for the "callerid", "uuie" and "dtmf" subfields. Notice that the "external" subfield does not accept a value.

An endpoint that can only become the passive party in the circuit-switched bearer setup MUST include all correlation mechanisms it supports in the "a=cs-correlation" attribute, but MUST NOT specify values for the subfields.

An endpoint that is willing to become either the active or passive party (by including the "a=setup:actpass" attribute in the Offer), MUST include all correlation mechanisms it supports in the "a=cs-correlation" attribute, and MUST also specify values for the "callerid", "uuie" and "dtmf" subfields. Notice that the "external" subfield does not accept a value.

5.3.3. Considerations on correlations

Passive endpoints should expect an incoming CS call for setting up the audio bearer. Passive endpoints MAY suppress the incoming CS alert during a certain time periods. Additional restrictions can be applied, such as the passive endpoint not alerting incoming calls originated from the number that was observed during the offer/answer negotiation.

There may be cases when an endpoint is not willing to include one or more correlation mechanisms in the "a=cs-correlation" attribute line even if it supports it. For example, some correlation mechanisms can be omitted if the endpoint is certain that the PSTN network does not support carrying the correlation identifier. Also, since using the DTMF based correlation mechanism requires the call to be accepted before DTMF tones can be sent, some endpoints may enforce a policy restricting this due to for example cost associated with received calls, making the DTMF based mechanism unusable.

Note that it cannot be guaranteed that the correlation mechanisms relying on caller identification, User-User Information Element and DTMF sending will succeed even if the usage of those was agreed beforehand. This is due to the fact that the correlation mechanisms require support from the circuit-switched bearer technology used.

Therefore, even a single positive indication using any of these mechanisms SHOULD be interpreted by the passive endpoint so that the circuit-switched bearer establishment is related to the ongoing session, even if the other correlation mechanisms fail.

If, after successfully negotiating any of the "callerid", "uuie" or "dtmf" correlation mechanisms in the SDP offer/answer exchange, an endpoint receives an incoming establishment of a circuit-switched bearer with no correlation information present, the endpoint first checks whether the offer/answer exchange was used to successfully negotiate also the "external" correlation mechanism. If it was, the endpoint should leave the decision to be made by this external means, typically the human user. If the "external" correlation mechanism was not successfully negotiated, the endpoint should treat the call as unrelated to the ongoing session in the IP domain.

5.4. Considerations for Usage of Existing SDP

5.4.1. Originator of the Session

According to SDP [RFC4566], the origin line in SDP has the following syntax:

```
o=<username> <sess-id> <sess-version> <nettype> <addrtype>
   <unicast-address>
```

Of interest here are the <nettype> and <addrtype> fields, which indicate the type of network and type of address, respectively. Typically, this field carries the IP address of the originator of the session. Even if the SDP was used to negotiate an audio or video media stream transported over a circuit-switched bearer, the originator is using SDP over an IP bearer. Therefore, <nettype> and <addrtype> fields in the "o=" line should be populated with the IP address identifying the source of the signaling.

5.4.2. Contact information

SDP [RFC4566] defines the "p=" line which may include the phone number of the person responsible for the conference. Even though this line can carry a phone number, it is not suited for the purpose of defining a connection address for the media. Therefore, we have selected to define the PSTN specific connection addresses in the "c=" line.

5.5. Considerations for Usage of Third Party Call Control (3PCC)

Best Current Practices for Third Party Call Control (3pcc) in the Session Initiation Protocol (SIP) [RFC3725] outlines several flows which are possible in third party call control scenarios and recommends some flows for specific situations.

One of the assumptions in [RFC3725] is that an SDP Offer may include a "black hole" connection address, which has the property that packets sent to it will never leave the host which sent them. For IPv4, this "black hole" connection address is 0.0.0.0, or a domain name within the .invalid DNS top level domain.

When using an E.164 address scheme in the context of third-party call control, when the User Agent needs to indicate an unknown phone number, it MUST populate the <addrtype> of the SDP "c=" line with a "-" string.

Note that this may result in the recipient of the initial offer rejecting such offer if the recipient of the offer was not aware of its own E.164 number. Consequently it will not be possible to establish a circuit-switched bearer, since neither party is aware of their E.164 number.

5.6. Offer/Answer mode extensions

In this section, we define extensions to the Offer/Answer model defined in The Offer/Answer Model in SDP [RFC3264] to allow for PSTN addresses to be used with the Offer/Answer model.

5.6.1. Generating the Initial Offer

The Offerer, wishing to use PSTN audio or video stream, MUST populate the "c=" and "m=" lines as follows.

The endpoint MUST set the <nettype> in the "c=" line to "PSTN", and the <addrtype> to "E164". Furthermore, the endpoint SHOULD set the <connection-address> field to its own international E.164 number (with a leading "+"). If the endpoint is not aware of its own E.164 number, it MUST set the <connection-address> to "-".

In the "m=" line, the endpoint MUST set the <media> subfield to "audio" or "video", depending on the media type, and the <proto> subfield to "PSTN". The <port> subfield SHOULD be set to "9" (the discard port). The values "audio" or "video" in the <media> subfield MUST NOT be set by the endpoint unless it has knowledge that these bearer types are available on the circuit-switched network.

The <fmt> subfield carries the payload type number(s) the endpoint is wishing to use. Payload type numbers in this case refer to the codecs that the endpoint wishes to use on the PSTN media stream. For example, if the endpoint wishes to use the GSM codec, it would add payload type number 3 in the list of codecs. The list of payload types MUST only contain those codecs the endpoint is able to use on the PSTN bearer. In case the endpoint is not aware of the codecs available for the circuit-switched media streams, it MUST include a dash ("-") in the <fmt> subfield.

The mapping table of static payload types numbers to payload types is initially specified in [RFC3551] and maintained by IANA. For dynamic payload types, the endpoint MUST define the set of valid encoding names and related parameters using the "a=rtpmap" attribute line. See Section 6 of RFC 4566 [RFC4566] for details.

When generating the Offer, the Offerer MUST include an attribute line "a=cs-correlation" in the SDP offer. The Offerer MUST NOT include more than one "cs-correlation" attribute per media description. The "a=cs-correlation" line SHOULD contain an enumeration of all the correlation mechanisms supported by the Offerer, in the format of subfields. See Section 5.3.3 for more information on usage of the correlation mechanisms.

The current list of subfields include "callerid", "uuie", "dtmf", and "external", and they refer to the correlation mechanisms defined in Section 5.2.3.2, Section 5.2.3.3, Section 5.2.3.4, and Section 5.2.3.5 respectively.

If the Offerer supports any of the correlation mechanisms defined in this memo, and is willing to become the active party, the Offerer MUST add the "callerid", "uuie", "dtmf" and/or "extern" subfields and MUST specify values for them as follows:

- o the international E.164 number as the value in the "callerid" subfield,
- o the contents of the User-User information element as the value of the "uuie" subfield, and/or
- o the DTMF tone string as the value of the "dtmf" subfield
- o for the "external" subfield, the endpoint MUST NOT specify any value.

If the Offerer is only able to become the passive party in the circuit-switched bearer setup, it MUST add at least one of the

possible correlation mechanisms, but MUST NOT specify values for those subfields.

For example, if the Offerer is willing to use the User-User Information element and DTMF digit sending mechanisms but can only become the passive party, and is also able to let the human user decide whether the correlation should be done or not, it includes the following lines in the SDP:

```
a=cs-correlation:uuie dtmf external
```

```
a=setup:passive
```

If, on the other hand, the Offerer is willing to use the User-User Information element and the DTMF correlation mechanisms and is able to become the active or passive side, and is also able to let the human user decide whether the correlation should be done or no, it includes the following lines in the SDP:

```
a=cs-correlation:uuie:56A390F3D2B7310023 dtmf:14D*3 external
```

```
a=setup:actpass
```

The negotiation of the value of the 'setup' attribute takes place as defined in Section 4.1 of RFC4145 [RFC4145].

The Offerer states which role or roles it is willing to perform; and the Answerer, taking the Offerer's willingness into consideration, chooses which roles both endpoints will actually perform during the circuit-switched bearer setup.

By 'active' endpoint, we refer to an endpoint that will establish the circuit-switched bearer; and by 'passive' endpoint, we refer to an endpoint that will receive a circuit-switched bearer.

If an Offerer does not know its international E.164 number, it MUST set the 'a=setup' attribute to the value 'active'. If the Offerer knows its international E.164 number, it SHOULD set the value to either 'actpass' or 'passive'.

Also 'holdconn' is a permissible value in the 'a=setup' attribute. It indicates that the connection should not be established for the time being.

The Offerer uses the "a=connection" attribute to decide whether a new circuit-switched bearer is to be established or not. For the initial Offer, the Offerer MUST use value 'new'.

5.6.2. Generating the Answer

If the Offer contained a circuit-switched audio or video stream, the Answerer first determines whether it is able to accept and use such streams on the circuit-switched network. If the Answerer does not support or is not willing to use circuit-switched media for the session, it **MUST** construct an Answer where the port number for such media stream(s) is set to zero, according to Section 6 of [RFC3264]. If the Answerer is willing to use circuit-switched media for the session, it **MUST** ignore the received port number (unless the port number is set to zero).

If the Offer included a "-" as the payload type number, it indicates that the Offerer is not willing or able to define any specific payload type. Most often, a "-" is expected to be used instead of the payload type when the endpoint is not aware of or not willing to define the codecs which will eventually be used on the circuit-switched bearer. The circuit-switched signaling protocols have their own means of negotiating or indicating the codecs, therefore an Answerer **SHOULD** accept such Offers, and **SHOULD** set the payload type to "-" also in the Answer.

If the Answerer explicitly wants to specify a codec for the circuit-switched media, it **MAY** set the respective payload numbers in the <fmt> subfield in the answer. This behavior, however, is **NOT RECOMMENDED**.

When receiving the Offer, the Answerer **MUST** determine whether it becomes the active or passive party.

If the SDP in the Offer indicates that the Offerer is only able to become the active party, the Answerer needs to determine whether it is able to become the passive party. If this is not possible e.g. due to the Answerer not knowing its international E.164 number, the Answerer **MUST** reject the circuit-switched media by setting the port number to zero on the Answer. If the Answerer is aware of its international E.164 number, it **MUST** include the "a=setup" attribute in the Answer and set it to value "passive" or "holdconn". The Answerer **MUST** also include its E.164 number in the "c=" line.

If the SDP in the Offer indicates that the Offerer is only able to become the passive party, the Answerer **MUST** verify that the Offerer's E.164 number is included in the "c=" line of the Offer. If the number is included, the Answerer **MUST** include the "a=setup" attribute in the Answer and set it to value "active" or "holdconn". If the number is not included, or the recipient of the Offer is not willing to establish a connection the E.164 based on a priori knowledge of cost, or other reasons, call establishment is not possible, and the

Answerer MUST reject the circuit-switched media by setting the port number to zero in the Answer.

If the SDP in the Offer indicates that the Offerer is able to become either the active or passive party, the Answerer determines which role it will take. If the Offer includes an international E.164 number in the "c=" line, the Answerer SHOULD become the active party. If the Answerer does not become the active party, and if the Answerer is aware of its E.164 number, it MUST become the passive party. If the Answerer does not become the active or the passive party, it MUST reject the circuit-switched media by setting the port number to zero in the Answer.

For each media description where the Offer includes a "a=cs-correlation" attribute, the Answerer MUST select from the Offer those correlation mechanisms it supports, and include in the Answer one "a=cs-correlation" attribute line containing those mechanisms it is willing to use. The Answerer MUST only add one "a=cs-correlation" attribute in those media descriptions where also the Offer included a "a=cs-correlation" attribute. The Answerer MUST NOT add any mechanisms which were not included in the offer. If there are more than one "cs-correlation" attributes per media description in the Offer, the Answerer MUST discard all but the first for any media description. Also, the Answerer MUST discard all unknown "cs-correlation" attribute values.

If the Answerer becomes the active party, it MUST add a value to any of the possible subfields.

If the Answerer becomes the passive party, it MUST NOT add any values to the subfields in the "cs-correlation" attribute.

After generating and sending the Answer, if the Answerer became the active party, it

- o MUST extract the E.164 number from the "c=" line of the Offer and MUST establish a circuit-switched bearer to that address.
- o if the SDP Answer contained a value for the "callerid" subfield, MUST set the Calling Party Number Information Element to that number,
- o if the SDP Answer contained a value for the "uuie" subfield, MUST send the User-User Information element according to the rules defined for the circuit-switched technology used, and set the value of the Information Element to that received in the SDP Offer,

- o if the SDP Answer contained a value for the "dtmf" subfield, MUST send those DTMF digits according to the circuit-switched technology used.

If, on the other hand, the Answerer became the passive party, it

- o MUST be prepared to receive a circuit-switched bearer,
- o if the Offer contained a value for the "callerid" subfield, MUST compare that value to the Calling Party Number Information Element of the circuit-switched bearer. If the received Calling Party Number Information Element matches the value of the "callerid" subfield, the call SHOULD be treated as correlated to the ongoing session.
- o if the Offer contained a value for the "dtmf" subfield, MUST be prepared to receive and collect DTMF digits once the circuit-switched bearer is set up. The Answerer MUST compare the received DTMF digits to the value of the "dtmf" subfield. If the received DTMF digits match the value of the "dtmf" subfield in the "cs-correlation" attribute, the call SHOULD be treated as correlated to the ongoing session.
- o if the Offer contained a value for the "uuie" subfield, MUST be prepared to receive a User-User Information element once the circuit-switched bearer is set up. The Answerer MUST compare the received UUI to the value of the "uuie" subfield. If the value of the received UUI matches the value of the "uuie" subfield, the call SHOULD be treated as correlated to the ongoing session.
- o if the Offer contained a "external" subfield, MUST be prepared to receive a circuit-switched call and use the external means (typically the human user) for accepting or rejecting the call.

If the Answerer becomes the active party, generates an SDP answer, and then it finds out that the circuit-switched call cannot be established, then the Answerer MUST create a new SDP offer where circuit-switched stream is removed from the session (actually, by setting the corresponding port in the m= line to zero) and send it to its counterpart. This is to synchronize both parties (and potential intermediaries) on the state of the session.

5.6.3. Offerer processing the Answer

When receiving the Answer, if the SDP does not contain "a=cs-correlation" attribute line, the Offerer should take that as an indication that the other party does not support or is not willing to

use the procedures defined in the document for this session, and MUST revert to normal processing of SDP.

When receiving the Answer, the Offerer MUST first determine whether it becomes the active or passive party, as described in Section 5.3.1.

If the Offerer becomes the active party, it

- o MUST extract the E.164 number from the "c=" line and MUST establish a circuit-switched bearer to that address.
- o if the SDP Answer contained a value for the "uuie" subfield, MUST send the User-User Information element according to the rules defined for the circuit-switched technology used, and set the value of the Information Element to that received in the SDP Answer,
- o if the SDP Answer contained a value for the "dtmf" subfield, MUST send those DTMF digits according to the circuit-switched technology used.

If the Offerer becomes the passive party, it

- o MUST be prepared to receive a circuit-switched bearer,
- o Note that if delivery of the Answer is delayed for some reason, the circuit-switched call attempt may arrive at the Offerer before the Answer has been processed. In this case, since the correlation mechanisms are negotiated as part of the Offer/Answer exchange, the Answerer cannot know whether or not the incoming circuit-switched call attempt is correlated with the session being negotiated, the Offerer SHOULD answer the circuit-switched call attempt only after it has received and processed the Answer.
- o If the Answer contained a value for the "dtmf" subfield, the Offerer MUST be prepared to receive and collect DTMF digits once the circuit-switched bearer is set up. The Offerer SHOULD compare the received DTMF digits to the value of the "dtmf" subfield. If the received DTMF digits match the value of the "dtmf" subfield in the "cs-correlation" attribute, the call SHOULD be treated as correlated to the ongoing session.
- o If the Answer contained a value for the "uuie" subfield, the Offerer MUST be prepared to receive a User-User Information element once the circuit-switched bearer is set up. The Offerer SHOULD compare the received UUI to the value of the "uuie" subfield. If the value of the received UUI matches the value of

the "uuie" subfield, the call SHOULD be treated as correlated to the ongoing session.

- o If the Answer contained a "external" subfield, the Offerer MUST be prepared to receive a circuit-switched call and use the external means (typically the human user) for accepting or rejecting the call.

According to the Offer/Answer Model with SDP [RFC3264], the Offerer needs to be ready to receive media as soon as the Offer has been sent. It may happen that the Answerer, if it became the active party, will initiate a circuit-switched bearer setup which will arrive at the Offerer before the Answer has arrived. However, the Offerer needs to receive the Answer and examine the information about the correlation mechanisms in order to successfully perform correlation of the circuit-switched call to the session. Therefore, if the Offerer receives an incoming circuit-switched call, it MUST NOT accept the call before the Answer has been received. If no Answer is received during an implementation specific time, the Offerer MUST either modify the session according to [RFC3264] or terminate it according to the session signaling procedures in question (for terminating a SIP session, see Section 15 of [RFC3261]).

5.6.4. Modifying the session

If, at a later time, one of the parties wishes to modify the session, e.g., by adding new media stream, or by changing properties used on an existing stream, it may do so via the mechanisms defined for an Offer/Answer Model with SDP [RFC3264].

If there is an existing circuit-switched bearer between the endpoints, and the Offerer wants to reuse that, the Offerer MUST set the value of the "a=connection" attribute to 'existing'.

If either party removes the circuit-switched media from the session (by setting the port number to zero), it MUST terminate the circuit-switched bearer using whatever mechanism is appropriate for the technology in question.

If either party wishes to drop and reestablish an existing call, that party MUST first remove the circuit-switched media from the session by setting the port number to zero, and then use another Offer/Answer exchange where it MUST set the "a=connection" attribute to 'new'. If the media types are different (for example, a different codec will be used for the circuit-switched bearer), the media descriptions for terminating the existing bearer and the new bearer can be in the same Offer.

If either party would like to remove existing RTP based media from the session and replace that with a circuit-switched bearer, it would create a new Offer to add the circuit-switched media as described in Section 5.6.1 above, replacing the RTP based media description by the circuit-switched media description, as specified in RFC 3264 [RFC3264].

Once the Offer/Answer exchange is done, but the circuit-switched bearer is not yet established, there may be a period of time when no media is available. Also, it may happen that correlating the circuit-switched call fails for reasons discussed in Section 5.3.3. In this case, even if the Offer/Answer exchange was successful, endpoints are not able to receive or send media. It is up to the implementation to decide the behavior in this case; if nothing else is done, the user most likely hangs up after a while if there was no other media in the session. Note that this may also happen when switching from RTP media to another RTP media (for example when firewall blocks the new media stream).

If either party would like to remove existing circuit-switched media from the session and replace that with RTP based media, it would modify the media description as per the procedures defined in RFC 3264 [RFC3264]. The endpoint MUST then terminate the circuit-switched bearer using whatever mechanism is appropriate for the technology in question.

5.7. Formal Syntax

The following is the formal Augmented Backus-Naur Form (ABNF) [RFC5234] syntax that supports the extensions defined in this specification. The syntax is built above the SDP [RFC4566] and the tel URI [RFC3966] grammars. Implementations according to this specification MUST be compliant with this syntax.

Figure 2 shows the formal syntax of the extensions defined in this memo.

```

; extension to the connection field originally specified
; in RFC4566

connection-field    = [%x63 "=" nettype SP addrtype SP
connection-address CRLF]
; CRLF defined in RFC5234

;nettype and addrtype are defined in RFC 4566

connection-address /= global-number-digits / "-"
; global-number-digits specified in RFC3966

;subrules for correlation attribute
attribute           /= cs-correlation-attr
; attribute defined in RFC4566
cs-correlation-attr = "cs-correlation:" corr-mechanisms
corr-mechanisms     = corr-mech *(SP corr-mech)
corr-mech           = caller-id-mech / uuie-mech /
                    dtmf-mech / external-mech /
                    ext-mech
caller-id-mech      = "callerid" [":" caller-id-value]
caller-id-value     = "+" 1*15DIGIT
; DIGIT defined in RFC5234
uuie-mech           = "uuie" [":" uuie-value]
uuie-value          = 1*65(HEXDIG HEXDIG)
                    ;This represents up to 130 HEXDIG
                    ; (65 octets)
                    ;HEXDIG defined in RFC5234
                    ;HEXDIG defined as 0-9, A-F
dtmf-mech           = "dtmf" [":" dtmf-value]
dtmf-value          = 1*32(DIGIT / %x41-44 / %x23 / %x2A )
                    ;0-9, A-D, '#' and '*'
external-mech       = "external"
ext-mech            = ext-mech-name [":" ext-mech-value]
ext-mech-name       = token
ext-mech-value      = token
; token is specified in RFC4566

```

Figure 2: Syntax of the SDP extensions

6. Examples

In the examples below, where an SDP line is too long to be displayed as a single line, a breaking character "\" indicates continuation in the following line. Note that this character is included for display purposes only. Implementations MUST write a single line without breaks.

6.1. Single PSTN audio stream

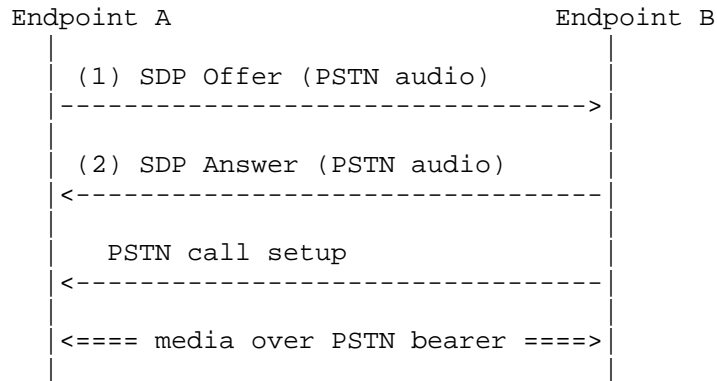


Figure 3: Basic flow

Figure 3 shows a basic example that describes a single audio media stream over a circuit-switched bearer. Endpoint A generates a SDP Offer which is shown in Figure 4. The Offer describes a PSTN circuit-switched bearer in the "m=" and "c=" line where it also indicates its international E.164 number format. Additionally, Endpoint A expresses that it can initiate the circuit-switched bearer or be the recipient of it in the "a=setup" attribute line. The SDP Offer also includes correlation identifiers that this endpoint will insert in the Calling Party Number and/or User-User Information Element of the PSTN call setup if eventually this endpoint initiates the PSTN call. Endpoint A also includes the "external" as one correlation mechanism indicating that it can use the human user to perform correlation in case other mechanisms fail.

```

v=0
o=alice 2890844526 2890842807 IN IP4 192.0.2.5
s=
t=0 0
m=audio 9 PSTN -
c=PSTN E164 +441134960123
a=setup:actpass
a=connection:new
a=cs-correlation:callerid:+441134960123 \
  uuie:56A390F3D2B7310023 external
  
```

Figure 4: SDP offer (1)

Endpoint B generates a SDP Answer (Figure 5), describing a PSTN audio media on port 9 without information on the media sub-type on the "m=" line. The "c=" line contains B's international E.164 number. In the

"a=setup" line Endpoint B indicates that it is willing to become the active endpoint when establishing the PSTN call, and it also includes the "a=cs-correlation" attribute line containing the values it is going to include in the Calling Party Number and User-User IE of the PSTN call establishment. Endpoint B is also able to perform correlation by external means, in case other correlation mechanisms fail.

```
v=0
o=- 2890973824 2890987289 IN IP4 192.0.2.7
s=
t=0 0
m=audio 9 PSTN -
c=PSTN E164 +441134960124
a=setup:active
a=connection:new
a=cs-correlation:callerid:+441134960124 \
    uuie:74B9027A869D7966A2 external
```

Figure 5: SDP Answer with circuit-switched media

When Endpoint A receives the Answer, it examines that B is willing to become the active endpoint when setting up the PSTN call. Endpoint A temporarily stores B's E.164 number and the User-User IE value of the "cs-correlation" attribute, and waits for a circuit-switched bearer establishment.

Endpoint B initiates a circuit-switched bearer using whatever circuit-switched technology is available for it. The called party number is set to A's number, and calling party number is set to B's own number. Endpoint B also sets the User-User Information Element value to the one contained in the SDP Answer.

When Endpoint A receives the circuit-switched bearer establishment, it examines the UUIE and the calling party number, and by comparing those received during O/A exchange determines that the call is related to the SDP session.

It may also be that neither the UUIE nor the calling party number is received by the called party, or the format of the calling party number is changed by the PSTN. Implementations may still accept such call establishment attempts as being related to the session that was established in the IP network. As it cannot be guaranteed that the values used for correlation are always passed intact through the network, they should be treated as additional hints that the circuit-switched bearer is actually related to the session.

6.2. Advanced SDP example: Circuit-Switched Audio and Video Streams

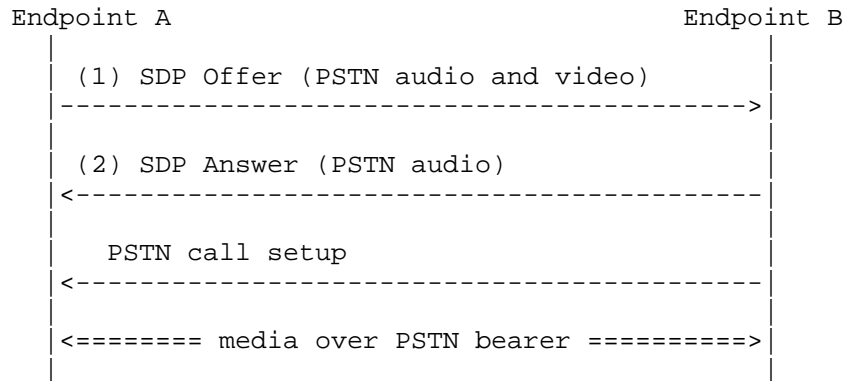


Figure 6: Circuit-Switched Audio and Video streams

Figure 6 shows an example of negotiating audio and video media streams over circuit-switched bearers.

```

v=0
o=alice 2890844526 2890842807 IN IP4 192.0.2.5
s=
t=0 0
a=setup:actpass
a=connection:new
c=PSTN E164 +441134960123
m=audio 9 PSTN -
a=cs-correlation:dtmf:1234536
m=video 9 PSTN 34
a=rtpmap:34 H263/90000
a=cs-correlation:callerid:+441134960123
  
```

Figure 7: SDP offer with circuit-switched audio and video (1)

Upon receiving the SDP offer described in Figure 7, Endpoint B rejects the video stream as the device does not currently support video, but accepts the circuit-switched audio stream. As Endpoint A indicated that it is able to become either the active or passive party, Endpoint B gets to select which role it would like to take. Since the Offer contained the international E.164 number of Endpoint A, Endpoint B decides that it becomes the active party in setting up the circuit-switched bearer. B includes a new value in the "dtmf" subfield of the "cs-correlation" attribute, which it is going to send as DTMF tones once the bearer setup is complete. The Answer is described in Figure 8

```
v=0
o=- 2890973824 2890987289 IN IP4 192.0.2.7
s=
t=0 0
a=setup:active
a=connection:new
c=PSTN E164 +441134960124
m=audio 9 PSTN -
a=cs-correlation:dtmf:654321
m=video 0 PSTN 34
a=cs-correlation:callerid:+441134960124
```

Figure 8: SDP answer with circuit-switched audio and video (2)

7. Security Considerations

This document provides an extension on top of RFC 4566 [RFC4566], and RFC 3264 [RFC3264]. As such, the security considerations of those documents apply.

This memo provides mechanisms to agree on a correlation identifier or identifiers that are used to evaluate whether an incoming circuit-switched bearer is related to an ongoing session in the IP domain. If an attacker replicates the correlation identifier and establishes a call within the time window the receiving endpoint is expecting a call, the attacker may be able to hijack the circuit-switched bearer. These types of attacks are not specific to the mechanisms presented in this memo. For example, caller ID spoofing is a well-known attack in the PSTN. Users are advised to use the same caution before revealing sensitive information as they would on any other phone call. Furthermore, users are advised that mechanisms that may be in use in the IP domain for securing the media, like Secure RTP (SRTP) [RFC3711], are not available in the CS domain.

For the purposes of establishing a circuit-switched bearer, the active endpoint needs to know the passive endpoint's phone number. Phone numbers are sensitive information, and some people may choose not to reveal their phone numbers when calling using supplementary services like Calling Line Identification Restriction (CLIR) in GSM. Implementations should take the caller's preferences regarding calling line identification into account if possible, by restricting the inclusion of the phone number in SDP "c=" line if the caller has chosen to use CLIR. If this is not possible, implementations may present a prompt informing the user that their phone number may be transmitted to the other party.

Similarly as with IP addresses, if there is a desire to protect the SDP containing phone numbers carried in SIP, implementers are advised to follow the security mechanisms defined in [RFC3261].

It is possible that an attacker creates a circuit-switched session whereby the attacked endpoint should dial a circuit-switched number, perhaps even a premium-rate telephone number. To mitigate the consequences of this attack, endpoints **MUST** authenticate and trust remote endpoints users who try to remain passive in the circuit-switched connection establishment. It is **RECOMMENDED** that endpoints have local policies precluding the active establishment of circuit switched connections to certain numbers (e.g., international, premium, long distance). Additionally, it is strongly **RECOMMENDED** that the end user is asked for consent prior to the endpoint initiating a circuit-switched connection.

8. IANA Considerations

This document instructs IANA to register a number of SDP tokens according to the following data.

8.1. Registration of new cs-correlation SDP attribute

Contact: Miguel Garcia <miguel.a.garcia@ericsson.com>

Attribute name: cs-correlation

Long-form attribute name: PSTN Correlation Identifier

Type of attribute: media level only

Subject to charset: No

Description: This attribute provides the Correlation Identifier used in PSTN signaling

Appropriate values: see Section 5.2.3.1

Specification: RFC XXXX

The IANA is requested to create a subregistry for 'cs-correlation' attribute under the Session Description Protocol (SDP) Parameters registry. The initial values for the subregistry are presented in the following, and IANA is requested to add them into its database:

Value of 'cs-correlation' attribute	Reference	Description
-----	-----	-----
callerid	RFC XXXX	Caller ID
uuie	RFC XXXX	User-User Information Element
dtmf	RFC XXXX	Dual-tone Multi-Frequency
external	RFC XXXX	External

Note for the RFC Editor: 'RFC XXXX' above should be replaced by a reference to the RFC number of this draft.

As per the terminology in [RFC5226], the registration policy for new values of 'cs-correlation' parameter is 'Specification Required'.

8.2. Registration of a new "nettype" value

This memo provides instructions to IANA to register a new "nettype" in the Session Description Protocol Parameters registry [1]. The registration data, according to RFC 4566 [RFC4566] follows.

Type	SDP Name	Reference
----	-----	-----
nettype	PSTN	[RFCxxxx]

8.3. Registration of new "addrtype" value

This memo provides instructions to IANA to register a new "addrtype" in the Session Description Protocol Parameters registry [2]. The registration data, according to RFC 4566 [RFC4566] follows.

Type	SDP Name	Reference
----	-----	-----
addrtype	E164	[RFCxxxx]

Note: RFC XXXX defines the "E164" addrtype in the context of the "PSTN" nettype only. Please refer to the relevant RFC for a description of that representation.

8.4. Registration of a new "proto" value

This memo provides instructions to IANA to register a new "proto" in the Session Description Protocol Parameters registry [3]. The registration data, according to RFC 4566 [RFC4566] follows.

Type	SDP Name	Reference
-----	-----	-----
proto	PSTN	[RFCxxxx]

The related "fmt" namespace re-uses the conventions and payload type number defined for RTP/AVP. In RFC XXXX, the RTP audio and video media types, when applied to PSTN circuit-switched bearers, represent merely an audio or video codec in its native format directly on top of a single PSTN bearer.

In some cases, the endpoint is not able to determine the list of available codecs for circuit-switched media streams. In this case, in order to be syntactically compliant with SDP [RFC4566], the endpoint MUST include a single dash ("-") in the <fmt> subfield.

9. Acknowledgments

The authors want to thank Paul Kyzivat, Flemming Andreassen, Thomas Belling, John Elwell, Jari Mutikainen, Miikka Poikselka, Jonathan Rosenberg, Ingemar Johansson, Christer Holmberg, Alf Heidermark, Tom Taylor, Thomas Belling, Keith Drage, and Andrew Allen for providing their insight and comments on this document.

10. References

10.1. Normative References

- [ITU.Q931.1998] "Digital Subscriber Signalling System No. 1 (DSS 1) - ISDN User - Network Interface Layer 3 Specification for Basic Call Control", ISO Standard 9594-1, May 1998.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.
- [RFC3966] Schulzrinne, H., "The tel URI for Telephone Numbers", RFC 3966, December 2004.
- [RFC4145] Yon, D. and G. Camarillo, "TCP-Based Media Transport in the Session Description Protocol (SDP)", RFC 4145, September 2005.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.

- [RFC4648] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", RFC 4648, October 2006.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

10.2. Informative References

- [I-D.ietf-cuss-sip-uu] Johnston, A. and J. Rafferty, "A Mechanism for Transporting User to User Call Control Information in SIP", draft-ietf-cuss-sip-uu-12 (work in progress), January 2014.
- [ITU.E164.1991] International Telecommunications Union, "The International Public Telecommunication Numbering Plan", ITU-T Recommendation E.164, 1991.
- [ITU.Q23.1988] International Telecommunications Union, "Technical features of push-button telephone sets", ITU-T Technical Recommendation Q.23, 1988.
- [RFC3108] Kumar, R. and M. Mostafa, "Conventions for the use of the Session Description Protocol (SDP) for ATM Bearer Connections", RFC 3108, May 2001.
- [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, June 2002.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.
- [RFC3551] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, July 2003.
- [RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.

- [RFC3725] Rosenberg, J., Peterson, J., Schulzrinne, H., and G. Camarillo, "Best Current Practices for Third Party Call Control (3pcc) in the Session Initiation Protocol (SIP)", BCP 85, RFC 3725, April 2004.
- [RFC4975] Campbell, B., Mahy, R., and C. Jennings, "The Message Session Relay Protocol (MSRP)", RFC 4975, September 2007.
- [TS.24.008] 3GPP, "Mobile radio interface Layer 3 specification; Core network protocols; Stage 3", 3GPP TS 24.008 3.20.0, December 2005.

10.3. URIs

- [1] <http://www.iana.org/assignments/sdp-parameters>
- [2] <http://www.iana.org/assignments/sdp-parameters>
- [3] <http://www.iana.org/assignments/sdp-parameters>

Authors' Addresses

Miguel A. Garcia-Martin
Ericsson
Calle Via de los Poblados 13
Madrid, ES 28033
Spain

Email: miguel.a.garcia@ericsson.com

Simo Veikkolainen
Nokia
P.O. Box 226
NOKIA GROUP, FI 00045
Finland

Phone: +358 50 486 4463
Email: simo.veikkolainen@nokia.com

MMUSIC
Internet-Draft
Intended status: Standards Track
Expires: May 3, 2012

R. Gilman
Independent
R. Even
Gesher Erove Ltd
F. Andreassen
Cisco Systems
October 31, 2011

SDP Media Mapabilities Negotiation
draft-ietf-mmusic-sdp-media-capabilities-12

Abstract

Session Description Protocol (SDP) capability negotiation provides a general framework for indicating and negotiating capabilities in SDP. The base framework defines only capabilities for negotiating transport protocols and attributes. In this document, we extend the framework by defining media capabilities that can be used to negotiate media types and their associated parameters.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on May 3, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect

to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	4
2. Terminology	5
3. SDP Media Capabilities	6
3.1. Requirements	6
3.2. Solution Overview	7
3.3. New Capability Attributes	13
3.3.1. The Media Format Capability Attributes	13
3.3.2. The Media Format Parameter Capability Attribute	15
3.3.3. The Media-Specific Capability Attribute	18
3.3.4. New Configuration Parameters	20
3.3.5. The Latent Configuration Attribute	21
3.3.6. Enhanced Potential Configuration Attribute	24
3.3.7. Substitution of Media Payload Type Numbers in Capability Attribute Parameters	27
3.3.8. The Session Capability Attribute	28
3.4. Offer/Answer Model Extensions	32
3.4.1. Generating the Initial Offer	32
3.4.2. Generating the Answer	33
3.4.3. Offerer Processing of the Answer	33
3.4.4. Modifying the Session	34
4. Examples	35
4.1. Alternative Codecs	35
4.2. Alternative Combinations of Codecs (Session Configurations)	38
4.3. Latent Media Streams	38
5. IANA Considerations	41
5.1. New SDP Attributes	41
5.2. New SDP Option Tag	42
5.3. New SDP Capability Negotiation Parameters	42
6. Security Considerations	43
7. Changes from previous versions	44
7.1. Changes from version 11	44
7.2. Changes from version 10	44
7.3. Changes from version 09	44
7.4. Changes from version 08	45
7.5. Changes from version 04	45
7.6. Changes from version 03	45
7.7. Changes from version 02	46
7.8. Changes from version 01	46
7.9. Changes from version 00	47
8. Acknowledgements	48
9. References	49
9.1. Normative References	49
9.2. Informative References	49
Authors' Addresses	50

1. Introduction

Session Description Protocol (SDP) capability negotiation [RFC5939] provides a general framework for indicating and negotiating capabilities in SDP[RFC4566]. The base framework defines only capabilities for negotiating transport protocols and attributes.

The [RFC5939] document lists some of the issues with the current SDP capability negotiation process. An additional real life case is to be able to offer one media stream (e.g. audio) but list the capability to support another media stream (e.g. video) without actually offering it concurrently.

In this document, we extend the framework by defining media capabilities that can be used to indicate and negotiate media types and their associated format parameters. This document also adds the ability to declare support for media streams, the use of which can be offered and negotiated later, and the ability to specify session configurations as combinations of media stream configurations. The definitions of new attributes for media capability negotiation are chosen to make the translation from these attributes to "conventional" SDP [RFC4566] media attributes as straightforward as possible in order to simplify implementation. This goal is intended to reduce processing in two ways: each proposed configuration in an offer may be easily translated into a conventional SDP media stream record for processing by the receiver; and the construction of an answer based on a selected proposed configuration is straightforward.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC2119 [RFC2119] and indicate requirement levels for compliant RTP implementations.

"Base Attributes": Conventional SDP attributes appearing in the base configuration of a media block.

"Base Configuration": The media configuration represented by a media block exclusive of all the capability negotiation attributes defined in this document, the base capability negotiation document [RFC5939], or any other capability negotiation document. In an offer SDP, the base configuration corresponds to the actual configuration as defined in [RFC5939].

"Conventional Attribute": Any SDP attribute other than those defined by the series of capability negotiation specifications.

"Conventional SDP": An SDP record devoid of capability negotiation attributes.

"Media Capability": A media encoding, typically a media subtype such as PCMU, H263-1998, or T38.

3. SDP Media Capabilities

The SDP capability negotiation [RFC5939] discusses the use of any SDP [RFC4566] attribute (a=) under the attribute capability "acap". The limitations of using acap for ftmp and rtpmap in a potential configuration are described in [RFC5939]; for example they can be used only at the media level since they are media level attributes. The [RFC5939] does not provide a way to exchange media-level capabilities prior to the actual offer of the associated media stream. This section provides an overview of extensions providing an SDP Media Capability negotiation solution offering more robust capabilities negotiation. This is followed by definitions of new SDP attributes for the solution and its associated updated offer/answer procedures [RFC3264]

3.1. Requirements

The capability negotiation extensions requirements considered herein are as follows.

- REQ-01: Support the specification of alternative (combinations of) media formats (codecs) in a single media block.
- REQ-02: Support the specification of alternative media format parameters for each media format.
- REQ-03: Retain backward compatibility with conventional SDP. Ensure that each and every offered configuration can be easily translated into a corresponding SDP media block expressed with conventional SDP lines.
- REQ-04: Ensure the scheme operates within the offer/answer model in such a way that media formats and parameters can be agreed upon with a single exchange.
- REQ-05: Provide the ability to express offers in such a way that the offerer can receive media as soon as the offer is sent. (Note that the offerer may not be able to render received media prior to exchange of keying material.)
- REQ-06: Provide the ability to offer latent media configurations for future negotiation.
- REQ-07: Provide reasonable efficiency in the expression of alternative media formats and/or format parameters, especially in those cases in which many combinations of options are offered.

REQ-08: Retain the extensibility of the base capability negotiation mechanism.

REQ-09: Provide the ability to specify acceptable combinations of media streams and media formats. For example, offer a PCMU audio stream with an H264 video stream, or a G729 audio stream with an H263 video stream. This ability would give the offerer a means to limit processing requirements for simultaneous streams. This would also permit an offer to include the choice of an audio/T38 stream or an image/T38 stream, but not both.

Other possible extensions have been discussed, but have not been treated in this document. They may be considered in the future. Three such extensions are:

FUT-01: Provide the ability to mix, or change, media types within a single media block. Conventional SDP does not support this capability explicitly; the usual technique is to define a media subtype that represents the actual format within the nominal media type. For example, T.38 FAX as an alternative to audio/PCMU within an audio stream is identified as audio/T38; a separate FAX stream would use image/T38.

FUT-02: Provide the ability to support multiple transport protocols within an active media stream without reconfiguration. This is not explicitly supported by conventional SDP.

FUT-03: Provide capability negotiation attributes for all media-level SDP line types in the same manner as already done for the attribute type, with the exception of the media line type itself. The media line type is handled in a special way to permit compact expression of media coding/format options. The line types are bandwidth ("b="), information ("i="), connection data ("c="), and, possibly, the deprecated encryption key ("k=").

3.2. Solution Overview

The solution consists of new capability attributes corresponding to conventional SDP line types, new parameters for the pcfg, acfg, and the new lcfg attributes extending the base attributes from [RFC5939], and a use of the pcfg attribute to return capability information in the SDP answer.

Several new attributes are defined in a manner that can be related to the capabilities specified in a media line, and its corresponding rtpmap and fmp4 attributes.

- o A new media attribute ("a=rmcap") defines RTP-based media capabilities in the form of a media subtype (e.g. "PCMU"), and its encoding parameters (e.g. "/8000/2"). Each resulting media format type/subtype capability has an associated handle called a media capability number. The encoding parameters are as specified for the rtpmap attribute defined in [RFC4566], without the payload type number part.
- o A new media attribute ("a=omcap") defines other (non RTP-based) media capabilities in the form of a media subtype only (e.g. "T38"). Each resulting media format type/subtype capability has an associated handle called a media capability number.
- o A new attribute ("a=mfcap") specifies media format parameters associated with one or more media capabilities. The mfcap attribute is used primarily to associate the formatting capabilities normally carried in the fmtp attribute. Note that media format parameters can be used with RTP and non-RTP based media formats.
- o A new attribute ("a=mscap") that specifies media parameters associated with one or more media capabilities. The mscap attribute is used to associate capabilities with attributes other than fmtp or rtpmap, for example, the rtcp-fb attribute defined in [RFC4585].
- o A new attribute ("a=lcfg") specifies latent media stream configurations when no corresponding media line ("m=") is offered. An example is the offer of latent configurations for video even though no video is currently offered. If the peer indicates support for one or more offered latent configurations, the corresponding media stream(s) may be added via a new offer/answer exchange.
- o A new attribute ("a=sescap") is used to specify an acceptable combination of simultaneous media streams and their configurations as a list of potential and/or latent configurations.

New parameters are defined for the potential configuration (pcfg), latent configuration (lcfg), and accepted configuration (acfg) attributes to associate the new attributes with particular configurations.

- o A new parameter type ("m=") is added to the potential configuration ("a=pcfg:") attribute and the actual configuration ("a=acfg:") attribute defined in [RFC5939], and to the new latent configuration ("a=lcfg:") attribute. This permits specification of media capabilities (including their associated parameters) and

combinations thereof for the configuration. For example, the "a=pcfg:" line might specify PCMU and telephone events [RFC4733] or G.729B and telephone events as acceptable configurations. The "a=acfg:" line in the answer would specify the configuration chosen.

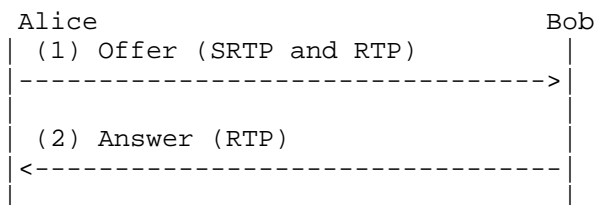
- o A new parameter type ("pt=") is added to the potential configuration, actual configuration, and latent configuration attributes. This parameter associates RTP payload type numbers with the referenced RTP-based media capabilities, and is appropriate only when the transport protocol uses RTP.
- o A new parameter type ("mt=") is used to specify the media type for latent configurations.

Special processing rules are defined for capability attribute arguments in order to reduce the need to replicate essentially-identical attribute lines for the base configuration and potential configurations.

- o A substitution rule is defined for any capability attribute to permit the replacement of the (escaped) media capability number with the media format identifier (e.g., the payload type number in audio/video profiles).
- o Replacement rules are defined for the conventional SDP equivalents of the mfcap and mscap capability attributes. This reduces the necessity to use the deletion qualifier in the a=pcfg parameter in order to ignore rtpmap, fmtp, and certain other attributes in the base configuration.
- o An argument concatenation rule is defined for mfcap attributes which refer to the same media capability number. This makes it convenient to combine format options concisely by associating multiple mfcap lines with multiple media capabilities.

This document extends the base protocol extensions to the offer/answer model that allow for capabilities and potential configurations to be included in an offer. Media capabilities constitute capabilities that can be used in potential and latent configurations. Whereas potential configurations constitute alternative offers that may be accepted by the answerer instead of the actual configuration(s) included in the "m=" line(s) and associated parameters, latent configurations merely inform the other side of possible configurations supported by the entity. Those latent configurations may be used to guide subsequent offer/answer exchanges, but they are not part of the current offer/answer exchange.

The mechanism is illustrated by the offer/answer exchange below, where Alice sends an offer to Bob:



Alice's offer includes RTP and SRTP as alternatives. RTP is the default, but SRTP is the preferred one (long lines are folded to fit the margins):

```

v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 3456 RTP/AVP 0 18
a=tcap:1 RTP/SAVP RTP/AVP
a=rtpmap:0 PCMU/8000/1
a=rtpmap:18 G729/8000/1
a=fmtp:18 annexb=yes
a=rmcap:1,4 g729/8000/1
a=rmcap:2 PCMU/8000/1
a=rmcap:5 telephone-event/8000
a=mfcap:1 annexb=no
a=mfcap:4 annexb=yes
a=mfcap:5 0-11
a=acap:1 crypto:1 AES_CM_128_HMAC_SHA1_32 \
inline:NzB4dlBINUAvLEw6UzF3WSJ+PSdFcGdUJShpX1Zj|2^20|1:32
a=pcfg:1 m=4,5|1,5 t=1 a=1 pt=1:100,4:101,5:102
a=pcfg:2 m=2 t=1 a=1 pt=2:103
a=pcfg:3 m=4 t=2 pt=4:18
  
```

The required base and extensions are provided by the "a=creq" attribute defined in [RFC5939], with the option tag "med-v0", which indicates that the extension framework defined here, must be supported. The Base level support is implied since it is required for the extensions.

The "m=" line indicates that Alice is offering to use plain RTP with PCMU or G.729B. The media line implicitly defines the default transport protocol (RTP/AVP in this case) and the default actual

configuration.

The "a=tcap:1" line, specified in the base protocol, defines transport protocol capabilities, in this case Secure RTP (SAVP profile) as the first option and RTP (AVP profile) as the second option.

The "a=rmcap:1,4" line defines two G.729 RTP-based media format capabilities, numbered 1 and 4, and their encoding rate. The capabilities are of media type "audio" and subtype G729. Note that the media subtype is explicitly specified here, rather than RTP payload type numbers. This permits the assignment of payload type numbers in the media stream configuration specification. In this example, two G.729 subtype capabilities are defined. This permits the declaration of two sets of formatting parameters for G.729.

The "a=rmcap:2" line defines a G.711 mu-law capability, numbered 2.

The "a=rmcap:5" line defines an audio telephone-event capability, numbered 5.

The "a=mfcap:1" line specifies the fmp format parameters for capability 1 (offerer will not accept G.729 Annex B packets).

The "a=mfcap:4" line specifies the fmp format parameters for capability 4 (offerer will accept G.729 Annex B packets).

The "a=mfcap:5" line specifies the fmp format parameters for capability 5 (the DTMF touchtones 0-9,*,#).

The "a=acap:1" line specified in the base protocol provides the "crypto" attribute which provides the keying material for SRTP using SDP security descriptions.

The "a=pcfg:" attributes provide the potential configurations included in the offer by reference to the media capabilities, transport capabilities, attribute capabilities and specified payload type number mappings. Three explicit alternatives are provided; the lowest-numbered one is the preferred one. The "a=pcfg:1 ..." line specifies media capabilities 4 and 5, i.e., G.729B and DTMF, or media capability 1 and 5, i.e., G.729 and DTMF. Furthermore, it specifies transport protocol capability 1 (i.e. the RTP/SAVP profile - secure RTP), and the attribute capability 1, i.e. the crypto attribute provided. Lastly, it specifies a payload type number mapping for (RTP-based) media capabilities 1, 4, and 5, thereby permitting the offerer to distinguish between encrypted media and unencrypted media received prior to receipt of the answer.

Use of unique payload type numbers in alternative configurations is not required; codecs such as AMR-WB [RFC4867] have the potential for so many combinations of options that it may be impractical to define unique payload type numbers for all supported combinations. If unique payload type numbers cannot be specified, then the offerer will be obliged to wait for the SDP answer before rendering received media. For SRTP using SDES inline keying [RFC4568], the offerer will still need to receive the answer before being able to decrypt the stream.

The second alternative ("a=pcfg:2 ...") specifies media capability 2, i.e. PCMU, under the RTP/SAVP profile, with the same SRTP key material.

The third alternative ("a=pcfg:3 ...") offers G.729B unsecured; its only purpose in this example is to show a preference for G.729B over PCMU.

The media line, with any qualifying attributes such as fmtp or rtpmap, is itself considered a valid configuration; it is assumed to be the lowest preference.

Bob receives the SDP offer from Alice. Bob supports G.729B, PCMU, and telephone events over RTP, but not SRTP, hence he accepts the potential configuration 3 for RTP provided by Alice. Bob generates the following answer:

```
v=0
o=- 24351 621814 IN IP4 192.0.2.2
s=
c=IN IP4 19x2.0.2.2
t=0 0
a=csup:med-v0
m=audio 4567 RTP/AVP 18
a=rtpmap:18 G729/8000
a=fmtp:18 annexb=yes
a=acfg:3 m=4 t=2 pt=4:18
```

Bob includes the "a=csup" and "a=acfg" attributes in the answer to inform Alice that he can support the med-v0 level of capability negotiations. Note that in this particular example, the answerer supported the capability extensions defined here, however had he not, he would simply have processed the offer based on the offered PCMU and G.729 codecs under the RTP/AVP profile only. Consequently, the answer would have omitted the "a=csup" attribute line and chosen one or both of the PCMU and G.729 codecs instead. The answer carries the accepted configuration in the "m=" line along with corresponding rtpmap and/or fmtp parameters, as appropriate.

Note that per the base protocol, after the above, Alice MAY generate a new offer with an actual configuration ("m=" line, etc.) corresponding to the actual configuration referenced in Bob's answer (not shown here).

3.3. New Capability Attributes

In this section, we present the new attributes associated with indicating the media capabilities for use by the SDP Capability negotiation. The approach taken is to keep things similar to the existing media capabilities defined by the existing media descriptions ("m=" lines) and the associated "rtpmap" and "fmtp" attributes. We use media subtypes and "media capability numbers" to link the relevant media capability parameters. This permits the capabilities to be defined at the session level and be used for multiple streams, if desired. For RTP-based media formats, payload types are then specified at the media level (see Section 3.3.4.2).

A media capability merely indicates possible support for the media type and media format(s) in question. In order to actually use a media capability in an offer/answer exchange, it MUST be referenced in a potential configuration.

Media capabilities can be provided at the session-level and/or the media-level. Media capabilities provided at the session level may be referenced in any pcfg or lcfc attribute at the media level (consistent with the media type), whereas media capabilities provided at the media level may be referenced only by the pcfg or lcfc attribute within that media stream only. In either case, the scope of the <med-cap-num> is the entire session description. This enables each media capability to be uniquely referenced across the entire session description (e.g. in a potential configuration).

3.3.1. The Media Format Capability Attributes

Media subtypes can be expressed as media format capabilities by use of the "a=rmcap" and "a=omcap" attributes. The "a=rmcap" attribute MUST be used for RTP-based media whereas the "a=omcap" attribute MUST be used for non-RTP-based (other) media formats. The two attributes are defined as follows:

```
a=rmcap:<media-cap-num-list> <encoding-name>/<clock-rate>
                               [/<encoding-parms>]
```

```
a=omcap:<media-cap-num-list> <format-name>
```

where <media-cap-num-list> is a (list of) media capability number(s) used to number a media format capability, the <encoding name> is the

media subtype e.g. H263-1998 or PCMU, <clock rate> is the encoding rate, and <encoding parms> are the media encoding parameters for the media subtype;. All media format capabilities in the list are assigned to the same media type/subtype. Each occurrence of the rmcap and omcap attribute MUST use unique values in their <media-cap-num-list>; the media capability numbers are shared between the two attributes and the numbers MUST be unique across the entire SDP session. In short, the rmcap and omcap attributes define media capabilities and associates them with a media capability number in the same manner as the rtpmap attribute defines them and associates them with a payload type number. Additionally, the attributes allow multiple capability numbers to be defined for the media format in question. This permits the media format to be associated with different media parameters in different configurations.

In ABNF, we have:

```
media-capability-line = rtp-mcap / non-rtp-mcap

rtp-mcap               = "a=rmcap:" media-cap-num-list
                        1*WSP encoding-name "/" clock-rate
                        ["/" encoding-parms]
non-rtp-mcap           = "a=omcap:" media-cap-num-list 1*WSP format-name
media-cap-num-list     = media-cap-num-element
                        *["," media-cap-num-element]
media-cap-num-element  = media-cap-num
                        / media-cap-num-range
media-cap-num-range    = media-cap-num "-" media-cap-num
media-cap-num          = 1*10(DIGIT)
encoding-name          = token ; defined in RFC4566
clock-rate             = 1*10(DIGIT)
encoding-parms         = token
format-name            = token ;defined in RFC4566
```

The encoding-name, clock-rate and encoding-parms are as defined to appear in an rtpmap attribute for each media type/subtype. Thus, it is easy to convert an rmcap attribute line into one or more rtpmap attribute lines, once a payload type number is assigned to a media-cap-num (see Section 3.3.5).

The format-name is a media format description for non-RTP based media as defined for the <fmt> part of the media description ("m=" line) in [RFC4566]. In simple terms, it's the name of the media format, e.g. "t38".

The "rmcap" and "omcap" attributes can be provided at the session-level and/or the media-level. There can be more than one rmcap plus one omcap attribute at the session or media level (i.e at most one of

each at the session-level and at most one of in each media description). Each media-cap-num MUST be unique within the entire SDP record; it is used to identify that media capability in potential, latent and actual configurations, and in other attribute lines as explained below. Note that the media-cap-num values are shared between the rmcap and omcap attributes, and hence the uniqueness requirement applies to the union of them. When the media capabilities are used in a potential, latent or actual configuration, the media formats referred by those configurations apply at the media level, irrespective of whether the media capabilities themselves were specified at the session or media level. In other words, the media capability applies to the specific media description associated with the configuration which invokes it.

For example:

```
v=0
a=rmcap:1 L16/8000/1
a=rmcap:2 L16/16000/2
a=rmcap:3,4 H263-1998/90000
m=audio 54320 RTP/AVP 0
a=pcfg:1 m=1|2, pt=1:99,2:98
m=video 66544 RTP/AVP 100
a=rtpmap:100 H264/90000
a=pcfg:10 m=3 pt=3:101
```

3.3.2. The Media Format Parameter Capability Attribute

This attribute is used to associate media format specific format parameters with one or more media capabilities. The form of the attribute is:

```
a=mfcap:<media-caps> <list of parameters>
```

where <media-caps> permits the list of parameters to be associated with one or more media capabilities and the format parameters are specific to the type of media format. The mfcap lines map to a single traditional SDP fmt parameter attribute line (one for each entry in <media-caps>) of the form

```
a=fmtp:<fmt> <list of parameters>
```

where <fmt> is the media format description defined in RFC 4566 [RFC4566], as appropriate for the particular media stream. The mfcap attribute MUST be used to encode attributes for media capabilities, which would conventionally appear in an fmtp attribute. The existing acap attribute MUST NOT be used to encode fmtp attributes.

The mfcap attribute adheres to [RFC4566] attribute production rules with

```
media-format-capability = "a=mfcap:" media-cap-num-list 1*WSP
                               fmt-specific-param-list
fmt-specific-param-list = text ; defined in RFC4566
```

Note that media format parameters can be used with RTP-based and non-RTP based media formats.

3.3.2.1. Media Format Parameter Concatenation Rule

The appearance of media subtypes with a large number of formatting options (e.g., AMR-WB [RFC4867]) coupled with the restriction that only a single fmt attribute can appear per media format, suggests that it is useful to create a combining rule for mfcap parameters which are associated with the same media capability number. Therefore, different mfcap lines MAY include the same media-cap-num in their media-cap-num-list. When a particular media capability is selected for processing, the parameters from each mfcap line which references the particular capability number in its media-cap-num-list are concatenated together via ";", in the order the mfcap attributes appear in the SDP record, to form the equivalent of a single fmt attribute line. This permits one to define a separate mfcap line for a single parameter and value that is to be applied to each media capability designated in the media-cap-num-list. This provides a compact method to specify multiple combinations of format parameters when using codecs with multiple format options. Note that order-dependent parameters SHOULD be placed in a single mfcap line to avoid possible problems with line rearrangement by a middlebox.

Format parameters are not parsed by SDP; their content is specific to the media type/subtype. When format parameters for a specific media capability are combined from multiple a=mfcap lines which reference that media capability, the format-specific parameters are concatenated together and separated by ";" for construction of the corresponding format attribute (a=fmtp). The resulting format attribute will look something like the following (without line breaks):

```
a=fmtp:<fmt> <fmt-specific-param-list1>;
           <fmt-specific-param-list2>;
           ...
```

where <fmt> depends on the transport protocol in the manner defined in RFC4566. SDP cannot assess the legality of the resulting parameter list in the "a=fmtp" line; the user must take care to

ensure that legal parameter lists are generated.

The "mfcap" attribute can be provided at the session-level and the media-level. There can be more than one mfcap attribute at the session or media level. The unique media-cap-num is used to associate the parameters with a media capability.

As a simple example, a G.729 capability is, by default, considered to support comfort noise as defined by Annex B. Capabilities for G.729 with and without comfort noise support may thus be defined by:

```
a=rmcap:1,2 audio G729/8000
a=mfcap:2 annexb:no
```

Media format capability 1 supports G.729 with Annex B, whereas media format capability 2 supports G.729 without Annex B.

Example for H.263 video:

```
a=rmcap:1 video H263-1998/90000
a=rmcap:2 video H263-2000/90000
a=mfcap:1 CIF=4;QCIF=2;F=1;K=1
a=mfcap:2 profile=2;level=2.2
```

Finally, for six format combinations of the Adaptive MultiRate codec:

```
a=rmcap:1-3 AMR/8000/1
a=rmcap:4-6 AMR-WB/16000/1
a=mfcap:1,2,3,4 mode-change-capability=1
a=mfcap:5,6 mode-change-capability=2
a=mfcap:1,2,3,5 max-red=220
a=mfcap:3,4,5,6 octet-align=1
a=mfcap:1,3,5 mode-set=0,2,4,7
a=mfcap:2,4,6 mode-set=0,3,5,6
```

So that AMR codec #1, when specified in a pcfg attribute within an audio stream block (and assigned payload type number 98) as in

```
a=pcfg:1 m=1 pt=1:98
```

is essentially equivalent to the following

```
m=audio 49170 RTP/AVP 98
a=rtpmap:98 AMR/8000/1
a=fmtp:98 mode-change-capability=1; \
max-red=220; mode-set=0,2,4,7
```

and AMR codec #4 with payload type number 99, depicted by the

potential configuration:

```
a=pcfg:4 m=4, pt=4:99
```

is equivalent to the following:

```
m=audio 49170 RTP/AVP 99
a=rtpmap:99 AMR-WB/16000/1
a=fmtp:99 mode-change-capability=1; octet-align=1; \
mode-set=0,3,5,6
```

and so on for the other four combinations. SDP could thus convert the media capabilities specifications into one or more alternative media stream specifications, one of which can be chosen for the answer.

3.3.3. The Media-Specific Capability Attribute

Media-specific attributes, beyond the rtpmap and fmtp attributes, may be associated with media capability numbers via a new media-specific attribute, mscap, of the following form:

```
a=mscap:<media caps star> <att field> <att value>
```

Where <media caps star> is a (list of) media capability number(s), <att field> is the attribute name, and <att value> is the value field for the named attribute. The media capability numbers may include a wildcard ("*"), which will be used instead of any payload type mappings. In ABNF, we have:

```
media-specific-capability = "a=mscap:"
                             media-caps-star
                             1*WSP att-field ; from RFC4566
                             1*WSP att-value ; from RFC4566
media-caps-star            = media-cap-star-element
                             *["," media-cap-star-element]
media-cap-star-element    = media-cap-num [wildcard]
                             / media-cap-num-range [wildcard]
wildcard                  = "*"
```

Given an association between a media capability and a payload type number as specified by the pt= parameters in an lcfg or pcfg attribute line, a mscap line may be translated easily into a conventional SDP attribute line of the form

a=<att field>":"<fmt> <att value> ; <fmt> defined in [RFC4566]

A resulting attribute that is not a legal SDP attribute as specified by RFC4566 MUST be ignored by the receiver.

If a media capability number (or range) contains a wildcard character at the end, any payload type mapping specified for that media specific capability will be use the wildcard character instead of the payload type.

A single mscap line may refer to multiple media capabilities; this is equivalent to multiple mscap lines, each with the same attribute values (but different media capability numbers), one line per media capability.

Multiple mscap lines may refer to the same media capability, but, unlike the mfcap attribute, no concatenation operation is defined. Hence, multiple mscap lines applied to the same media capability is equivalent to multiple lines of the specified attribute in a conventional media record.

Here is an example with the rtcp-fb attribute, modified from an example in [RFC5104] (with the session-level and audio media omitted). If the offer contains a media block like the following (note the wildcard character),

```
m=video 51372 RTP/AVP 98
a=rtpmap:98 H263-1998/90000
a=tcap:1 RTP/AVPF
a=rmcap:1 H263-1998/90000
a=mscap:1 rtcp-fb ccm tstr
a=mscap:1 rtcp-fb ccm fir
a=mscap:1* rtcp-fb ccm tmmbr smaxpr=120
a=pcfg:1 t=1 m=1 pt=1:98
```

and if the proposed configuration is chosen, then the equivalent media block would look like

```
m=video 51372 RTP/AVPF 98
a=rtpmap:98 H263-1998/90000
a=rtcp-fb:98 ccm tstr
a=rtcp-fb:98 ccm fir
a=rtcp-fb:* ccm tmmbr smaxpr=120
```

3.3.4. New Configuration Parameters

Along with the new attributes for media capabilities, new extension parameters are defined for use in the potential configuration, the actual configuration, and/or the new latent configuration defined in Section 3.3.5.

3.3.4.1. The Media Configuration Parameter (m=)

The media configuration parameter is used to specify the media encoding(s) and related parameters for a potential, actual, or latent configuration. Adhering to the ABNF for extension-config-list in [RFC5939] with

```
ext-cap-name = "m"
ext-cap-list = media-cap-num-list
               [*(BAR media-cap-num-list)]
```

we have

```
media-config-list = ["+" ]"m=" media-cap-num-list
                   [*(BAR media-cap-num-list)]
                   ; BAR is defined in RFC5939
                   ; media-cap-num-list is defined above
```

Alternative media configurations are separated by a vertical bar ("|"). The alternatives are ordered by preference, most-preferred first. When media capabilities are not included in a potential configuration at the media level, the media type and media format from the associated "m=" line will be used. The use of the plus sign ("+") is described in RFC5939.

3.3.4.2. The Payload Type Number Mapping Parameter (pt=)

The payload type number mapping parameter is used to specify the payload type number to be associated with each media type in a potential, actual, or latent configuration. We define the payload type number mapping parameter, payload-number-config-list, in accordance with the extension-config-list format defined in [RFC5939]. In ABNF:

```
payload-number-config-list = ["+" ]"pt=" media-map-list
media-map-list = media-map *[" ," media-map]
media-map = media-cap-num ":" payload-type-number
           ; media-cap-num is defined in 3.3.1
payload-type-number = 1*3(DIGIT) ; RTP payload type number
                    / "*"       ; dummy
```

The example in Section 3.3.7 shows how the parameters from the `rmcap` line are mapped to payload type numbers from the `pcfg` "pt" parameter. The use of the plus sign "+" is described in RFC5939.

The "*" value for payload-type-number is used in cases such as BFCP [RFC4583] in which the `fmt` list in the `m`-line is ignored.

A latent configuration represents a future capability, hence the `pt` parameter is not directly meaningful in the `lcfg` attribute because no actual media session is being offered or accepted; it is permitted in order to tie any payload type number parameters within attributes to the proper media format. A primary example is the case of format parameters for the Redundant Audio Data (RED) payload, which are payload type numbers. Specific payload type numbers used in a latent configuration MAY be interpreted as suggestions to be used in any future offer based on the latent configuration, but they are not binding; the offerer and/or answerer may use any payload type numbers each deems appropriate. The use of explicit payload type numbers for latent configurations can be avoided by use of the parameter substitution rule of Section 3.3.7. Future extensions are also permitted.

3.3.4.3. The Media Type Parameter

When a latent configuration is specified (always at the media level), indicating the ability to support an additional media stream, it is necessary to specify the media type (audio, video, etc.) as well as the format and transport type. The media type parameter is defined in ABNF as

```
media-type = ["+"] "mt=" media; media defined in RFC4566
```

At present, the `media-type` parameter is used only in the latent configuration attribute, and the use of the "+" prefix to specify that the entire attribute line is to be ignored if the `mt` parameter is not understood, is unnecessary. However, if the `media-type` parameter is later added to an existing capability attribute such as `pcfg`, then the "+" would be useful. The media format(s) and transport type(s) are specified using the media configuration parameter ("m=") defined above, and the transport parameter ("t=") defined in [RFC5939], respectively.

3.3.5. The Latent Configuration Attribute

One of the goals of this work is to permit the exchange of supportable media configurations in addition to those offered or accepted for immediate use. Such configurations are referred to as "latent configurations". For example, a party may offer to establish

a session with an audio stream, and, at the same time, announce its ability to support a video stream as part of the same session. The offerer can supply its video capabilities by offering one or more latent video configurations along with the media stream for audio; the responding party may indicate its ability and willingness to support such a video session by returning a corresponding latent configuration.

Latent configurations returned in SDP answers must match offered latent configurations (or parameter subsets thereof). Therefore, it is appropriate for the offering party to announce most, if not all, of its capabilities in the initial offer. This choice has been made in order to keep the size of the answer more compact by not requiring acap, rmcap, tcap, etc. lines in the answer.

Latent configurations may be announced by use of the latent configuration attribute, which is defined in a manner very similar to the potential configuration attribute. The latent configuration attribute combines the properties of a media line and a potential configuration. The media type (mt=) and the transport protocol(s) (t=) MUST be specified since the latent configuration is independent of any media line present. In most cases, the media configuration (m=) parameter MUST be present as well (see Section 4 for examples). The lcfg attribute is a media level attribute and, like a media line, it ends the session level of the session description if it appears before any media line.

Each media line in an SDP description represents an offered simultaneous media stream, whereas each latent configuration represents an additional stream which may be negotiated in a future offer/answer exchange. Session capability attributes may be used to determine whether a latent configuration may be used to form an offer for an additional simultaneous stream or to reconfigure an existing stream in a subsequent offer/answer exchange.

The latent configuration attribute is of the form:

```
a=lcfg:<config-number> <latent-cfg-list>
```

which adheres to the [RFC4566] "attribute" production with att-field and att-value defined as:

```
att-field   = "lcfg"
att-value   = config-number 1*WSP lcfg-cfg-list
config-number = 1*10(DIGIT) ; defined in RFC5234
lcfg-cfg-list = media-type 1*WSP pot-cfg-list
               ; as defined in RFC5939
               ; and extended herein
```

The media-type (mt=) parameter identifies the media type (audio, video, etc.) to be associated with the latent media stream, and MUST be present. The pot-cfg-list MUST contain a transport-protocol-config-list (t=) parameter and a media-config-list (m=) parameter. The pot-cfg-list MUST NOT contain more than one instance of each type of parameter list. As specified in [RFC5939], the use of the "+" prefix with a parameter indicates that the entire configuration MUST be ignored if the parameter is not understood; otherwise, the parameter itself may be ignored.

Media stream payload numbers are not assigned by a latent configuration. Assignment will take place if and when the corresponding stream is actually offered via an m-line in a later exchange. The payload-number-config-list is included as a parameter to the lcfg attribute in case it is necessary to tie payload numbers in attribute capabilities to specific media capabilities.

If an lcfg attribute invokes an acap attribute that appears at the session level, then that attribute will be expected to appear at the session level of a subsequent offer when and if a corresponding media stream is offered. Otherwise, acap attributes which appear at the media level represent media-level attributes. Note, however, that rmcap, omcap, mfcap, mscap, and tcap attributes may appear at the session level because they always result in media-level attributes or m-line parameters.

The configuration numbers for latent configurations do not imply a preference; the offerer will imply a preference when actually offering potential configurations derived from latent configurations negotiated earlier. Note however that the offerer of latent configurations MAY specify preferences for combinations of potential and latent configurations by use of the sescap attribute defined in Section 3.3.8. For example, if an SDP offer contains, say, an audio stream with pcfg:1, and two latent video configurations, lcfg:2, and lcfg:3, then a session with one audio stream and one video stream could be specified by including "a=sescap:1 1,2|3". One audio stream and two video streams could be specified by including "a=sescap:2 1,2,3" in the offer. In order to permit combinations of latent and potential configurations in session capabilities, latent configuration numbers MUST be different from those used for potential configurations. This restriction is especially important if the offerer does not require cmed-v0 capability and the recipient of the offer doesn't support it. If the lcfg attribute is not recognized, the capability attributes intended to be associated with it may be confused with those associated with a potential configuration of some other media stream.

If a cryptographic attribute, such as the SDES "a=crypto:" attribute

[RFC4568], is referenced by a latent configuration through an acap attribute, any keying material required in the conventional attribute, such as the SDES key/salt string, MUST be included in order to satisfy formatting rules for the attribute. The actual value(s) of the keying material SHOULD be meaningless, and the receiver of the lcfg attribute MUST ignore the values.

3.3.6. Enhanced Potential Configuration Attribute

The present work requires new extensions (parameters) for the pcfg attribute defined in the base protocol [RFC5939]. The parameters and their definitions are "borrowed" from the definitions provided for the latent configuration attribute in Section 3.3.5. The expanded ABNF definition of the pcfg attribute is

```
a=pcfg: <config-number> [<pot-cfg-list>]
```

where

```
config-number = 1*DIGIT ;defined in [RFC5234]
pot-cfg-list = pot-config *(1*WSP pot-config)
pot-config = / attribute-config-list / ;def in [RFC5939]
transport-protocol-config-list / ;defined in [RFC5939]
extension-config-list / ;[RFC5939]
media-config-list / ; Section 3.3.4.1
payload-number-config-list ; Section 3.3.4.2
```

Except for the extension-config-list, the pot-cfg-list MUST NOT contain more than one instance of each parameter list.

3.3.6.1. Returning Capabilities in the Answer

Potential and/or latent configuration attributes may be returned within an answer SDP to indicate the ability of the answerer to support alternative configurations of the corresponding stream(s). For example, an offer may include multiple potential configurations for a media stream and/or latent configurations for additional streams; the corresponding answer will indicate (via an acfg attribute) the configuration accepted and used to construct the base configuration for each active media stream in the reply, but the reply MAY also contain potential and/or latent configuration attributes, with parameters, to indicate which other offered configurations would be acceptable. This information is useful if it becomes desirable to reconfigure a media stream, e.g., to reduce resource consumption.

When potential and/or latent configurations are returned in an answer, all numbering MUST refer to the configuration and capability

attribute numbering of the offer. The offered capability attributes need not be returned in the answer. The answer MAY include additional capability attributes and/or configurations (with distinct numbering). The parameter values of any returned pcfg or lcfg attributes MUST be a subset of those included in the offered configurations or those added by the answerer; values may be omitted only if they were indicated as alternative sets, or optional, in the original offer. The parameter set indicated in the returned acfg attribute need not be repeated in a returned pcfg attribute. The answerer may return more than one pcfg attribute with the same configuration number if it is necessary to describe selected combinations of optional or alternative parameters.

Similarly, one or more session capability attributes (a=sescap) may be returned to indicate which of the offered session capabilities is/are supportable by the answerer (see Section 3.3.8.)

Note that, although the answerer MAY return capabilities beyond those included by the offerer, these capabilities MUST NOT be used to form any base level media description in the answer. For this reason, it is advisable for the offerer to include most, if not all, potential and latent configurations it can support in the initial offer, unless the size of the resulting SDP is a concern. Either party MAY later announce additional capabilities by renegotiating the session in a second offer/answer exchange.

3.3.6.2. Payload Type Number Mapping

When media capabilities defined in rmcap attributes are used in potential configuration lines, the transport protocol uses RTP and it is necessary to assign payload type numbers. In some cases, it is desirable to assign different payload type numbers to the same media capability when used in different potential configurations. One example is when configurations for AVP and SAVP are offered: the offerer would like the answerer to use different payload type numbers for encrypted and unencrypted media so that it (the offerer) can decide whether or not to render early media which arrives before the answer is received. This association of distinct payload type number(s) with different transport protocols requires a separate pcfg line for each protocol. Clearly, this technique cannot be used if the number of potential configurations exceeds the number of possible payload type numbers.

3.3.6.3. Processing of Media-Format-Related Conventional Attributes for Potential Configurations

In cases in which media capabilities negotiation is employed, SDP records are likely to contain conventional attributes such as rtpmap,

fmtp, and other media-format-related lines, as well as capability attributes such as rmcap, omcap, mfcap, and mscap which map into those conventional attributes when invoked by a potential configuration. In such cases, it MAY be appropriate to employ the delete-attributes option [RFC5939] in the attribute configuration list parameter in order to avoid the generation of conflicting fmtp attributes for a particular configuration. Any media-specific attributes in the media block which refer to media formats not used by the potential configuration MUST be ignored.

For example:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 3456 RTP/AVP 0 18 100
a=rtpmap:100 telephone-events
a=fmtp:100 0-11
a=rmcap:1 PCMU/8000
a=rmcap:2 g729/8000
a=rmcap:3 telephone-events/8000
a=mfcap:3 0-15
a=pcfg:1 m=2,3|1,3 a=-m pt=1:0,2:18,3:100
a=pcfg:2
```

In this example, PCMU is media capability 1, G729 is media capability 2, and telephone-event is media capability 3. The a=pcfg:1 line specifies that the preferred configuration is G.729 with extended dtmf events, second is G.711 mu-law with extended dtmf events, and the base media-level attributes are to be deleted. Intermixing of G.729, G.711, and "commercial" dtmf events is least preferred (the base configuration provided by the "m=" line, which is, by default, the least preferred configuration). The rtpmap and fmtp attributes of the base configuration are replaced by the rmcap and mfcap attributes when invoked by the proposed configuration.

If the preferred configuration is selected, the SDP answer will look like

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=csup:med-v0
```

```
m=audio 6543 RTP/AVP 18 100
a=rtpmap:100 telephone-events/8000
a=fmtp:100 0-15
a=acfg:1 m=2,3 pt=1:0,2:18,3:100
```

3.3.7. Substitution of Media Payload Type Numbers in Capability Attribute Parameters

In some cases, for example, when an RFC 2198 redundancy audio subtype (RED) capability is defined in an mfcap attribute, the parameters to an attribute may contain payload type numbers. Two options are available for specifying such payload type numbers. They may be expressed explicitly, in which case they are bound to actual payload types by means of the payload type number parameter (pt=) in the appropriate potential or latent configuration. For example, the following SDP fragment defines a potential configuration with redundant G.711 mu-law:

```
m=audio 45678 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=rmcap:1 PCMU/8000
a=rmcap:2 RED/8000
a=mfcap:2 0/0
a=pcfg:1 m=2,1 pt=2:98,1:0
```

The potential configuration is then equivalent to

```
m=audio 45678 RTP/AVP 98 0
a=rtpmap:0 PCMU/8000
a=rtpmap:98 RED/8000
a=fmtp:98 0/0
```

A more general mechanism is provided via the parameter substitution rule. When an mfcap, mscap, or acap attribute is processed, its arguments will be scanned for a payload type number escape sequences of the following form (in ABNF):

ptn-esc = "%m=" media-cap-num "%" ; defined in 3.3.1

If the sequence is found, the sequence is replaced by the payload type number assigned to the media capability number, as specified by the pt= parameter in the selected potential configuration; only actual payload type numbers are supported - wildcards are excluded. The sequence "%%" (null digit string) is replaced by a single percent sign and processing continues with the next character, if any.

For example, the above offer sequence could have been written as

```

m=audio 45678 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=rmcap:1 PCMU/8000
a=rmcap:2 RED/8000
a=mfcap:2 %m=1%/m=1%
a=pcfg:1 m=2,1 pt=2:98,1:0

```

and the equivalent SDP is the same as above.

3.3.8. The Session Capability Attribute

The session capability attribute provides a means for the offerer and/or the answerer to specify combinations of specific media stream configurations which it is willing and able to support. Each session capability in an offer or answer MAY be expressed as a list of required potential configurations, and MAY include a list of optional potential and/or latent configurations.

The choices of session capabilities may be based on processing load, total bandwidth, or any other criteria of importance to the communicating parties. If the answerer supports media capabilities negotiation, and session configurations are offered, it MUST accept one of the offered configurations, or it MUST refuse the session. Therefore, if the offer includes any session capabilities, it SHOULD include all the session capabilities the offerer is willing to support.

The session capability attribute is described by:

```
"a=sescap:" <session num> <list of configs>
```

which corresponds to the standard value attribute definition with

```

att-field      = "sescap"
att-value      = session-num 1*WSP list-of-configs
                  [1*WSP optional-configs]
session-num    = 1*10(DIGIT) ; defined in RFC5234
list-of-configs = alt-config *["|" alt-config]
optional-configs = "[" list-of-configs "]"
alt-config     = config-number *["|" config-number]
                  ; config-number defined in RFC5939

```

The session-num identifies the session; a lower-number session is preferred over a higher-numbered session. Each alt-config list specifies alternative media configurations within the session; preference is based on config-num as specified in [RFC5939]. Note that the session preference order, when present, takes precedence over the individual media stream configuration preference order.

Use of session capability attributes requires that configuration numbers assigned to potential and latent configurations MUST be unique across the entire session; [RFC5939] requires only that pcfg configuration numbers be unique within a media description.

As an example, consider an endpoint that is capable of supporting an audio stream with either one H.264 video stream or two H.263 video streams with a floor control stream. The SDP offer might look like the following (offering audio, two H.263 video streams and BFCP)- the empty lines are added for readability only (not part of valid SDP):

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=sescap:2 1,2,3,5
a=sescap:1 1,4

m=audio 54322 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=pcfg:1

m=video 22344 RTP/AVP 102
a=rtpmap:102 H263-1998/90000
a=fmtp:102 CIF=4;QCIF=2;F=1;K=1
i= main video stream
a=label:11
a=pcfg:2
a=rmcap:1 H264/90000
a=mfcap:1 profile-level-id=42A01E; packetization-mode=2
a=acap:1 label:13
a=pcfg:4 m=1 a=1 pt=1:104

m=video 33444 RTP/AVP 103
a=rtpmap:103 H263-1998/90000
a=fmtp:103 CIF=4;QCIF=2;F=1;K=1
i= secondary video (slides)
a=label:12
a=pcfg:3

m=application 33002 TCP/BFCP *
a=setup:passive
a=connection:new
a=floorid:1 m-stream:11 12
a=floor-control:s-only
a=confid:4321
```

```
a=userid:1234
a=pcfg:5
```

If the answerer understands MediaCapNeg, but cannot support the Binary Floor Control Protocol, then it would respond with (invalid empty lines in SDP included again for readability):

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.22
t=0 0
a=csup:med-v0
a=sescap:1 1,4

m=audio 23456 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=acfg:1

m=video 41234 RTP/AVP 104
a=rtpmap:100 H264/90000
a=fmtp:104 profile-level-id=42A01E; packetization-mode=2
a=acfg:4 m=1 a=1 pt=1:104

m=video 0 RTP/AVP 103
a=acfg:3

m=application 0 TCP/BFCP *
a=acfg:5
```

An endpoint that doesn't support Media capabilities negotiation, but does support H.263 video, would respond with one or two H.263 video streams. In the latter case, the answerer may issue a second offer to reconfigure the session to one audio and one video channel using H.264 or H.263.

Session capabilities can include latent capabilities as well. Here's a similar example in which the offerer wishes to initially establish an audio stream, and prefers to later establish two video streams with chair control. If the answerer doesn't understand Media CapNeg, or cannot support the dual video streams or flow control, then it may support a single H.264 video stream. Note that establishment of the most favored configuration will require two offer/answer exchanges.

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
```

```
t=0 0
a=creq:med-v0
a=sescap:1 1,3,4,5
a=sescap:2 1,2
a=sescap:3 1
a=rmcap:1 H263-1998/90000
a=mfcap:1 CIF=4;QCIF=2;F=1;K=1
a=tcap:1 RTP/AVP TCP/BFCP
m=audio 54322 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=label:11
a=pcfg:1
m=video 22344 RTP/AVP 102
a=rtpmap:102 H264/90000
a=fmtp:102 profile-level-id=42A01E; packetization-mode=2
a=label:11
a=content:main
a=pcfg:2
a=lcfg:3 mt=video t=1 m=1 a=31,32 i=3
a=acap:31 label:12
a=acap:32 content:main
a=lcfg:4 mt=video t=1 m=1 a=41,42 i=4
a=acap:41 label:13
a=acap:42 content:slides
a=lcfg:5 mt=application m=51 t=51
a=tcap:51 TCP/BFCP
a=rmcap:51 *
a=acap:51 setup:passive
a=acap:52 connection:new
a=acap:53 floorid:1 m-stream:12 13
a=acap:54 floor-control:s-only
a=acap:55 confid:4321
a=acap:56 userid:1234
```

In this example, the default offer, as seen by endpoints which do not understand capabilities negotiation, proposes a PCMU audio stream and an H.264 video stream. Note that the offered lcfg lines for the video streams don't carry pt= parameters because they're not needed (payload type numbers will be assigned in the offer/answer exchange that establishes the streams). Note also that the three rmcap, mfcap, and tcap attributes used by lcfg:3 and lcfg:4 are included at the session level so they may be referenced by both latent configurations. As per Section 3.3, the media attributes generated from the rmcap, mfcap, and tcap attributes are always media-level attributes. If the answerer supports Media CapNeg, and supports the most desired configuration, it would return the following SDP:


```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.22
t=0 0
a=csup:med-v0
a=sescap:1 1,3,4,5
a=sescap:2 1,2
a=sescap:3 1
m=audio 23456 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=acfg:1
m=video 0 RTP/AVP 102
a=pcfg:2
a=lcfg:3 mt=video t=1 m=1 a=31,32
a=lcfg:4 mt=video t=1 m=1 a=41,42
a=lcfg:5 mt=application t=2
```

This exchange supports immediate establishment of an audio stream for preliminary conversation. This exchange would presumably be followed at the appropriate time with a "reconfiguration" offer/answer exchange to add the video and chair control streams.

3.4. Offer/Answer Model Extensions

EDITOR'S NOTE: SECTION NEEDS MORE ELABORATE PROCEDURES

In this section, we define extensions to the offer/answer model defined in RFC3264 [RFC3264] and [RFC5939] to allow for media capabilities, bandwidth capabilities, and latent configurations to be used with the SDP Capability Negotiation framework.

The [RFC5939] provides a relatively compact means to offer the equivalent of an ordered list of alternative media stream configurations (as would be described by separate m= lines and associated attributes). The attributes acap, mscap, mfcap and rmcap are designed to map somewhat straightforwardly into equivalent m= lines and conventional attributes when invoked by a pcfg, lcfg, or acfg attribute with appropriate parameters. The a=pcfg: lines, along with the m= line itself, represent offered media configurations. The a=lcfg: lines represent alternative capabilities for future use.

3.4.1. Generating the Initial Offer

When an endpoint generates an initial offer and wants to use the functionality described in the current document, it should identify and define the codecs it can support via rmcap, mfcap and mscap attributes. The SDP media line(s) should be made up with the

configuration to be used if the other party does not understand capability negotiations (by default, this is the least preferred configuration). Typically, the media line configuration will contain the minimum acceptable capabilities. The offer MUST include the level of capability negotiation extensions needed to support this functionality in a "creq" attribute.

Preferred configurations for each media stream are identified following the media line. The present offer may also include latent configuration (lcfg) attributes, at the media level, describing media streams and/or configurations the offerer is not now offering, but which it is willing to support in a future offer/answer exchange. A simple example might be the inclusion of a latent video configuration in an offer for an audio stream.

3.4.2. Generating the Answer

When the answering party receives the offer and if it supports the required capability negotiation extensions, it should select the most-preferred configuration it can support for each media stream, and build its answer accordingly. The configuration selected for each accepted media stream is placed into the answer as a media line with associated parameters and attributes. If a proposed configuration is chosen, the answer must include the supported extension attribute and each media stream for which a proposed configuration was chosen must contain an actual configuration (acfg) attribute to indicate just which pcfg attribute was used to build the answer. The answer should also include any potential or latent configurations the answerer can support, especially any configurations compatible with other potential or latent configurations received in the offer. The answerer should make note of those configurations it might wish to offer in the future.

3.4.3. Offerer Processing of the Answer

When the offerer receives the answer, it should make note of any capabilities and/or latent configurations for future use. The media line(s) must be processed in the normal way to identify the media stream(s) accepted by the answer, if any. The acfg attribute, if present, may be used to verify the proposed configuration used to form the answer, and to infer the lack of acceptability of higher-preference configurations that were not chosen. Note that the base specification [RFC5939] requires the answerer to choose the highest preference configuration it can support, subject to local policies.

3.4.4. Modifying the Session

If, at a later time, one of the parties wishes to modify the operating parameters of a session, e.g., by adding a new media stream, or by changing the properties used on an existing stream, it may do so via the mechanisms defined for offer/answer [RFC3264]. If the initiating party has remembered the codecs, potential configurations, and latent configurations announced by the other party in the earlier negotiation, it may use this knowledge to maximize the likelihood of a successful modification of the session. Alternatively, the initiator may perform a new capabilities exchange as part of the reconfiguration. In such a case, the new capabilities will replace the previously-negotiated capabilities. This may be useful if conditions change on the endpoint.

4. Examples

In this section, we provide examples showing how to use the Media Capabilities with the SDP Capability Negotiation.

4.1. Alternative Codecs

This example provide a choice of one of six variations of the adaptive multirate codec. In this example, the default configuration as specified by the media line is the same as the most preferred configuration. Each configuration uses a different payload type number so the offerer can interpret early media.

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 54322 RTP/AVP 96
rtptime:96 AMR-WB/16000/1
a=fmtp:96 mode-change-capability=1; max-red=220; \
mode-set=0,2,4,7
a=rmcap:1,3,5 audio AMR-WB/16000/1
a=rmcap:2,4,6 audio AMR/8000/1
a=mfcap:1,2,3,4 mode-change-capability=1
a=mfcap:5,6 mode-change-capability=2
a=mfcap:1,2,3,5 max-red=220
a=mfcap:3,4,5,6 octet-align=1
a=mfcap:1,3,5 mode-set=0,2,4,7
a=mfcap:2,4,6 mode-set=0,3,5,6
a=pcfg:1 m=1 pt=1:96
a=pcfg:2 m=2 pt=2:97
a=pcfg:3 m=3 pt=3:98
a=pcfg:4 m=4 pt=4:99
a=pcfg:5 m=5 pt=5:100
a=pcfg:6 m=6 pt=6:101
```

In the above example, media capability 1 could have been excluded from the first rmcap declaration and from the corresponding mfcap attributes, and the pcfg:1 attribute line could have been simply "pcfg:1".

The next example offers a video stream with three options of H.264 and 4 transports. It also includes an audio stream with different audio qualities: four variations of AMR, or AC3. The offer looks something like:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=An SDP Media NEG example
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=ice-pwd:speEc3QGZiNWpVLFJhQX
m=video 49170 RTP/AVP 100
c=IN IP4 192.0.2.56
a=maxprate:1000
a=rtcp:51540
a=sendonly
a=candidate 12345 1 UDP 9 192.0.2.56 49170 host
a=candidate 23456 2 UDP 9 192.0.2.56 51540 host
a=candidate 34567 1 UDP 7 198.51.100.1 41345 srflx raddr \
192.0.2.56 rport 49170
a=candidate 45678 2 UDP 7 198.51.100.1 52567 srflx raddr \
192.0.2.56 rport 51540
a=candidate 56789 1 UDP 3 192.0.2.100 49000 relay raddr \
192.0.2.56 rport 49170
a=candidate 67890 2 UDP 3 192.0.2.100 49001 relay raddr \
192.0.2.56 rport 51540
b=AS:10000
b=TIAS:10000000
b=RR:4000
b=RS:3000
a=rtpmap:100 H264/90000
a=fmtp:100 profile-level-id=42A01E; packetization-mode=2; \
sprop-parameter-sets=Z0IACpZTBmI,aMljiA==; \
sprop-interleaving-depth=45; sprop-deint-buf-req=64000; \
sprop-init-buf-time=102478; deint-buf-cap=128000
a=tcap:1 RTP/SAVPF RTP/SAVP RTP/AVPF
a=rmcap:1-3,7-9 H264/90000
a=rmcap:4-6 rtx/90000
a=mfcap:1-9 profile-level-id=42A01E
a=mfcap:1-9 aMljiA==
a=mfcap:1,4,7 packetization-mode=0
a=mfcap:2,5,8 packetization-mode=1
a=mfcap:3,6,9 packetization-mode=2
a=mfcap:1-9 sprop-parameter-sets=Z0IACpZTBmI
a=mfcap:1,7 sprop-interleaving-depth=45; \
sprop-deint-buf-req=64000; sprop-init-buf-time=102478; \
deint-buf-cap=128000
a=mfcap:4 apt=100
a=mfcap:5 apt=99
a=mfcap:6 apt=98
a=mfcap:4-6 rtx-time=3000
a=mscap:1-6 rtcp-fb nack
```

```
a=acap:1 crypto:1 AES_CM_128_HMAC_SHA1_80 \
inline:d0RmdmcmVCspeEc3QGZiNWpVLFJhQXlcfHAWJSoj|220|1:32
a=pcfg:1 t=1 m=1,4 a=1 pt=1:100,4:97
a=pcfg:2 t=1 m=2,5 a=1 pt=2:99,4:96
a=pcfg:3 t=1 m=3,6 a=1 pt=3:98,6:95
a=pcfg:4 t=2 m=7 a=1 pt=7:100
a=pcfg:5 t=2 m=8 a=1 pt=8:99
a=pcfg:6 t=2 m=9 a=1 pt=9:98
a=pcfg:7 t=3 m=1,3 pt=1:100,4:97
a=pcfg:8 t=3 m=2,4 pt=2:99,4:96
a=pcfg:9 t=3 m=3,6 pt=3:98,6:95
m=audio 49176 RTP/AVP 101 100 99 98
c=IN IP4 192.0.2.56
a=ptime:60
a=maxptime:200
a=rtcp:51534
a=sendonly
a=candidate 12345 1 UDP 9 192.0.2.56 49176 host
a=candidate 23456 2 UDP 9 192.0.2.56 51534 host
a=candidate 34567 1 UDP 7 198.51.100.1 41348 srflx \
raddr 192.0.2.56 rport 49176
a=candidate 45678 2 UDP 7 198.51.100.1 52569 srflx \
raddr 192.0.2.56 rport 51534
a=candidate 56789 1 UDP 3 192.0.2.100 49002 relay \
raddr 192.0.2.56 rport 49176
a=candidate 67890 2 UDP 3 192.0.2.100 49003 relay \
raddr 192.0.2.56 rport 51534
b=AS:512
b=TIAS:512000
b=RR:4000
b=RS:3000
a=maxprate:120
a=rtpmap:98 AMR-WB/16000
a=fmtp:98 octet-align=1; mode-change-capability=2
a=rtpmap:99 AMR-WB/16000
a=fmtp:99 octet-align=1; crc=1; mode-change-capability=2
a=rtpmap:100 AMR-WB/16000/2
a=fmtp:100 octet-align=1; interleaving=30
a=rtpmap:101 AMR-WB+/72000/2
a=fmtp:101 interleaving=50; int-delay=160000;
a=rmcap:14 ac3/48000/6
a=acap:23 crypto:1 AES_CM_128_HMAC_SHA1_80 \
inline:d0RmdmcmVCspeEc3QGZiNWpVLFJhQXlcfHAWJSoj|220|1:32
a=tcap:4 RTP/SAVP
a=pcfg:10 t=4 a=23
a=pcfg:11 t=4 m=14 a=23 pt=14:102
```

This offer illustrates the advantage in compactness that arises if

one can avoid deleting the base configuration attributes and recreating them in acap attributes for the potential configurations.

4.2. Alternative Combinations of Codecs (Session Configurations)

If an endpoint has limited signal processing capacity, it might be capable of supporting, say, a G.711 mu-law audio stream in combination with an H.264 video stream, or a G.729B audio stream in combination with an H.263-1998 video stream. It might then issue an offer like the following:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=sescap:1 2,4
a=sescap:2 1,3
m=audio 54322 RTP/AVP 18
a=rtpmap:18 G729/8000
a=fmtp:18 annexb=yes
a=rmcap:1 PCMU/8000
a=pcfg:1 m=1 pt=1:0
a=pcfg:2
m=video 54344 RTP/AVP 100
a=rtpmap:100 H263-1998/90000
a=rmcap:2 H264/90000
a=mfcap:2 profile-level-id=42A01E; packetization-mode=2
a=pcfg:3 m=2 pt=2:101
a=pcfg:4
```

Note that the preferred session configuration (and the default as well) is G.729B with H.263. This overrides the individual media stream preferences which are PCMU and H.264 by the potential configuration numbering rule.

4.3. Latent Media Streams

Consider a case in which the offerer can support either G.711 mu-law, or G.729B, along with DTMF telephony events for the 12 common touchtone signals, but is willing to support simple G.711 mu-law audio as a last resort. In addition, the offerer wishes to announce its ability to support video in the future, but does not wish to offer a video stream at present. The offer might look like the following:

```

v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 23456 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=rmcap:1 PCMU/8000
a=rmcap:2 g729/8000
a=rmcap:3 telephone-event/8000
a=mfcap:3 0-11
a=pcfg:1 m=1,3|2,3 pt=1:0,2:18,3:100 a=lcfg:10 mt=video t=1
m=10|11
a=rmcap:10 H263-1998/90000
a=rmcap:11 H264/90000
a=tcap:1 RTP/AVP

```

The lcfg attribute line announces support for H.263 and H.264 video (H.263 preferred) for future reference. The m-line and the rtpmap attribute offer an audio stream and provide the lowest precedence configuration (PCMU without any DTMF encoding). The rmcap lines define the media capabilities (PCMU, G729, and telephone-event) to be offered in potential configurations. The mfcap attribute provides the format parameters for telephone-events, specifying the 12 commercial DTMF 'digits'. The pcfg attribute line defines the most-preferred media configuration as PCMU plus DTMF events and the next-most-preferred configuration as G.729B plus DTMF events.

If the answerer is able to support all the potential configurations, and also support H.263 video (but not H.264), it would reply with an answer like:

```

v=0
o=- 24351 621814 IN IP4 192.0.2.2
s=
c=IN IP4 192.0.2.2
t=0 0
a=csup:med-v0
m=audio 54322 RTP/AVP 0 100
a=rtpmap:0 PCMU/8000
a=rtpmap:100 telephone-event/8000
a=fmtp:100 0-11
a=acfg:1 m=1,3 pt=1:0,3:100
a=pcfg:1 m=2,3 pt=2:18,3:100 a=lcfg:1 mt=video t=1 m=10

```

The lcfg attribute line announces the capability to support H.263 video at a later time. The media line and subsequent rtpmap and fmtp

attribute lines present the selected configuration for the media stream. The `acfg` attribute line identifies the potential configuration from which it was taken, and the `pcfg` attribute line announces the potential capability to support G.729 with DTMF events as well. If, at some later time, congestion becomes a problem in the network, either party may, with expectation of success, offer a reconfiguration of the media stream to use G.729 in order to reduce packet sizes.

5. IANA Considerations

5.1. New SDP Attributes

The IANA is hereby requested to register the following new SDP attributes:

Attribute name: rmcap
Long form name: RTP-based media capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate RTP-based media capability number(s) with media subtype and encoding parameters
Appropriate Values: see Section 3.3.1

Attribute name: omcap
Long form name: Non RTP-based media capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate non RTP-based media capability number(s) with media subtype and encoding parameters
Appropriate Values: see Section 3.3.1

Attribute name: mfcap
Long form name: media format capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate media format attributes and parameters with media format capabilities
Appropriate Values: see Section 3.3.2

Attribute name: mscap
Long form name: media-specific capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate media-specific attributes and parameters with media capabilities
Appropriate Values: see Section 3.3.3

Attribute name: lcfcg
Long form name: latent configuration
Type of attribute: media-level
Subject to charset: no
Purpose: to announce supportable media streams without offering them for immediate use.
Appropriate Values: see Section 3.3.5

Attribute name: sescap
Long form name: session capability
Type of attribute: session-level
Subject to charset: no
Purpose: to specify and prioritize acceptable combinations of media stream configurations.
Appropriate Values: see Section 3.3.8

5.2. New SDP Option Tag

The IANA is hereby requested to add the new option tag "med-v0", defined in this document, to the SDP Capability Option Negotiation Capability registry created for [RFC5939].

5.3. New SDP Capability Negotiation Parameters

The IANA is hereby requested to expand the SDP Capability Negotiation Potential Configuration Parameter Registry established by [RFC5939] to become the SDP Capability Negotiation Configuration Parameter Registry and to include parameters for the potential, actual and latent configuration attributes. The new parameters to be registered are the "m" for "media", "pt" for "payload type number", and "mt" for "media type" parameters. Note that the "mt" parameter is defined for use only in the latent configuration attribute.

6. Security Considerations

EDITOR'S NOTE: SECTION NEEDS TO BE EXPANDED

The security considerations of [RFC5939] apply for this document.

The addition of negotiable media encoding, bandwidth attributes, and connection data in this specification can cause problems for middleboxes which attempt to control bandwidth utilization, media flows, and/or processing resource consumption as part of network policy, but which do not understand the media capability negotiation feature. As for the initial CapNeg work, the SDP answer is formulated in such a way that it always carries the selected media encoding and bandwidth parameters for every media stream selected. Pending an understanding of capabilities negotiation, the middlebox should examine the answer SDP to obtain the best picture of the media streams being established.

As always, middleboxes can best do their job if they fully understand media capabilities negotiation.

7. Changes from previous versions

7.1. Changes from version 11

- o Corrected several statements implying lcfg was a session-level attribute.
- o Added non-RTP based media format capabilities ("a=omcap") and renamed "mcap" to "rmcap"

7.2. Changes from version 10

- o Defined the latent configuration attribute as a media-level attribute because it specifies a possible future media stream. Added text to clarify how to specify alternative configurations of a single latent stream and/or multiple streams.
- o Improved the definition of the session capability attribute to permit both required configurations and optional configurations - latent configurations cannot be required because they have not yet been offered.
- o Removed the special-case treatment of conflicts between base-level fmtp attributes and fmtp attributes generated for a configuration via invoked mcap and mfcap attributes.
- o Removed reference to bandwidth capability (bcap) attribute.
- o Changed various "must", etc., terms to normative terms ("MUST", etc.) as appropriate, in Section 3.3.5Section 3.3.6.1
Section 3.3.6.3 and Section 3.3.8
- o Attempted to clarify the substitution mechanism in Section 3.3.7 and improve its uniqueness.
- o Made various editorial changes, including changing the title in the header, and removing numbering from some SDP examples.

7.3. Changes from version 09

- o Additional corrections to latent media stream example in Section 4.3
- o Fixed up attribute formatting examples and corresponding ABNF.
- o Removed preference rule for latent configurations.

- o Various spelling and other editorial changes were made.
- o updated crossreferences.

7.4. Changes from version 08

The major change is in Section 4.3, Latent Media Streams, fixing the syntax of the answer. All the other changes are editorial.

7.5. Changes from version 04

- o The definitions for bcap, ccap, icap, and kcap attributes have been removed, and are to be defined in another document.
- o Corrected formatting of m= and p= configuration parameters to conform to extension-config-list form defined in [RFC5939]
- o Reorganized definitions of new parameters to make them easier to find in document.
- o Added ability to renegotiate capabilities when modifying the session (Section 3.4.4).
- o Made various editorial changes, clarifications, and typo corrections.

7.6. Changes from version 03

- o A new session capability attribute (sescap) has been added to permit specification of acceptable media stream combinations.
- o Capability attribute definitions corresponding to the i, c, b, and k SDP line types have been added for completeness.
- o Use of the pcfg: attribute in SDP answers has been included in order to conveniently return information in the answer about acceptable configurations in the media stream offer.
- o The use of the lcfg: attribute(s) in SDP answers has been restricted to indicate just which latent configuration offers would be acceptable to the answerer.
- o A suggestion for "naive" middleboxes has been added to the Security Considerations.
- o Various editorial changes have been made.

- o Several errors/omissions have been corrected.
- o The description of the mscap attribute has been modified to make it clear that it should not be used to generate undefined SDP attributes, or to "extend" existing attributes.
- o <ms-parameters> are made optional in the mscap attribute definition.
- o "AMR" changed to "AMR-WB" in cases in which the sample rate is 16000.

7.7. Changes from version 02

This version contains several detail changes intended to simplify capability processing and mapping into conventional SDP media blocks.

- o The "mcap" attribute is enhanced to include the role of the "ecap" attribute; the latter is eliminated.
- o The "fcap" attribute has been renamed "mfcap". New replacement rules vis-a-vis fmp attributes in the base media specification have been added.
- o A new "mscap" attribute is defined to handle the problem of attributes (other than rtpmap and fmp) that are specific to a particular payload type.
- o New rules for processing the mcap, mfcap, and mscap attributes, and overriding standard rtpmap, fmp, or other media-specific attributes, are put forward to reduce the need to use the deletion option in the a= parameter of the potential configuration (pcfg) attribute.
- o A new parameter, "mt=" is added to the latent configuration attribute (lcfg) to specify the media stream type (audio, video, etc.) when the lcfg is declared at the session level.
- o The examples are expanded.
- o Numerous typos and misspellings have been corrected.

7.8. Changes from version 01

The documents adds a new attribute for specifying bandwidth capability and a parametr to list in the potential configuration. Other changes are to align the document with the terminolgy and attribute names from draft-ietf-mmusic-sdp-capability-negotiation-07.

The document also clarifies some previous open issues.

7.9. Changes from version 00

The major changes include taking out the "mcap" and "cptmap" parameter. The mapping of payload type is now in the "pt" parameter of "pcfg". Media subtype need to explicitly defined in the "cmed" attribute if referenced in the "pcfg"

8. Acknowledgements

This document is heavily influenced by the discussions and work done by the SDP Capability Negotiation Design team. The following people in particular provided useful comments and suggestions to either the document itself or the overall direction of the solution defined herein: Cullen Jennings, Matt Lepinski, Joerg Ott, Colin Perkins, and Thomas Stach.

We thank Ingemar Johansson and Magnus Westerlund for examples that stimulated this work, and for critical reading of the document.

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.
- [RFC5939] Andreasen, F., "SDP Capability Negotiation", RFC 5939, September 2010.

9.2. Informative References

- [RFC4568] Andreasen, F., Baugher, M., and D. Wing, "Session Description Protocol (SDP) Security Descriptions for Media Streams", RFC 4568, July 2006.
- [RFC4583] Camarillo, G., "Session Description Protocol (SDP) Format for Binary Floor Control Protocol (BFCP) Streams", RFC 4583, November 2006.
- [RFC4585] Ott, J., Wenger, S., Sato, N., Burmeister, C., and J. Rey, "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)", RFC 4585, July 2006.
- [RFC4733] Schulzrinne, H. and T. Taylor, "RTP Payload for DTMF Digits, Telephony Tones, and Telephony Signals", RFC 4733, December 2006.
- [RFC4867] Sjoberg, J., Westerlund, M., Lakaniemi, A., and Q. Xie, "RTP Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs", RFC 4867, April 2007.
- [RFC5104] Wenger, S., Chandra, U., Westerlund, M., and B. Burman, "Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)", RFC 5104, February 2008.

Authors' Addresses

Robert R Gilman
Independent
3243 W. 11th Ave. Dr.
Broomfield, CO 80020
USA

Email: bob_gilman@comcast.net

Roni Even
Gesher Erova Ltd
14 David Hamelech
Tel Aviv 64953
Israel

Email: ron.even.tlv@gmail.com

Flemming Andreassen
Cisco Systems
Iselin, NJ
USA

Email: fandreas@cisco.com

MMUSIC
Internet-Draft
Updates: 5939 (if approved)
Intended status: Standards Track
Expires: July 8, 2013

R. Gilman
Independent
R. Even
Gesher Erove Ltd
F. Andreassen
Cisco Systems
January 4, 2013

Session Description Protocol (SDP) Media Capabilities Negotiation
draft-ietf-mmusic-sdp-media-capabilities-17

Abstract

Session Description Protocol (SDP) capability negotiation provides a general framework for indicating and negotiating capabilities in SDP. The base framework defines only capabilities for negotiating transport protocols and attributes. In this document, we extend the framework by defining media capabilities that can be used to negotiate media types and their associated parameters.

This document updates the IANA Considerations of RFC 5939.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on July 8, 2013.

Copyright Notice

Copyright (c) 2013 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of

publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

This document may contain material from IETF Documents or IETF Contributions published or made publicly available before November 10, 2008. The person(s) controlling the copyright in some of this material may not have granted the IETF Trust the right to allow modifications of such material outside the IETF Standards Process. Without obtaining an adequate license from the person(s) controlling the copyright in such materials, this document may not be modified outside the IETF Standards Process, and derivative works of it may not be created outside the IETF Standards Process, except to format it for publication as an RFC or to translate it into languages other than English.

Table of Contents

1. Introduction	5
2. Terminology	6
3. SDP Media Capabilities	8
3.1. Requirements	8
3.2. Solution Overview	9
3.3. New Capability Attributes	15
3.3.1. The Media Format Capability Attributes	15
3.3.2. The Media Format Parameter Capability Attribute	18
3.3.3. The Media-Specific Capability Attribute	20
3.3.4. New Configuration Parameters	22
3.3.5. The Latent Configuration Attribute	24
3.3.6. Enhanced Potential Configuration Attribute	26
3.3.7. Substitution of Media Payload Type Numbers in Capability Attribute Parameters	30
3.3.8. The Session Capability Attribute	31
3.4. Offer/Answer Model Extensions	35
3.4.1. Generating the Initial Offer	36
3.4.2. Generating the Answer	39
3.4.3. Offerer Processing of the Answer	43
3.4.4. Modifying the Session	44
4. Examples	45
4.1. Alternative Codecs	45
4.2. Alternative Combinations of Codecs (Session Configurations)	48
4.3. Latent Media Streams	48
5. IANA Considerations	51
5.1. New SDP Attributes	51
5.2. New SDP Capability Negotiation Option Tag	52
5.3. SDP Capability Negotiation Configuration Parameters Registry	52
5.4. SDP Capability Negotiation Configuration Parameter Registrations	53
6. Security Considerations	55
7. Changes from previous versions	56
7.1. Changes from version 16	56
7.2. Changes from version 15	56
7.3. Changes from version 14	56
7.4. Changes from version 13	56
7.5. Changes from version 12	56
7.6. Changes from version 11	57
7.7. Changes from version 10	57
7.8. Changes from version 09	58
7.9. Changes from version 08	58
7.10. Changes from version 04	58
7.11. Changes from version 03	58
7.12. Changes from version 02	59

7.13. Changes from version 01	60
7.14. Changes from version 00	60
8. Acknowledgements	61
9. References	62
9.1. Normative References	62
9.2. Informative References	62
Authors' Addresses	64

1. Introduction

Session Description Protocol (SDP) capability negotiation [RFC5939] provides a general framework for indicating and negotiating capabilities in SDP [RFC4566]. The base framework defines only capabilities for negotiating transport protocols and attributes.

RFC 5939 [RFC5939] lists some of the issues with the current SDP capability negotiation process. An additional real life case is to be able to offer one media stream (e.g. audio) but list the capability to support another media stream (e.g. video) without actually offering it concurrently.

In this document, we extend the framework by defining media capabilities that can be used to indicate and negotiate media types and their associated format parameters. This document also adds the ability to declare support for media streams, the use of which can be offered and negotiated later, and the ability to specify session configurations as combinations of media stream configurations. The definitions of new attributes for media capability negotiation are chosen to make the translation from these attributes to "conventional" SDP [RFC4566] media attributes as straightforward as possible in order to simplify implementation. This goal is intended to reduce processing in two ways: each proposed configuration in an offer may be easily translated into a conventional SDP media stream record for processing by the receiver; and the construction of an answer based on a selected proposed configuration is straightforward.

This document updates RFC 5939 [RFC5939] by updating the IANA Considerations. All other extensions defined in this document are considered extensions above and beyond RFC 5939 [RFC5939].

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119] and indicate requirement levels for compliant implementations.

"Actual Configuration": An actual configuration specifies which combinations of SDP session parameters and media stream components can be used in the current offer/answer exchange and with what parameters. Use of an actual configuration does not require any further negotiation in the offer/answer exchange. See RFC 5939 [RFC5939] for further details.

"Base Attributes": Conventional SDP attributes appearing in the base configuration of a media block.

"Base Configuration": The media configuration represented by a media block exclusive of all the capability negotiation attributes defined in this document, the base capability negotiation document [RFC5939], or any other capability negotiation document. In an offer SDP, the base configuration corresponds to the actual configuration as defined in RFC 5939 [RFC5939].

"Conventional Attribute": Any SDP attribute other than those defined by the series of capability negotiation specifications.

"Conventional SDP": An SDP record devoid of capability negotiation attributes.

"Media Format Capability": A media format, typically a media subtype such as PCMU, H263-1998, or T38, expressed in the form of a capability.

"Media Format Parameter Capability": A media format parameter ("a=fmtp" in conventional SDP) expressed in the form of a capability. The media format parameter capability is associated with a media format capability.

"Media Capability": The combined set of capabilities associated with expressing a media format and its relevant parameters (e.g. media format parameters and media specific parameters).

"Potential Configuration": A potential configuration indicates which combinations of capabilities can be used for the session and its associated media stream components. Potential configurations are not ready for use, however they are offered for potential use in the current offer/answer exchange. They provide an alternative that may

be used instead of the actual configuration, subject to negotiation in the current offer/answer exchange. See RFC 5939 [RFC5939] for further details.

"Latent Configuration": A latent configuration indicates which combinations of capabilities could be used in a future negotiation for the session and its associated media stream components. Latent configurations are neither ready for use, nor are they offered for actual or potential use in the current offer/answer exchange. Latent configurations merely inform the other side of possible configurations supported by the entity. Those latent configurations may be used to guide subsequent offer/answer exchanges, but they are not offered for use as part of the current offer/answer exchange.

3. SDP Media Capabilities

The SDP capability negotiation [RFC5939] discusses the use of any SDP [RFC4566] attribute (a=) under the attribute capability "acap". The limitations of using acap for ftmp and rtpmap in a potential configuration are described in RFC 5939 [RFC5939]; for example they can be used only at the media level since they are media level attributes. RFC 5939 [RFC5939] does not provide a way to exchange media-level capabilities prior to the actual offer of the associated media stream. This section provides an overview of extensions providing an SDP Media Capability negotiation solution offering more robust capabilities negotiation. This is followed by definitions of new SDP attributes for the solution and its associated updated offer/answer procedures [RFC3264]

3.1. Requirements

The capability negotiation extensions requirements considered herein are as follows.

- REQ-01: Support the specification of alternative (combinations of) media formats (codecs) in a single media block.
- REQ-02: Support the specification of alternative media format parameters for each media format.
- REQ-03: Retain backward compatibility with conventional SDP. Ensure that each and every offered configuration can be easily translated into a corresponding SDP media block expressed with conventional SDP lines.
- REQ-04: Ensure the scheme operates within the offer/answer model in such a way that media formats and parameters can be agreed upon with a single exchange.
- REQ-05: Provide the ability to express offers in such a way that the offerer can receive media as soon as the offer is sent. (Note that the offerer may not be able to render received media prior to exchange of keying material.)
- REQ-06: Provide the ability to offer latent media configurations for future negotiation.
- REQ-07: Provide reasonable efficiency in the expression of alternative media formats and/or format parameters, especially in those cases in which many combinations of options are offered.

REQ-08: Retain the extensibility of the base capability negotiation mechanism.

REQ-09: Provide the ability to specify acceptable combinations of media streams and media formats. For example, offer a PCMU audio stream with an H264 video stream, or a G729 audio stream with an H263 video stream. This ability would give the offerer a means to limit processing requirements for simultaneous streams. This would also permit an offer to include the choice of an audio/T38 stream or an image/T38 stream, but not both.

Other possible extensions have been discussed, but have not been treated in this document. They may be considered in the future. Three such extensions are:

FUT-01: Provide the ability to mix, or change, media types within a single media block. Conventional SDP does not support this capability explicitly; the usual technique is to define a media subtype that represents the actual format within the nominal media type. For example, T.38 FAX as an alternative to audio/PCMU within an audio stream is identified as audio/T38; a separate FAX stream would use image/T38.

FUT-02: Provide the ability to support multiple transport protocols within an active media stream without reconfiguration. This is not explicitly supported by conventional SDP.

FUT-03: Provide capability negotiation attributes for all media-level SDP line types in the same manner as already done for the attribute type, with the exception of the media line type itself. The media line type is handled in a special way to permit compact expression of media coding/format options. The line types are bandwidth ("b="), information ("i="), connection data ("c="), and, possibly, the deprecated encryption key ("k=").

3.2. Solution Overview

The solution consists of new capability attributes corresponding to conventional SDP line types, new parameters for the pcfg, acfg, and the new lcfg attributes extending the base attributes from RFC 5939 [RFC5939], and a use of the pcfg attribute to return capability information in the SDP answer.

Several new attributes are defined in a manner that can be related to the capabilities specified in a media line, and its corresponding rtpmap and fmpmap attributes.

- o A new attribute ("a=rmcap") defines RTP-based media format capabilities in the form of a media subtype (e.g. "PCMU"), and its encoding parameters (e.g. "/8000/2"). Each resulting media format type/subtype capability has an associated handle called a media capability number. The encoding parameters are as specified for the rtpmap attribute defined in SDP [RFC4566], without the payload type number part.
- o A new attribute ("a=omcap") defines other (non RTP-based) media format capabilities in the form of a media subtype only (e.g. "T38"). Each resulting media format type/subtype capability has an associated handle called a media capability number.
- o A new attribute ("a=mfcap") specifies media format parameters associated with one or more media format capabilities. The mfcap attribute is used primarily to associate the media format parameters normally carried in the fmpv attribute. Note that media format parameters can be used with RTP and non-RTP based media formats.
- o A new attribute ("a=mscap") that specifies media parameters associated with one or more media format capabilities. The mscap attribute is used to associate capabilities with attributes other than fmpv or rtpmap, for example, the rtcp-fb attribute defined in RFC 4585 [RFC4585].
- o A new attribute ("a=lcfg") specifies latent media stream configurations when no corresponding media line ("m=") is offered. An example is the offer of latent configurations for video even though no video is currently offered. If the peer indicates support for one or more offered latent configurations, the corresponding media stream(s) may be added via a new offer/answer exchange.
- o A new attribute ("a=sescap") is used to specify an acceptable combination of simultaneous media streams and their configurations as a list of potential and/or latent configurations.

New parameters are defined for the potential configuration (pcfg), latent configuration (lcfg), and accepted configuration (acfg) attributes to associate the new attributes with particular configurations.

- o A new parameter type ("m=") is added to the potential configuration ("a=pcfg:") attribute and the actual configuration ("a=acfg:") attribute defined in RFC 5939 [RFC5939], and to the new latent configuration ("a=lcfg:") attribute. This permits specification of media capabilities (including their associated

parameters) and combinations thereof for the configuration. For example, the "a=pcfg:" line might specify PCMU and telephone events [RFC4733] or G.729B and telephone events as acceptable configurations. The "a=acfg:" line in the answer would specify the configuration chosen.

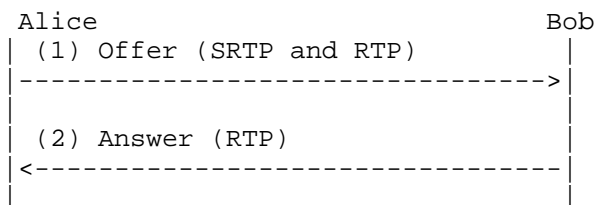
- o A new parameter type ("pt=") is added to the potential configuration, actual configuration, and latent configuration attributes. This parameter associates RTP payload type numbers with the referenced RTP-based media format capabilities, and is appropriate only when the transport protocol uses RTP.
- o A new parameter type ("mt=") is used to specify the media type for latent configurations.

Special processing rules are defined for capability attribute arguments in order to reduce the need to replicate essentially-identical attribute lines for the base configuration and potential configurations.

- o A substitution rule is defined for any capability attribute to permit the replacement of the (escaped) media capability number with the media format identifier (e.g., the payload type number in audio/video profiles).
- o Replacement rules are defined for the conventional SDP equivalents of the mfcap and mscap capability attributes. This reduces the necessity to use the deletion qualifier in the a=pcfg parameter in order to ignore rtpmap, fmtp, and certain other attributes in the base configuration.
- o An argument concatenation rule is defined for mfcap attributes which refer to the same media capability number. This makes it convenient to combine format options concisely by associating multiple mfcap lines with multiple media format capabilities.

This document extends the base protocol extensions to the offer/answer model that allow for capabilities and potential configurations to be included in an offer. Media capabilities constitute capabilities that can be used in potential and latent configurations. Whereas potential configurations constitute alternative offers that may be accepted by the answerer instead of the actual configuration(s) included in the "m=" line(s) and associated parameters, latent configurations merely inform the other side of possible configurations supported by the entity. Those latent configurations may be used to guide subsequent offer/answer exchanges, but they are not part of the current offer/answer exchange.

The mechanism is illustrated by the offer/answer exchange below, where Alice sends an offer to Bob:



Alice's offer includes RTP and SRTP as alternatives. RTP is the default, but SRTP is the preferred one (long lines are folded to fit the margins):

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 3456 RTP/AVP 0 18
a=tcap:1 RTP/SAVP RTP/AVP
a=rtpmap:0 PCMU/8000/1
a=rtpmap:18 G729/8000/1
a=fmtp:18 annexb=yes
a=rmcap:1,4 G729/8000/1
a=rmcap:2 PCMU/8000/1
a=rmcap:5 telephone-event/8000
a=mfcap:1 annexb=no
a=mfcap:4 annexb=yes
a=mfcap:5 0-11
a=acap:1 crypto:1 AES_CM_128_HMAC_SHA1_32 \
inline:NzB4dlBINUAvLEw6UzF3WSJ+PSdFcGdUJShpXlZj|2^20|1:32
a=pcfg:1 m=4,5|1,5 t=1 a=1 pt=1:100,4:101,5:102
a=pcfg:2 m=2 t=1 a=1 pt=2:103
a=pcfg:3 m=4 t=2 pt=4:18
```

The required base and extensions are provided by the "a=creq" attribute defined in RFC 5939 [RFC5939], with the option tag "med-v0", which indicates that the extension framework defined here, must be supported. The base level capability negotiation support ("cap-v0" [RFC5939]) is implied since it is required for the extensions.

The "m=" line indicates that Alice is offering to use plain RTP with PCMU or G.729B. The media line implicitly defines the default

transport protocol (RTP/AVP in this case) and the default actual configuration.

The "a=tcap:1" line, specified in the SDP Capability Negotiation base protocol [RFC5939], defines transport protocol capabilities, in this case Secure RTP (SAVP profile) as the first option and RTP (AVP profile) as the second option.

The "a=rmcap:1,4" line defines two G.729 RTP-based media format capabilities, numbered 1 and 4, and their encoding rate. The capabilities are of media type "audio" and subtype G729. Note that the media subtype is explicitly specified here, rather than RTP payload type numbers. This permits the assignment of payload type numbers in the media stream configuration specification. In this example, two G.729 subtype capabilities are defined. This permits the declaration of two sets of formatting parameters for G.729.

The "a=rmcap:2" line defines a G.711 mu-law capability, numbered 2.

The "a=rmcap:5" line defines an audio telephone-event capability, numbered 5.

The "a=mfcap:1" line specifies the fmltp formatting parameters for capability 1 (offerer will not accept G.729 Annex B packets).

The "a=mfcap:4" line specifies the fmltp formatting parameters for capability 4 (offerer will accept G.729 Annex B packets).

The "a=mfcap:5" line specifies the fmltp formatting parameters for capability 5 (the DTMF touchtones 0-9,*,#).

The "a=acap:1" line specified in the base protocol provides the "crypto" attribute which provides the keying material for SRTP using SDP security descriptions.

The "a=pcfg:" attributes provide the potential configurations included in the offer by reference to the media capabilities, transport capabilities, attribute capabilities and specified payload type number mappings. Three explicit alternatives are provided; the lowest-numbered one is the preferred one. The "a=pcfg:1 ..." line specifies media capabilities 4 and 5, i.e., G.729B and DTMF (incl. their associated media format parameters), or media capability 1 and 5, i.e., G.729 and DTMF (incl. their associated media format parameters). Furthermore, it specifies transport protocol capability 1 (i.e. the RTP/SAVP profile - secure RTP), and the attribute capability 1, i.e. the crypto attribute provided. Lastly, it specifies a payload type number mapping for (RTP-based) media capabilities 1, 4, and 5, thereby permitting the offerer to

distinguish between encrypted media and unencrypted media received prior to receipt of the answer.

Use of unique payload type numbers in alternative configurations is not required; codecs such as AMR-WB [RFC4867] have the potential for so many combinations of options that it may be impractical to define unique payload type numbers for all supported combinations. If unique payload type numbers cannot be specified, then the offerer will be obliged to wait for the SDP answer before rendering received media. For SRTP using SDES inline keying [RFC4568], the offerer will still need to receive the answer before being able to decrypt the stream.

The second alternative ("a=pcfg:2 ...") specifies media capability 2, i.e., PCMU, under the RTP/SAVP profile, with the same SRTP key material.

The third alternative ("a=pcfg:3 ...") offers G.729B unsecured; its only purpose in this example is to show a preference for G.729B over PCMU.

Per RFC 5939 [RFC5939], the media line, with any qualifying attributes such as fmtp or rtpmap, is itself considered a valid configuration (the current actual configuration); it has the lowest preference (per RFC 5939 [RFC5939]).

Bob receives the SDP offer from Alice. Bob supports G.729B, PCMU, and telephone events over RTP, but not SRTP, hence he accepts the potential configuration 3 for RTP provided by Alice. Bob generates the following answer:

```
v=0
o=- 24351 621814 IN IP4 192.0.2.2
s=
c=IN IP4 192.0.2.2
t=0 0
a=csup:med-v0
m=audio 4567 RTP/AVP 18
a=rtpmap:18 G729/8000
a=fmtp:18 annexb=yes
a=acfg:3 m=4 t=2 pt=4:18
```

Bob includes the "a=csup" and "a=acfg" attributes in the answer to inform Alice that he can support the med-v0 level of capability negotiations. Note that in this particular example, the answerer supported the capability extensions defined here, however had he not, he would simply have processed the offer based on the offered PCMU and G.729 codecs under the RTP/AVP profile only. Consequently, the

answer would have omitted the "a=csup" attribute line and chosen one or both of the PCMU and G.729 codecs instead. The answer carries the accepted configuration in the "m=" line along with corresponding rtpmap and/or fmpmap parameters, as appropriate.

Note that per the base protocol, after the above, Alice MAY generate a new offer with an actual configuration ("m=" line, etc.) corresponding to the actual configuration referenced in Bob's answer (not shown here).

3.3. New Capability Attributes

In this section, we present the new attributes associated with indicating the media capabilities for use by the SDP Capability negotiation. The approach taken is to keep things similar to the existing media capabilities defined by the existing media descriptions ("m=" lines) and the associated "rtpmap" and "fmpmap" attributes. We use media subtypes and "media capability numbers" to link the relevant media capability parameters. This permits the capabilities to be defined at the session level and be used for multiple streams, if desired. For RTP-based media formats, payload types are then specified at the media level (see Section 3.3.4.2).

A media capability merely indicates possible support for the media type and media format(s) and parameters in question. In order to actually use a media capability in an offer/answer exchange, it MUST be referenced in a potential configuration.

Media capabilities, i.e. the attributes associated with expressing media capability formats, parameters, etc., can be provided at the session-level and/or the media-level. Media capabilities provided at the session level may be referenced in any pcfp or lcfp attribute at the media level (consistent with the media type), whereas media capabilities provided at the media level may be referenced only by the pcfp or lcfp attribute within that media stream. In either case, the scope of the <med-cap-num> is the entire session description. This enables each media capability to be uniquely referenced across the entire session description (e.g. in a potential configuration).

3.3.1. The Media Format Capability Attributes

Media subtypes can be expressed as media format capabilities by use of the "a=rmcap" and "a=omcap" attributes. The "a=rmcap" attribute MUST be used for RTP-based media whereas the "a=omcap" attribute MUST be used for non-RTP-based (other) media formats. The two attributes are defined as follows:

```
a=rmcap:<media-cap-num-list> <encoding-name>/<clock-rate>
[</encoding-parms>]
```

```
a=omcap:<media-cap-num-list> <format-name>
```

where <media-cap-num-list> is a (list of) media capability number(s) used to number a media format capability, the <encoding name> or <format name> is the media subtype, e.g., H263-1998, PCMU, or T38, <clock rate> is the encoding rate, and <encoding parms> are the media encoding parameters for the media subtype. All media format capabilities in the list are assigned to the same media type/subtype. Each occurrence of the rmcap and omcap attribute MUST use unique values in their <media-cap-num-list>; the media capability numbers are shared between the two attributes and the numbers MUST be unique across the entire SDP session. In short, the rmcap and omcap attributes define media format capabilities and associate them with a media capability number in the same manner as the rtpmap attribute defines them and associates them with a payload type number. Additionally, the attributes allow multiple capability numbers to be defined for the media format in question by specifying a range of media capability numbers. This permits the media format to be associated with different media parameters in different configurations. When a range of capability numbers is specified, the first (leftmost) capability number MUST be strictly smaller than the second (rightmost), i.e. the range increases and covers at least two numbers.

In ABNF [RFC5234], we have:

```
media-capability-line = rtp-mcap / non-rtp-mcap

rtp-mcap               = "a=rmcap:" media-cap-num-list
                        1*WSP encoding-name "/" clock-rate
                        ["/" encoding-parms]
non-rtp-mcap           = "a=omcap:" media-cap-num-list 1*WSP format-name
media-cap-num-list     = media-cap-num-element
                        *("," media-cap-num-element)
media-cap-num-element  = media-cap-num
                        / media-cap-num-range
media-cap-num-range    = media-cap-num "-" media-cap-num
media-cap-num          = NonZeroDigit *9(DIGIT)
encoding-name          = token ;defined in RFC4566
clock-rate             = NonZeroDigit *9(DIGIT)
encoding-parms         = token
format-name            = token ;defined in RFC4566
NonZeroDigit           = %x31-39      ; 1-9
```

The encoding-name, clock-rate and encoding-params are as defined to

appear in an rtpmap attribute for each media type/subtype. Thus, it is easy to convert an rmcap attribute line into one or more rtpmap attribute lines, once a payload type number is assigned to a media-cap-num (see Section 3.3.5).

The format-name is a media format description for non-RTP based media as defined for the <fmt> part of the media description ("m=" line) in SDP [RFC4566]. In simple terms, it is the name of the media format, e.g. "t38". This form can also be used in cases such as BFCP [RFC4585] where the fmt list in the m-line is effectively ignored (BFCP uses "*").

The "rmcap" and "omcap" attributes can be provided at the session-level and/or the media-level. There can be more than one rmcap and more than one omcap attribute at both the session and media level (i.e., more than one of each at the session-level and more than one of each in each media description). Media capability numbers cannot include leading zeroes, and each media-cap-num MUST be unique within the entire SDP record; it is used to identify that media capability in potential, latent and actual configurations, and in other attribute lines as explained below. Note that the media-cap-num values are shared between the rmcap and omcap attributes, and hence the uniqueness requirement applies to the union of them. When the media capabilities are used in a potential, latent or actual configuration, the media formats referred by those configurations apply at the media level, irrespective of whether the media capabilities themselves were specified at the session or media level. In other words, the media capability applies to the specific media description associated with the configuration which invokes it.

For example:

```
v=0
o=- 24351 621814 IN IP4 192.0.2.2
s=
c=IN IP4 192.0.2.2
t=0 0
a=rmcap:1 L16/8000/1
a=rmcap:2 L16/16000/2
a=rmcap:3 H263-1998/90000
a=omcap:4 example
m=audio 54320 RTP/AVP 0
a=pcfg:1 m=1|2, pt=1:99,2:98
m=video 66544 RTP/AVP 100
a=rtpmap:100 H264/90000
a=pcfg:10 m=3 pt=3:101
a=tcap:1 TCP
a=pcfg:11 m=4 t=1
```

3.3.2. The Media Format Parameter Capability Attribute

This attribute is used to associate media format specific parameters with one or more media format capabilities. The form of the attribute is:

```
a=mfcap:<media-caps> <list of parameters>
```

where <media-caps> permits the list of parameters to be associated with one or more media format capabilities and the format parameters are specific to the type of media format. The mfcap lines map to a single traditional SDP fmt attribute line (one for each entry in <media-caps>) of the form

```
a=fmt:<fmt> <list of parameters>
```

where <fmt> is the media format parameter defined in RFC 4566 [RFC4566], as appropriate for the particular media stream. The mfcap attribute MUST be used to encode attributes for media capabilities, which would conventionally appear in an fmt attribute. The existing acap attribute MUST NOT be used to encode fmt attributes.

The mfcap attribute adheres to SDP [RFC4566] attribute production rules with

```
media-format-parameter-capability =  
    "a=mfcap:" media-cap-num-list 1*WSP fmt-specific-param-list  
    fmt-specific-param-list = text ; defined in RFC4566
```

Note that media format parameters can be used with RTP-based and non-RTP based media formats.

3.3.2.1. Media Format Parameter Concatenation Rule

The appearance of media subtypes with a large number of formatting options (e.g., AMR-WB [RFC4867]) coupled with the restriction that only a single fmt attribute can appear per media format, suggests that it is useful to create a combining rule for mfcap parameters which are associated with the same media capability number. Therefore, different mfcap lines MAY include the same media-cap-num in their media-cap-num-list. When a particular media capability is selected for processing, the parameters from each mfcap line which references the particular capability number in its media-cap-num-list are concatenated together via ";", in the order the mfcap attributes appear in the SDP record, to form the equivalent of a single fmt attribute line. This permits one to define a separate mfcap line for a single parameter and value that is to be applied to each media capability designated in the media-cap-num-list. This provides a

compact method to specify multiple combinations of format parameters when using codecs with multiple format options. Note that order-dependent parameters SHOULD be placed in a single mfcap line to avoid possible problems with line rearrangement by a middlebox.

Format parameters are not parsed by SDP; their content is specific to the media type/subtype. When format parameters for a specific media capability are combined from multiple a=mfcap lines which reference that media capability, the format-specific parameters are concatenated together and separated by ";" for construction of the corresponding format attribute (a=fmtp). The resulting format attribute will look something like the following (without line breaks):

```
a=fmtp:<fmt> <fmt-specific-param-list1>;  
          <fmt-specific-param-list2>;  
          ...
```

where <fmt> depends on the transport protocol in the manner defined in RFC4566. SDP cannot assess the legality of the resulting parameter list in the "a=fmtp" line; the user must take care to ensure that legal parameter lists are generated.

The "mfcap" attribute can be provided at the session-level and the media-level. There can be more than one mfcap attribute at the session or media level. The unique media-cap-num is used to associate the parameters with a media capability.

As a simple example, a G.729 capability is, by default, considered to support comfort noise as defined by Annex B. Capabilities for G.729 with and without comfort noise support may thus be defined by:

```
a=rmcap:1,2 G729/8000  
a=mfcap:2 annexb:no
```

Media capability 1 supports G.729 with Annex B, whereas media capability 2 supports G.729 without Annex B.

Example for H.263 video:

```
a=rmcap:1 H263-1998/90000  
a=rmcap:2 H263-2000/90000  
a=mfcap:1 CIF=4;QCIF=2;F=1;K=1  
a=mfcap:2 profile=2;level=2.2
```

Finally, for six format combinations of the Adaptive MultiRate codec:

```
a=rmcap:1-3 AMR/8000/1
a=rmcap:4-6 AMR-WB/16000/1
a=mfcap:1,2,3,4 mode-change-capability=1
a=mfcap:5,6 mode-change-capability=2
a=mfcap:1,2,3,5 max-red=220
a=mfcap:3,4,5,6 octet-align=1
a=mfcap:1,3,5 mode-set=0,2,4,7
a=mfcap:2,4,6 mode-set=0,3,5,6
```

So that AMR codec #1, when specified in a pcfg attribute within an audio stream block (and assigned payload type number 98) as in

```
a=pcfg:1 m=1 pt=1:98
```

is essentially equivalent to the following

```
m=audio 49170 RTP/AVP 98
a=rtpmap:98 AMR/8000/1
a=fmtp:98 mode-change-capability=1; \
max-red=220; mode-set=0,2,4,7
```

and AMR codec #4 with payload type number 99, depicted by the potential configuration:

```
a=pcfg:4 m=4, pt=4:99
```

is equivalent to the following:

```
m=audio 49170 RTP/AVP 99
a=rtpmap:99 AMR-WB/16000/1
a=fmtp:99 mode-change-capability=1; octet-align=1; \
mode-set=0,3,5,6
```

and so on for the other four combinations. SDP could thus convert the media capabilities specifications into one or more alternative media stream specifications, one of which can be chosen for the answer.

3.3.3. The Media-Specific Capability Attribute

Attributes and parameters associated with a media format are typically specified using the "rtpmap" and "fmtp" attributes in SDP, and the similar "rmcap" and "mfcap" attributes in SDP Media Capabilities. Some SDP extensions define other attributes that need to be associated with media formats, for example the "rtcp-fb" attribute defined in RFC 4585 [RFC4585]. Such media-specific attributes, beyond the rtpmap and fmtp attributes, may be associated with media capability numbers via a new media-specific attribute,

mscap, of the following form:

```
a=mscap:<media caps star> <att field> <att value>
```

where <media caps star> is a (list of) media capability number(s), <att field> is the attribute name, and <att value> is the value field for the named attribute. Note that the media capability numbers refer to media format capabilities specified elsewhere in the SDP ("rmcap" and/or "omcap"). If a range of capability numbers is specified, the first (leftmost) capability number MUST be strictly smaller than the second (rightmost). The media capability numbers may include a wildcard ("*"), which will be used instead of any payload type mappings in the resulting SDP (see, e.g. RFC 4585 [RFC4585] and the example below). In ABNF, we have:

```
media-specific-capability = "a=mscap:"
                           media-caps-star
                           1*WSP att-field ; from RFC4566
                           1*WSP att-value ; from RFC4566
media-caps-star            = media-cap-star-element
                           *("," media-cap-star-element)
media-cap-star-element    = (media-cap-num [wildcard])
                           / (media-cap-num-range [wildcard])
wildcard                  = "*"
```

Given an association between a media capability and a payload type number as specified by the pt= parameters in a pcfg attribute line, a mscap line may be translated easily into a conventional SDP attribute line of the form

```
a=<att field>:"<fmt> <att value> ; <fmt> defined in SDP
[RFC4566]
```

A resulting attribute that is not a legal SDP attribute as specified by RFC4566 MUST be ignored by the receiver.

If a media capability number (or range) contains a wildcard character at the end, any payload type mapping specified for that media specific capability (or range of capabilities) will use the wildcard character in the resulting SDP instead of the payload type specified in the payload type mapping ("pt" parameter) in the configuration attribute.

A single mscap line may refer to multiple media capabilities by use of a capability number range; this is equivalent to multiple mscap lines, each with the same attribute values (but different media capability numbers), one line per media capability.

Multiple mscap lines may refer to the same media capability, but, unlike the mfcap attribute, no concatenation operation is defined. Hence, multiple mscap lines applied to the same media capability is equivalent to multiple lines of the specified attribute in a conventional media record.

Here is an example with the rtcp-fb attribute, modified from an example in RFC 5104 [RFC5104] (with the session-level and audio media omitted). If the offer contains a media block like the following (note the wildcard character),

```
m=video 51372 RTP/AVP 98
a=rtpmap:98 H263-1998/90000
a=tcap:1 RTP/AVPF
a=rmcap:1 H263-1998/90000
a=mscap:1 rtcp-fb ccm tstr
a=mscap:1 rtcp-fb ccm fir
a=mscap:1* rtcp-fb ccm tmmbr smaxpr=120
a=pcfg:1 t=1 m=1 pt=1:98
```

and if the proposed configuration is chosen, then the equivalent media block would look like

```
m=video 51372 RTP/AVPF 98
a=rtpmap:98 H263-1998/90000
a=rtcp-fb:98 ccm tstr
a=rtcp-fb:98 ccm fir
a=rtcp-fb:* ccm tmmbr smaxpr=120
```

3.3.4. New Configuration Parameters

Along with the new attributes for media capabilities, new extension parameters are defined for use in the potential configuration, the actual configuration, and/or the new latent configuration defined in Section 3.3.5.

3.3.4.1. The Media Configuration Parameter (m=)

The media configuration parameter is used to specify the media format(s) and related parameters for a potential, actual, or latent configuration. Adhering to the ABNF for extension-config-list in RFC 5939 [RFC5939] with

```
ext-cap-name = "m"
ext-cap-list = media-cap-num-list
               [*(BAR media-cap-num-list)]
```

we have

```

media-config-list = ["+"] "m=" media-cap-num-list
                    *(BAR media-cap-num-list)
                    ;BAR is defined in RFC5939
                    ;media-cap-num-list is defined above

```

Alternative media configurations are separated by a vertical bar ("|"). The alternatives are ordered by preference, most-preferred first. When media capabilities are not included in a potential configuration at the media level, the media type and media format from the associated "m=" line will be used. The use of the plus sign ("+") is described in RFC5939.

3.3.4.2. The Payload Type Number Mapping Parameter (pt=)

The payload type number mapping parameter is used to specify the payload type number to be associated with each RTP-based media format in a potential, actual, or latent configuration. We define the payload type number mapping parameter, payload-number-config-list, in accordance with the extension-config-list format defined in RFC 5939 [RFC5939]. In ABNF:

```

payload-number-config-list = ["+"] "pt=" media-map-list
media-map-list             = media-map *("," media-map)
media-map                  = media-cap-num ":" payload-type-number
                           ; media-cap-num is defined in 3.3.1
payload-type-number = NonZeroDigit *2(DIGIT) ; RTP payload
                                                ; type number

```

The example in Section 3.3.7 shows how the parameters from the rmcap line are mapped to payload type numbers from the pcfg "pt" parameter. The use of the plus sign ("+") is described in RFC 5939 [RFC5939].

A latent configuration represents a future capability, hence the pt= parameter is not directly meaningful in the lcfg attribute because no actual media session is being offered or accepted; it is permitted in order to tie any payload type number parameters within attributes to the proper media format. A primary example is the case of format parameters for the Redundant Audio Data (RED) payload, which are payload type numbers. Specific payload type numbers used in a latent configuration MAY be interpreted as suggestions to be used in any future offer based on the latent configuration, but they are not binding; the offerer and/or answerer may use any payload type numbers each deems appropriate. The use of explicit payload type numbers for latent configurations can be avoided by use of the parameter substitution rule of Section 3.3.7. Future extensions are also permitted. Note that leading zeroes are not permitted.

3.3.4.3. The Media Type Parameter

When a latent configuration is specified (always at the media level), indicating the ability to support an additional media stream, it is necessary to specify the media type (audio, video, etc.) as well as the format and transport type. The media type parameter is defined in ABNF as

```
media-type = ["+"] "mt=" media; media defined in RFC4566
```

At present, the media-type parameter is used only in the latent configuration attribute, and the use of the "+" prefix to specify that the entire attribute line is to be ignored if the mt= parameter is not understood, is unnecessary. However, if the media-type parameter is later added to an existing capability attribute such as pcfg, then the "+" would be useful. The media format(s) and transport type(s) are specified using the media configuration parameter ("m=") defined above, and the transport parameter ("t=") defined in RFC 5939 [RFC5939], respectively.

3.3.5. The Latent Configuration Attribute

One of the goals of this work is to permit the exchange of supportable media configurations in addition to those offered or accepted for immediate use. Such configurations are referred to as "latent configurations". For example, a party may offer to establish a session with an audio stream, and, at the same time, announce its ability to support a video stream as part of the same session. The offerer can supply its video capabilities by offering one or more latent video configurations along with the media stream for audio; the responding party may indicate its ability and willingness to support such a video session by returning a corresponding latent configuration.

Latent configurations returned in SDP answers MUST match offered latent configurations (or parameter subsets thereof). Therefore, it is appropriate for the offering party to announce most, if not all, of its capabilities in the initial offer. This choice has been made in order to keep the size of the answer more compact by not requiring acap, rmcap, tcap, etc. lines in the answer.

Latent configurations may be announced by use of the latent configuration attribute, which is defined in a manner very similar to the potential configuration attribute. The latent configuration attribute combines the properties of a media line and a potential configuration. A latent configuration MUST include a media type (mt=) and a transport protocol configuration parameter since the latent configuration is independent of any media line present. In

most cases, the media configuration (m=) parameter needs to be present as well (see Section 4 for examples). The lcfcg attribute is a media level attribute.

The lcfcg attribute is defined as a media level attribute since it specifies a possible future media stream. However the lcfcg attribute is not necessarily related to the media description within which it is provided. Session capabilities ("sescap") may be used to indicate this.

Each media line in an SDP description represents an offered simultaneous media stream, whereas each latent configuration represents an additional stream which may be negotiated in a future offer/answer exchange. Session capability attributes may be used to determine whether a latent configuration may be used to form an offer for an additional simultaneous stream or to reconfigure an existing stream in a subsequent offer/answer exchange.

The latent configuration attribute is of the form:

```
a=lcfcg:<config-number> <latent-cfg-list>
```

which adheres to the SDP [RFC4566] "attribute" production with att-field and att-value defined as:

```
att-field   = "lcfcg"
att-value   = config-number 1*WSP lcfcg-cfg-list
config-number = NonZeroDigit *9(DIGIT) ; DIGIT defined in RFC5234
lcfcg-cfg-list = media-type 1*WSP pot-cfg-list
                  ; as defined in RFC5939
                  ; and extended herein
```

The media-type (mt=) parameter identifies the media type (audio, video, etc.) to be associated with the latent media stream, and MUST be present. The pot-cfg-list MUST contain a transport-protocol-config-list (t=) parameter and a media-config-list (m=) parameter. The pot-cfg-list MUST NOT contain more than one instance of each type of parameter list. As specified in RFC 5939 [RFC5939], the use of the "+" prefix with a parameter indicates that the entire configuration MUST be ignored if the parameter is not understood; otherwise, the parameter itself may be ignored.

Media stream payload numbers are not assigned by a latent configuration. Assignment will take place if and when the corresponding stream is actually offered via an m-line in a later exchange. The payload-number-config-list is included as a parameter to the lcfcg attribute in case it is necessary to tie payload numbers in attribute capabilities to specific media capabilities.

If an `lcfg` attribute invokes an `acap` attribute that appears at the session level, then that attribute will be expected to appear at the session level of a subsequent offer when and if a corresponding media stream is offered. Otherwise, `acap` attributes which appear at the media level represent media-level attributes. Note, however, that `rmcap`, `omcap`, `mfcap`, `mscap`, and `tcap` attributes may appear at the session level because they always result in media-level attributes or m-line parameters.

The configuration numbers for latent configurations do not imply a preference; the offerer will imply a preference when actually offering potential configurations derived from latent configurations negotiated earlier. Note however that the offerer of latent configurations MAY specify preferences for combinations of potential and latent configurations by use of the `sescap` attribute defined in Section 3.3.8. For example, if an SDP offer contains, say, an audio stream with `pcfg:1`, and two latent video configurations, `lcfg:2`, and `lcfg:3`, then a session with one audio stream and one video stream could be specified by including "`a=sescap:1 1,2|3`". One audio stream and two video streams could be specified by including "`a=sescap:2 1,2,3`" in the offer. In order to permit combinations of latent and potential configurations in session capabilities, latent configuration numbers MUST be different from those used for potential configurations. This restriction is especially important if the offerer does not require `cmmed-v0` capability and the recipient of the offer doesn't support it. If the `lcfg` attribute is not recognized, the capability attributes intended to be associated with it may be confused with those associated with a potential configuration of some other media stream. Note also that leading zeroes are not permitted in configuration numbers.

If a cryptographic attribute, such as the SDES "`a=crypto:`" attribute [RFC4568], is referenced by a latent configuration through an `acap` attribute, any keying material required in the conventional attribute, such as the SDES key/salt string, MUST be included in order to satisfy formatting rules for the attribute. Since the keying material will be visible but not actually used at this stage (since it's a latent configuration), the value(s) of the keying material MUST NOT be a real value used for real exchange of media, and the receiver of the `lcfg` attribute MUST ignore the values.

3.3.6. Enhanced Potential Configuration Attribute

The present work requires new extensions (parameters) for the `pcfg` attribute defined in the SDP Capability Negotiation base protocol [RFC5939]. The parameters and their definitions are "borrowed" from the definitions provided for the latent configuration attribute in Section 3.3.5. The expanded ABNF definition of the `pcfg` attribute is

a=pcfg: <config-number> [<pot-cfg-list>]

where

```
config-number = 1*DIGIT ;defined in [RFC5234]
pot-cfg-list  = pot-config *(1*WSP pot-config)
pot-config    = attribute-config-list / ;def in [RFC5939]
               transport-protocol-config-list / ;defined in [RFC5939]
               extension-config-list / ;[RFC5939]
               media-config-list / ; Section 3.3.4.1
               payload-number-config-list ; Section 3.3.4.2
```

Except for the extension-config-list, the pot-cfg-list MUST NOT contain more than one instance of each parameter list.

3.3.6.1. Returning Capabilities in the Answer

Potential and/or latent configuration attributes may be returned within an answer SDP to indicate the ability of the answerer to support alternative configurations of the corresponding stream(s). For example, an offer may include multiple potential configurations for a media stream and/or latent configurations for additional streams; the corresponding answer will indicate (via an acfg attribute) the configuration accepted and used to construct the base configuration for each active media stream in the reply, but the reply MAY also contain potential and/or latent configuration attributes, with parameters, to indicate which other offered configurations would be acceptable. This information is useful if it becomes desirable to reconfigure a media stream, e.g., to reduce resource consumption.

When potential and/or latent configurations are returned in an answer, all numbering MUST refer to the configuration and capability attribute numbering of the offer. The offered capability attributes need not be returned in the answer. The answer MAY include additional capability attributes and/or configurations (with distinct numbering). The parameter values of any returned pcfg or lcfc attributes MUST be a subset of those included in the offered configurations and/or those added by the answerer; values MAY be omitted only if they were indicated as alternative sets, or optional, in the original offer. The parameter set indicated in the returned acfg attribute need not be repeated in a returned pcfg attribute. The answerer MAY return more than one pcfg attribute with the same configuration number if it is necessary to describe selected combinations of optional or alternative parameters.

Similarly, one or more session capability attributes (a=sescap) MAY be returned to indicate which of the offered session capabilities is/

are supportable by the answerer (see Section 3.3.8.)

Note that, although the answerer MAY return capabilities beyond those included by the offerer, these capabilities MUST NOT be used to form any base level media description in the answer. For this reason, it is advisable for the offerer to include most, if not all, potential and latent configurations it can support in the initial offer, unless the size of the resulting SDP is a concern. Either party MAY later announce additional capabilities by renegotiating the session in a second offer/answer exchange.

3.3.6.2. Payload Type Number Mapping

When media format capabilities defined in `rmcap` attributes are used in potential configuration lines, the transport protocol uses RTP and it is necessary to assign payload type numbers. In some cases, it is desirable to assign different payload type numbers to the same media format capability when used in different potential configurations. One example is when configurations for AVP and SAVP are offered: the offerer would like the answerer to use different payload type numbers for encrypted and unencrypted media, so the offerer can decide whether or not to render early media which arrives before the answer is received.

For example, if use of AVP was selected by the answerer, then media received by the offerer is not encrypted and hence can be played out prior to receiving the answer. Conversely, if SAVP was selected, cryptographic parameters and keying material present in the answer may be needed to decrypt received media. If the offer configuration indicated that AVP media uses one set of payload types and SAVP a different set, then the offerer will know whether media received prior to the answer is encrypted or not by simply looking at the RTP payload type number in the received packet.

This association of distinct payload type number(s) with different transport protocols requires a separate `pcfg` line for each protocol. Clearly, this technique cannot be used if the number of potential configurations exceeds the number of possible payload type numbers.

3.3.6.3. Processing of Media-Format-Related Conventional Attributes for Potential Configurations

When media capabilities negotiation is employed, SDP records are likely to contain conventional attributes such as `rtpmap`, `fntp`, and other media-format-related lines, as well as capability attributes such as `rmcap`, `omcap`, `mfcap`, and `mscap` which map into those conventional attributes when invoked by a potential configuration. In such cases, it MAY be appropriate to employ the `delete-attributes`

option [RFC5939] in the attribute configuration list parameter in order to avoid the generation of conflicting fmtp attributes for a particular configuration. Any media-specific attributes in the media block which refer to media formats not used by the potential configuration MUST be ignored.

For example:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 3456 RTP/AVP 0 18 100
a=rtpmap:100 telephone-event
a=fmtp:100 0-11
a=rmcap:1 PCMU/8000
a=rmcap:2 G729/8000
a=rmcap:3 telephone-event/8000
a=mfcap:3 0-15
a=pcfg:1 m=2,3|1,3 a=-m pt=1:0,2:18,3:100
a=pcfg:2
```

In this example, PCMU is media capability 1, G729 is media capability 2, and telephone-event is media capability 3. The a=pcfg:1 line specifies that the preferred configuration is G.729 with extended dtmf events, second is G.711 mu-law with extended dtmf events, and the base media-level attributes are to be deleted. Intermixing of G.729, G.711, and "commercial" dtmf events is least preferred (the base configuration provided by the "m=" line, which is, by default, the least preferred configuration). The rtpmap and fmtp attributes of the base configuration are replaced by the rmcap and mfcap attributes when invoked by the proposed configuration.

If the preferred configuration is selected, the SDP answer will look like

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=csup:med-v0
m=audio 3456 RTP/AVP 18 100
a=rtpmap:100 telephone-event/8000
a=fmtp:100 0-15
a=acfg:1 m=2,3 pt=1:0,2:18,3:100
```

3.3.7. Substitution of Media Payload Type Numbers in Capability Attribute Parameters

In some cases, for example, when an RFC 2198 [RFC2198] redundancy audio subtype (RED) capability is defined in an mfcap attribute, the parameters to an attribute may contain payload type numbers. Two options are available for specifying such payload type numbers. They may be expressed explicitly, in which case they are bound to actual payload types by means of the payload type number parameter (pt=) in the appropriate potential or latent configuration. For example, the following SDP fragment defines a potential configuration with redundant G.711 mu-law:

```
m=audio 45678 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=rmcap:1 PCMU/8000
a=rmcap:2 RED/8000
a=mfcap:2 0/0
a=pcfg:1 m=2,1 pt=2:98,1:0
```

The potential configuration is then equivalent to

```
m=audio 45678 RTP/AVP 98 0
a=rtpmap:0 PCMU/8000
a=rtpmap:98 RED/8000
a=fmtp:98 0/0
```

A more general mechanism is provided via the parameter substitution rule. When an mfcap, mscap, or acap attribute is processed, its arguments will be scanned for a payload type number escape sequences of the following form (in ABNF):

```
ptn-esc = "%m=" media-cap-num "%" ; defined in 3.3.1
```

If the sequence is found, the sequence is replaced by the payload type number assigned to the media capability number, as specified by the pt= parameter in the selected potential configuration; only actual payload type numbers are supported - wildcards are excluded. The sequence "%" (null digit string) is replaced by a single percent sign and processing continues with the next character, if any.

For example, the above offer sequence could have been written as

```
m=audio 45678 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=rmcap:1 PCMU/8000
a=rmcap:2 RED/8000
a=mfcap:2 %m=1%/m=1%
```

a=pcfg:1 m=2,1 pt=2:98,1:0

and the equivalent SDP is the same as above.

3.3.8. The Session Capability Attribute

Potential and latent configurations enable offerers and answerers to express a wide range of alternative configurations for current and future negotiation. However in practice, it may not be possible to support all combinations of these configurations.

The session capability attribute provides a means for the offerer and/or the answerer to specify combinations of specific media stream configurations which it is willing and able to support. Each session capability in an offer or answer MAY be expressed as a list of required potential configurations, and MAY include a list of optional potential and/or latent configurations.

The choices of session capabilities may be based on processing load, total bandwidth, or any other criteria of importance to the communicating parties. If the answerer supports media capabilities negotiation, and session configurations are offered, it MUST accept one of the offered configurations, or it MUST refuse the session. Therefore, if the offer includes any session capabilities, it SHOULD include all the session capabilities the offerer is willing to support.

The session capability attribute is a session-level attribute described by:

"a=sescap:" <session num> <list of configs>

which corresponds to the standard value attribute definition with

```
att-field      = "sescap"
att-value      = session-num 1*WSP list-of-configs
                  [1*WSP optional-configs]
session-num    = NonZeroDigit *9(DIGIT) ; DIGIT defined
                  ; in RFC5234
list-of-configs = alt-config *("," alt-config)
optional-configs = "[" list-of-configs "]"
alt-config     = config-number *("|" config-number)
```

The session-num identifies the session: a lower-number session is preferred over a higher-number session, and leading zeroes are not permitted. Each alt-config list specifies alternative media configurations within the session; preference is based on config-num as specified in RFC 5939 [RFC5939]. Note that the session preference

order, when present, takes precedence over the individual media stream configuration preference order.

Use of session capability attributes requires that configuration numbers assigned to potential and latent configurations MUST be unique across the entire session; RFC 5939 [RFC5939] requires only that pcfg configuration numbers be unique within a media description. Also, leading zeroes are not permitted.

As an example, consider an endpoint that is capable of supporting an audio stream with either one H.264 video stream or two H.263 video streams with a floor control stream. In the latter case, the second video stream is optional. The SDP offer might look like the following (offering audio, an H.263 video streams, BFCP and another optional H.263 video stream)- the empty lines are added for readability only (not part of valid SDP):

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=sescap:2 1,2,5,[3]
a=sescap:1 1,4

m=audio 54322 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=pcfg:1

m=video 22344 RTP/AVP 102
a=rtpmap:102 H263-1998/90000
a=fmtp:102 CIF=4;QCIF=2;F=1;K=1
i=main video stream
a=label:11
a=pcfg:2
a=rmcap:1 H264/90000
a=mfcap:1 profile-level-id=42A01E; packetization-mode=2
a=acap:1 label:13
a=pcfg:4 m=1 a=1 pt=1:104

m=video 33444 RTP/AVP 103
a=rtpmap:103 H263-1998/90000
a=fmtp:103 CIF=4;QCIF=2;F=1;K=1
i=secondary video (slides)
a=label:12
a=pcfg:3
```

```
m=application 33002 TCP/BFCP *
a=setup:passive
a=connection:new
a=floorid:1 m-stream:11 12
a=floor-control:s-only
a=confid:4321
a=userid:1234
a=pcfg:5
```

If the answerer understands MediaCapNeg, but cannot support the Binary Floor Control Protocol, then it would respond with (invalid empty lines in SDP included again for readability):

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.22
t=0 0
a=csup:med-v0
a=sescap:1 1,4

m=audio 23456 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=acfg:1

m=video 41234 RTP/AVP 104
a=rtpmap:104 H264/90000
a=fmtp:104 profile-level-id=42A01E; packetization-mode=2
a=acfg:4 m=1 a=1 pt=1:104

m=video 0 RTP/AVP 103
a=acfg:3

m=application 0 TCP/BFCP *
a=acfg:5
```

An endpoint that doesn't support Media capabilities negotiation, but does support H.263 video, would respond with one or two H.263 video streams. In the latter case, the answerer may issue a second offer to reconfigure the session to one audio and one video channel using H.264 or H.263.

Session capabilities can include latent capabilities as well. Here's a similar example in which the offerer wishes to initially establish an audio stream, and prefers to later establish two video streams with chair control. If the answerer doesn't understand Media CapNeg, or cannot support the dual video streams or flow control, then it may support a single H.264 video stream. Note that establishment of the

most favored configuration will require two offer/answer exchanges.

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=sescap:1 1,3,4,5
a=sescap:2 1,2
a=sescap:3 1
a=rmcap:1 H263-1998/90000
a=mfcap:1 CIF=4;QCIF=2;F=1;K=1
a=tcap:1 RTP/AVP TCP/BFCP
m=audio 54322 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=pcfg:1
m=video 22344 RTP/AVP 102
a=rtpmap:102 H264/90000
a=fmtp:102 profile-level-id=42A01E; packetization-mode=2
a=label:11
a=content:main
a=pcfg:2
a=lcfg:3 mt=video t=1 m=1 a=31,32
a=acap:31 label:12
a=acap:32 content:main
a=lcfg:4 mt=video t=1 m=1 a=41,42
a=acap:41 label:13
a=acap:42 content:slides
a=lcfg:5 mt=application m=51 t=51
a=tcap:51 TCP/BFCP
a=omcap:51 *
a=acap:51 setup:passive
a=acap:52 connection:new
a=acap:53 floorid:1 m-stream:12 13
a=acap:54 floor-control:s-only
a=acap:55 confid:4321
a=acap:56 userid:1234
```

In this example, the default offer, as seen by endpoints which do not understand capabilities negotiation, proposes a PCMU audio stream and an H.264 video stream. Note that the offered lcfg lines for the video streams don't carry pt= parameters because they're not needed (payload type numbers will be assigned in the offer/answer exchange that establishes the streams). Note also that the three rmcap, mfcap, and tcap attributes used by lcfg:3 and lcfg:4 are included at the session level so they may be referenced by both latent configurations. As per Section 3.3, the media attributes generated

from the `rmcap`, `mfcap`, and `tcap` attributes are always media-level attributes. If the answerer supports Media CapNeg, and supports the most desired configuration, it would return the following SDP:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.22
t=0 0
a=csup:med-v0
a=sescap:1 1,3,4,5
a=sescap:2 1,2
a=sescap:3 1
m=audio 23456 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=acfg:1
m=video 0 RTP/AVP 102
a=pcfg:2
a=lcfg:3 mt=video t=1 m=1 a=31,32
a=lcfg:4 mt=video t=1 m=1 a=41,42
a=lcfg:5 mt=application t=2
```

This exchange supports immediate establishment of an audio stream for preliminary conversation. This exchange would presumably be followed at the appropriate time with a "reconfiguration" offer/answer exchange to add the video and chair control streams.

3.4. Offer/Answer Model Extensions

In this section, we define extensions to the offer/answer model defined in RFC 3264 [RFC3264] and RFC 5939 [RFC5939] to allow for media format and associated parameter capabilities, latent configurations and acceptable combinations of media stream configurations to be used with the SDP Capability Negotiation framework. Note that the procedures defined in this section extend the offer/answer procedures defined in RFC 5939 [RFC5939] Section 6; those procedures form a baseline set of capability negotiation offer/answer procedures that MUST be followed, subject to the extensions defined here.

SDP Capability Negotiation [RFC5939] provides a relatively compact means to offer the equivalent of an ordered list of alternative configurations for offered media streams (as would be described by separate `m=` lines and associated attributes). The attributes `acap`, `mscap`, `mfcap`, `omcap` and `rmcap` are designed to map somewhat straightforwardly into equivalent `m=` lines and conventional attributes when invoked by a `pcfg`, `lcfg`, or `acfg` attribute with appropriate parameters. The `a=pcfg:` lines, along with the `m=` line

itself, represent offered media configurations. The a=lcfg: lines represent alternative capabilities for future use.

3.4.1. Generating the Initial Offer

The Media Capabilities negotiation extensions defined in this document cover the following categories of features:

- o Media Format Capabilities and associated parameters (rmcap, omcap, mfcap, and mscap attributes)
- o Potential configurations using those media format capabilities and associated parameters
- o Latent media streams (lcfg attribute)
- o Acceptable combinations of media stream configurations (sescap attribute).

The high-level description of the operation is as follows:

When an endpoint generates an initial offer and wants to use the functionality described in the current document, it SHOULD identify and define the media formats and associated parameters it can support via the rmcap, omcap, mfcap and mscap attributes. The SDP media line(s) ("m=") should be made up with the actual configuration to be used if the other party does not understand capability negotiations (by default, this is the least preferred configuration). Typically, the media line configuration will contain the minimum acceptable configuration from the offerer's point of view.

Preferred configurations for each media stream are identified following the media line. The present offer may also include latent configuration (lcfg) attributes, at the media level, describing media streams and/or configurations the offerer is not now offering, but which it is willing to support in a future offer/answer exchange. A simple example might be the inclusion of a latent video configuration in an offer for an audio stream.

Lastly, if the offerer wishes to impose restrictions on the combinations of potential configurations to be used, it will include session capability (sescap) attributes indicating those.

If the offerer requires the answerer to understand the media capability extensions, the offerer MUST include a creq attribute containing the value "med-v0". If media capability negotiation is required only for specific media descriptions, the "med-v0" value MUST be provided only in creq attributes within those media

descriptions, as described in RFC 5939 [RFC5939].

Below, we provide a more detailed description of how to construct the offer SDP.

3.4.1.1. Offer with Media Capabilities

For each RTP-based media format the offerer wants to include as a media format capability, the offer MUST include an "rmcap" attribute for the media format as defined in Section 3.3.1.

For each non RTP-based media format the offer wants to include as a media format capability, the offer MUST include an "omcap" attribute for the media format as defined in Section 3.3.1.

Since the media capability number space is shared between the rmcap and omcap attributes, each media capability number provided (including ranges) MUST be unique in the entire SDP.

If an "fmt" parameter value is needed for a media format (whether RTP-based or not) in a media capability, then the offer MUST include one or more "mfcap" parameters with the relevant fmt parameter values for that media format as defined in Section 3.3.2. When multiple "mfcap" parameters are provided for a given media capability, they MUST be provided in accordance with the concatenation rules in Section 3.3.2.1.

For each of the media format capabilities above, the offer MAY include one or more "mscap" parameters with attributes needed for those specific media formats as defined in Section 3.3.3. Such attributes will be instantiated at the media-level, and hence session-level only attributes MUST NOT be used in the "mscap" parameter. The "mscap" parameter MUST NOT include an "rtpmap" or "fmt" attribute (rmcap and mfcap are used instead).

If the offerer wants to limit the relevance (and use) of a media format capability or parameter to a particular media stream, the media format capability or parameter MUST be provided within the corresponding media description. Otherwise, the media format capabilities and parameters MUST be provided at the session level. Note however, that the attribute or parameter embedded in these will always be instantiated at the media-level.

This is due to those parameters being effectively media-level parameters. If session-level attributes are needed, the "acap" attribute defined in RFC 5939 [RFC5939] can be used, however it does not provide for media format-specific instantiation.

Inclusion of the above does not constitute an offer to use the capabilities; a potential configuration is needed for that. If the offerer wants to offer one or more of the media capabilities above, they MUST be included as part of a potential configuration (pcfg) attribute as defined in Section 3.3.4. Each potential configuration MUST include a config-number, and each config-number MUST be unique in the entire SDP (note that this differs from RFC 5939 [RFC5939], which only requires uniqueness within a media description). Also, the config-number MUST NOT overlap with any config-number used by a latent configuration in the SDP. As described in RFC 5939 [RFC5939], lower config-numbers indicate a higher preference; the ordering still applies within a given media description only though.

For a media capability to be included in a potential configuration, there MUST be an "m=" parameter in the pcfg attribute referencing the media capability number in question. When one or more media capabilities are included in an offered potential configuration (pcfg), they completely replace the list of media formats offered in the actual configuration (m= line). Any attributes included for those formats remain in the SDP though (e.g., rtpmap, fmp, etc.). For non-RTP based media formats, the format-name (from the "omcap" media capability) is simply added to the "m=" line as a media format (e.g. t38). For RTP-based media, payload type mappings MUST be provided by use of the "pt" parameter in the potential configuration (see Section 3.3.4.2); payload type escaping may be used in mfcap, mscap, and acap attributes as defined in Section 3.3.7.

Note that the "mt" parameter MUST NOT be used with the pcfg attribute (since it is defined for the lcfcg attribute only); the media type in a potential configuration cannot be changed from that of the encompassing media description.

3.4.1.2. Offer with Latent Configuration

If the offerer wishes to offer one or more latent configurations for future use, the offer MUST include a latent configuration attribute (lcfcg) for each as defined in Section 3.3.5.

Each lcfcg attribute

- o MUST be specified at the media level
- o MUST include a config-number that is unique in the entire SDP (incl. for any potential configuration attributes). Note that config-numbers in latent configurations do not indicate any preference order

- o MUST include a media type ("mt")
- o MUST reference a valid transport capability ("t")

Each `lcfg` attribute MAY include additional capability references, which may refer to capabilities anywhere in the session description, subject to any restrictions normally associated with such capabilities. For example, a media-level attribute capability must be present at the media-level in some media description in the SDP. Note that this differs from the potential configuration attribute, which cannot validly refer to media-level capabilities in another media description (per RFC 5939 [RFC5939], Section 3.5.1).

Potential configurations constitute an actual offer and hence may instantiate a referenced capability. Latent configurations are not actual offers and hence cannot instantiate a referenced capability; it is therefore safe for those to refer to capabilities in another media description.

3.4.1.3. Offer with Configuration Combination Restrictions

If the offerer wants to indicate restrictions or preferences among combinations of potential and/or latent configuration, a session capability (`sescap`) attribute MUST be provided at the session-level for each such combination as described in Section 3.3.8. Each `sescap` attribute MUST include a session-num that is unique in the entire SDP; the lower the session-num the more preferred that combination is. Furthermore, `sescap` preference order takes precedence over any order specified in individual `pcfg` attributes.

For example, if we have `pcfg-1` and `pcfg-2`, and `sescap-1` references `pcfg-2`, whereas `sescap-2` references `pcfg-1`, then `pcfg-2` will be the most preferred potential configuration. Without the `sescap`, `pcfg-1` would be the most preferred.

3.4.2. Generating the Answer

When receiving an offer, the answerer MUST check the offer for `creg` attributes containing the value "med-v0"; answerers compliant with this specification will support this value in accordance with the procedures specified in RFC 5939 [RFC5939].

The SDP MAY contain

- o Media format capabilities and associated parameters (`rmcap`, `omcap`, `mfcap`, and `mscap` attributes)

- o Potential configurations using those media format capabilities and associated parameters
- o Latent media streams (lcfg attribute)
- o Acceptable combinations of media stream configurations (sescap attribute)

The high-level informative description of the operation is as follows:

When the answering party receives the offer and if it supports the required capability negotiation extensions, it should select the most-preferred configuration it can support for each media stream, and build its answer accordingly. The configuration selected for each accepted media stream is placed into the answer as a media line with associated parameters and attributes. If a proposed configuration is chosen for a given media stream, the answer must contain an actual configuration (acfg) attribute for that media stream to indicate which offered pcfg attribute was used to build the answer. The answer should also include any potential or latent configurations the answerer can support, especially any configurations compatible with other potential or latent configurations received in the offer. The answerer should make note of those configurations it might wish to offer in the future.

Below we provide a more detailed normative description of how the answerer processes the offer SDP and generates an answer SDP.

3.4.2.1. Processing Media Capabilities and Potential Configurations

The answerer **MUST** first determine if it needs to perform media capability negotiation by examining the SDP for valid and preferred potential configuration attributes that include media configuration parameters (i.e., an "m" parameter in the pcfg attribute).

Such a potential configuration is valid if:

1. It is valid according to the rules defined in RFC 5939 [RFC5939]
2. It contains a config-number that is unique in the entire SDP and does not overlap with any latent configuration config-numbers
3. All media format capabilities (rmcap or omcap), media format parameter capabilities (mfcap), and media-specific capabilities (mscap) referenced by the potential configuration ("m" parameter) are valid themselves (as defined in Section 3.3.1, 3.3.2, and 3.3.3) and each of them is provided either at the session level

or within this particular media description.

4. All RTP-based media format capabilities (rmcap) have a corresponding payload type ("pt") parameter in the potential configuration that result in mapping to a valid payload type that is unique within the resulting SDP.
5. Any concatenation (see Section 3.3.2.1) and substitution (see Section 3.3.7) applied to any capability (mfcap, mscap, or acap) referenced by this potential configuration results in a valid SDP.

Note that, since SDP does not interpret the value of fmtpparameters, any resulting fmtpparameter value will be considered valid.

Secondly, the answerer MUST determine the order in which potential configurations are to be negotiated. In the absence of any Session Capability ("sescap") attributes, this simply follows the rules of RFC 5939 [RFC5939], with a lower config-number within a media description being preferred over a higher one. If a valid "sescap" attribute is present, the preference order provided in the "sescap" attribute MUST take precedence. A "sescap" attribute is considered valid if:

1. It adheres to the rules provided in Section 3.3.8.
2. All the configurations referenced by the "sescap" attribute are valid themselves (note that this can include the actual, potential and latent configurations).

The answerer MUST now process the offer for each media stream based on the most preferred valid potential configuration in accordance with the procedures specified in RFC 5939 [RFC5939], Section 3.6.2, and further extended below:

- o If one or more media format capabilities are included in the potential configuration, then they replace all media formats provided in the "m=" line for that media description. For non-RTP-based media formats (omcap), the format-name is added. For RTP-based media formats (rmcap), the payload-type specified in the payload-type mapping ("pt") is added and a corresponding "rtpmap" attribute is added to the media description.
- o If one or more media format parameter capabilities are included in the potential configuration, then the corresponding "fmtpparameters" attributes are added to the media description. Note that this inclusion is done indirectly via the media format capability.

- o If one or more media-specific capabilities are included in the potential configuration, then the corresponding attributes are added to the media description. Note that this inclusion is done indirectly via the media format capability.
- o When checking to see if the answerer supports a given potential configuration that includes one or more media format capabilities, the answerer MUST support at least one of the media formats offered. If he does not, the answerer MUST proceed to the next potential configuration based on the preference order that applies.
- o If Session Capability ("sescap") preference ordering is included, then the potential configuration selection process MUST adhere to the ordering provided. Note that this may involve coordinated selection of potential configurations between media descriptions. The answerer MUST accept one of the offered "sescap" combinations (i.e. all the required potential configurations specified) or it MUST reject the entire session.

Once the answerer has selected a valid and supported offered potential configuration for all of the media streams (or has fallen back to the actual configuration plus any added session attributes), the answerer MUST generate a valid answer SDP as described in RFC 5939 [RFC5939], Section 3.6.2, and further extended below:

- o Additional answer capabilities and potential configurations MAY be returned in accordance with Section 3.3.6.1. Capability numbers and configuration numbers for those MUST be distinct from the ones used in the offer SDP.
- o Latent configuration processing and answer generation MUST be performed, as specified below.
- o Session capability specification for the potential and latent configurations in the answer MAY be included (see Section 3.3.8).

3.4.2.2. Latent Configuration Processing

The answerer MUST determine if it needs to perform any latent configuration processing by examining the SDP for valid latent configuration attributes (lcfg). An lcfg attribute is considered valid if:

- o It adheres to the description in Section 3.3.5.
- o It includes a config-number that is unique in the entire SDP and does not overlap with any potential configuration config-number

- o It includes a valid media type ("mt=")
- o It references a valid transport capability ("t=")
- o All other capabilities referenced by it are valid.

For each such valid latent configuration in the offer, the answerer checks to see if it could support the latent configuration in a subsequent offer/answer exchange. If so, it includes the latent configuration with the same configuration number in the answer, similar to the way potential configurations are processed and the selected one returned in an actual configuration attribute (see RFC 5939 [RFC5939]). If the answerer supports only a (non-mandatory) subset of the parameters offered in a latent configuration, the answer latent configuration will include only those parameters supported (similar to "acfg" processing). Note that latent configurations do not constitute an actual offer at this point in time; they merely indicate additional configurations that could be supported.

If a Session Capability ("sescap") attribute is included and it references a latent configuration, then the answerer processing of that latent configuration must be done within the constraints specified by that Session Capability, i.e. it must be possible to support it at the same time as any required (i.e. non-optional) potential configurations in the session capability. The answerer may in turn add his own "sescap" indications in the answer as well.

3.4.3. Offerer Processing of the Answer

The offerer MUST process the answer in accordance with RFC 5939 [RFC5939] Section 3.6.3, and further explained below.

When the offerer processes the answer SDP based on a valid actual configuration attribute in the answer, and that valid configuration includes one or more media capabilities, the processing MUST furthermore be done as if the offer was sent using those media capabilities instead of the actual configuration. In particular, the media formats in the "m=" line, and any associated payload type mappings (rtpmap), fmp parameters (mfcap) and media-specific attributes (mscap) MUST be used. Note that this may involve use of concatenation and substitution rules (see Section 3.3.2.1 and 3.3.7). The actual configuration attribute may also be used to infer the lack of acceptability of higher-preference configurations that were not chosen, subject to any constraints provided by a Session Capability attribute ("sescap") in the offer. Note that the SDP Capability Negotiation base specification [RFC5939] requires the answerer to choose the highest preference configuration it can support, subject

to local policies.

When the offerer receives the answer, it SHOULD furthermore make note of any capabilities and/or latent configurations included for future use, and any constraints on how those may be combined.

3.4.4. Modifying the Session

If, at a later time, one of the parties wishes to modify the operating parameters of a session, e.g., by adding a new media stream, or by changing the properties used on an existing stream, it can do so via the mechanisms defined for offer/answer [RFC3264]. If the initiating party has remembered the codecs, potential configurations, latent configurations and session capabilities provided by the other party in the earlier negotiation, it MAY use this knowledge to maximize the likelihood of a successful modification of the session. Alternatively, the initiator MAY perform a new capabilities exchange as part of the reconfiguration. In such a case, the new capabilities will replace the previously-negotiated capabilities. This may be useful if conditions change on the endpoint.

4. Examples

In this section, we provide examples showing how to use the Media Capabilities with the SDP Capability Negotiation.

4.1. Alternative Codecs

This example provides a choice of one of six variations of the adaptive multirate codec. In this example, the default configuration as specified by the media line is the same as the most preferred configuration. Each configuration uses a different payload type number so the offerer can interpret early media.

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 54322 RTP/AVP 96
a=rtpmap:96 AMR-WB/16000/1
a=fmtp:96 mode-change-capability=1; max-red=220; \
mode-set=0,2,4,7
a=rmcap:1,3,5 audio AMR-WB/16000/1
a=rmcap:2,4,6 audio AMR/8000/1
a=mfcap:1,2,3,4 mode-change-capability=1
a=mfcap:5,6 mode-change-capability=2
a=mfcap:1,2,3,5 max-red=220
a=mfcap:3,4,5,6 octet-align=1
a=mfcap:1,3,5 mode-set=0,2,4,7
a=mfcap:2,4,6 mode-set=0,3,5,6
a=pcfg:1 m=1 pt=1:96
a=pcfg:2 m=2 pt=2:97
a=pcfg:3 m=3 pt=3:98
a=pcfg:4 m=4 pt=4:99
a=pcfg:5 m=5 pt=5:100
a=pcfg:6 m=6 pt=6:101
```

In the above example, media capability 1 could have been excluded from the first rmcap declaration and from the corresponding mfcap attributes, and the pcfg:1 attribute line could have been simply "pcfg:1".

The next example offers a video stream with three options of H.264 and 4 transports. It also includes an audio stream with different audio qualities: four variations of AMR, or AC3. The offer looks something like:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=An SDP Media NEG example
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=ice-pwd:speEc3QGZiNWpVLFJhQX
m=video 49170 RTP/AVP 100
c=IN IP4 192.0.2.56
a=maxprate:1000
a=rtcp:51540
a=sendonly
a=candidate 12345 1 UDP 9 192.0.2.56 49170 host
a=candidate 23456 2 UDP 9 192.0.2.56 51540 host
a=candidate 34567 1 UDP 7 198.51.100.1 41345 srflx raddr \
192.0.2.56 rport 49170
a=candidate 45678 2 UDP 7 198.51.100.1 52567 srflx raddr \
192.0.2.56 rport 51540
a=candidate 56789 1 UDP 3 192.0.2.100 49000 relay raddr \
192.0.2.56 rport 49170
a=candidate 67890 2 UDP 3 192.0.2.100 49001 relay raddr \
192.0.2.56 rport 51540
b=AS:10000
b=TIAS:10000000
b=RR:4000
b=RS:3000
a=rtpmap:100 H264/90000
a=fmtp:100 profile-level-id=42A01E; packetization-mode=2; \
sprop-parameter-sets=Z0IACpZTBmI,aMljiA==; \
sprop-interleaving-depth=45; sprop-deint-buf-req=64000; \
sprop-init-buf-time=102478; deint-buf-cap=128000
a=tcap:1 RTP/SAVPF RTP/SAVP RTP/AVPF
a=rmcap:1-3,7-9 H264/90000
a=rmcap:4-6 rtx/90000
a=mfcap:1-9 profile-level-id=42A01E
a=mfcap:1-9 aMljiA==
a=mfcap:1,4,7 packetization-mode=0
a=mfcap:2,5,8 packetization-mode=1
a=mfcap:3,6,9 packetization-mode=2
a=mfcap:1-9 sprop-parameter-sets=Z0IACpZTBmI
a=mfcap:1,7 sprop-interleaving-depth=45; \
sprop-deint-buf-req=64000; sprop-init-buf-time=102478; \
deint-buf-cap=128000
a=mfcap:4 apt=100
a=mfcap:5 apt=99
a=mfcap:6 apt=98
a=mfcap:4-6 rtx-time=3000
a=mscap:1-6 rtcp-fb nack
```

```
a=acap:1 crypto:1 AES_CM_128_HMAC_SHA1_80 \
inline:d0RmdmcmVCspeEc3QGZiNWpVLFJhQXlcfHAWJSoj|220|1:32
a=pcfg:1 t=1 m=1,4 a=1 pt=1:100,4:97
a=pcfg:2 t=1 m=2,5 a=1 pt=2:99,4:96
a=pcfg:3 t=1 m=3,6 a=1 pt=3:98,6:95
a=pcfg:4 t=2 m=7 a=1 pt=7:100
a=pcfg:5 t=2 m=8 a=1 pt=8:99
a=pcfg:6 t=2 m=9 a=1 pt=9:98
a=pcfg:7 t=3 m=1,3 pt=1:100,4:97
a=pcfg:8 t=3 m=2,4 pt=2:99,4:96
a=pcfg:9 t=3 m=3,6 pt=3:98,6:95
m=audio 49176 RTP/AVP 101 100 99 98
c=IN IP4 192.0.2.56
a=ptime:60
a=maxptime:200
a=rtcp:51534
a=sendonly
a=candidate 12345 1 UDP 9 192.0.2.56 49176 host
a=candidate 23456 2 UDP 9 192.0.2.56 51534 host
a=candidate 34567 1 UDP 7 198.51.100.1 41348 srflx \
raddr 192.0.2.56 rport 49176
a=candidate 45678 2 UDP 7 198.51.100.1 52569 srflx \
raddr 192.0.2.56 rport 51534
a=candidate 56789 1 UDP 3 192.0.2.100 49002 relay \
raddr 192.0.2.56 rport 49176
a=candidate 67890 2 UDP 3 192.0.2.100 49003 relay \
raddr 192.0.2.56 rport 51534
b=AS:512
b=TIAS:512000
b=RR:4000
b=RS:3000
a=maxprate:120
a=rtpmap:98 AMR-WB/16000
a=fmtp:98 octet-align=1; mode-change-capability=2
a=rtpmap:99 AMR-WB/16000
a=fmtp:99 octet-align=1; crc=1; mode-change-capability=2
a=rtpmap:100 AMR-WB/16000/2
a=fmtp:100 octet-align=1; interleaving=30
a=rtpmap:101 AMR-WB+/72000/2
a=fmtp:101 interleaving=50; int-delay=160000;
a=rmcap:14 ac3/48000/6
a=acap:23 crypto:1 AES_CM_128_HMAC_SHA1_80 \
inline:d0RmdmcmVCspeEc3QGZiNWpVLFJhQXlcfHAWJSoj|220|1:32
a=tcap:4 RTP/SAVP
a=pcfg:10 t=4 a=23
a=pcfg:11 t=4 m=14 a=23 pt=14:102
```

This offer illustrates the advantage in compactness that arises if

one can avoid deleting the base configuration attributes and recreating them in acap attributes for the potential configurations.

4.2. Alternative Combinations of Codecs (Session Configurations)

If an endpoint has limited signal processing capacity, it might be capable of supporting, say, a G.711 mu-law audio stream in combination with an H.264 video stream, or a G.729B audio stream in combination with an H.263-1998 video stream. It might then issue an offer like the following:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
a=sescap:1 2,4
a=sescap:2 1,3
m=audio 54322 RTP/AVP 18
a=rtpmap:18 G729/8000
a=fmtp:18 annexb=yes
a=rmcap:1 PCMU/8000
a=pcfg:1 m=1 pt=1:0
a=pcfg:2
m=video 54344 RTP/AVP 100
a=rtpmap:100 H263-1998/90000
a=rmcap:2 H264/90000
a=mfcap:2 profile-level-id=42A01E; packetization-mode=2
a=pcfg:3 m=2 pt=2:101
a=pcfg:4
```

Note that the preferred session configuration (and the default as well) is G.729B with H.263. This overrides the individual media stream preferences which are PCMU and H.264 by the potential configuration numbering rule.

4.3. Latent Media Streams

Consider a case in which the offerer can support either G.711 mu-law, or G.729B, along with DTMF telephony events for the 12 common touchtone signals, but is willing to support simple G.711 mu-law audio as a last resort. In addition, the offerer wishes to announce its ability to support video and MSRP in the future, but does not wish to offer a video stream or an MSRP stream at present. The offer might look like the following:

```
v=0
o=- 25678 753849 IN IP4 192.0.2.1
s=
c=IN IP4 192.0.2.1
t=0 0
a=creq:med-v0
m=audio 23456 RTP/AVP 0
a=rtpmap:0 PCMU/8000
a=rmcap:1 PCMU/8000
a=rmcap:2 G729/8000
a=rmcap:3 telephone-event/8000
a=mfcap:3 0-11
a=pcfg:1 m=1,3|2,3 pt=1:0,2:18,3:100
a=lcfg:2 mt=video t=1 m=10|11
a=rmcap:10 H263-1998/90000
a=rmcap:11 H264/90000
a=tcap:1 RTP/AVP
a=lcfg:3 mt=message t=2 m=20
a=tcap:2 TCP/MSRP
a=omcap:20 *
```

The first lcfg attribute line ("lcfg:2") announces support for H.263 and H.264 video (H.263 preferred) for future negotiation. The second lcfg attribute line ("lcfg:3") announces support for MSRP for future negotiation. The m-line and the rtpmap attribute offer an audio stream and provide the lowest precedence configuration (PCMU without any DTMF encoding). The rmcap lines define the RTP-based media format capabilities (PCMU, G729, telephone-event, H263-1998 and H264) and the omcap line defines the non-RTP based media format capability (wildcard). The mfcap attribute provides the format parameters for telephone-event, specifying the 12 commercial DTMF 'digits'. The pcfg attribute line defines the most-preferred media configuration as PCMU plus DTMF events and the next-most-preferred configuration as G.729B plus DTMF events.

If the answerer is able to support all the potential configurations, and also support H.263 video (but not H.264), it would reply with an answer like:

```
v=0
o=- 24351 621814 IN IP4 192.0.2.2
s=
c=IN IP4 192.0.2.2
t=0 0
a=csup:med-v0
m=audio 54322 RTP/AVP 0 100
a=rtpmap:0 PCMU/8000
a=rtpmap:100 telephone-event/8000
```

```
a=fmtp:100 0-11
a=acfg:1 m=1,3 pt=1:0,3:100
a=pcfg:1 m=2,3 pt=2:18,3:100
a=lcfg:2 mt=video t=1 m=10
```

The lcfg attribute line announces the capability to support H.263 video at a later time. The media line and subsequent rtpmap and fmtp attribute lines present the selected configuration for the media stream. The acfg attribute line identifies the potential configuration from which it was taken, and the pcfg attribute line announces the potential capability to support G.729 with DTMF events as well. If, at some later time, congestion becomes a problem in the network, either party may, with expectation of success, offer a reconfiguration of the media stream to use G.729 in order to reduce packet sizes.

5. IANA Considerations

5.1. New SDP Attributes

The IANA is hereby requested to register the following new SDP attributes:

Attribute name: rmcap
Long form name: RTP-based media format capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate RTP-based media capability number(s) with media subtype and encoding parameters
Appropriate Values: see Section 3.3.1
Contact name: Flemming Andreassen, fandres@cisco.com

Attribute name: omcap
Long form name: Non RTP-based media format capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate non RTP-based media capability number(s) with media subtype and encoding parameters
Appropriate Values: see Section 3.3.1
Contact name: Flemming Andreassen, fandreas@cisco.com

Attribute name: mfcap
Long form name: media format parameter capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate media format attributes and parameters with media format capabilities
Appropriate Values: see Section 3.3.2
Contact name: Flemming Andreassen, fandreas@cisco.com

Attribute name: mscap
Long form name: media-specific capability
Type of attribute: session-level and media-level
Subject to charset: no
Purpose: associate media-specific attributes and parameters with media capabilities
Appropriate Values: see Section 3.3.3
Contact name: Flemming Andreassen, fandreas@cisco.com

Attribute name: lcfg
Long form name: latent configuration
Type of attribute: media-level
Subject to charset: no
Purpose: to announce supportable media streams

without offering them for immediate use.

Appropriate Values: see Section 3.3.5

Contact name: Flemming Andreassen, fandreas@cisco.com

Attribute name: sescap

Long form name: session capability

Type of attribute: session-level

Subject to charset: no

Purpose: to specify and prioritize acceptable combinations of media stream configurations.

Appropriate Values: see Section 3.3.8

Contact name: Flemming Andreassen, fandreas@cisco.com

5.2. New SDP Capability Negotiation Option Tag

The IANA is hereby requested to add the new option tag "med-v0", defined in this document, to the SDP Capability Negotiation Option Capability registry created for RFC 5939 [RFC5939].

5.3. SDP Capability Negotiation Configuration Parameters Registry

The IANA is hereby requested to change the "SDP Capability Negotiation Potential Configuration Parameters" registry currently registered and defined by RFC 5939 [RFC5939] as follows:

The name of the registry should be "SDP Capability Negotiation Configuration Parameters Registry" and it should contain a table with the following column headings:

- o Encoding Name: The syntactical value used for the capability negotiation configuration parameter, as defined in RFC 5939 [RFC5939], Section 3.5.
- o Descriptive Name: The name commonly used to refer to the capability negotiation configuration parameter.
- o Potential Configuration Definition: A reference to the RFC that defines the configuration parameter in the context of a potential configuration attribute. If the configuration parameter is not defined for potential configurations, the string "N/A" (Not Applicable) MUST be present instead.
- o Actual Configuration Definition: A reference to the RFC that defines the configuration parameter in the context of an actual configuration attribute. If the configuration parameter is not defined for actual configurations, the string "N/A" (Not Applicable) MUST be present instead.

- o Latent Configuration Definition: A reference to the RFC that defines the configuration parameter in the context of a latent configuration attribute. If the configuration parameter is not defined for latent configurations, the string "N/A" (Not Applicable) MUST be present instead.

An IANA SDP Capability Negotiation Configuration registration MUST be documented in an RFC in accordance with the IETF Review policy [RFC5226]. Furthermore:

- o The RFC MUST define the syntax and semantics of each new potential configuration parameter.
- o The syntax MUST adhere to the syntax provided for extension configuration lists in RFC 5939 [RFC5939] Section 3.5.1 and the semantics MUST adhere to the semantics provided for extension configuration lists in RFC 5939 [RFC5939] Section 3.5.1 and 3.5.2.
- o Configuration parameters that apply to latent configurations MUST furthermore adhere to the syntax provided in Section 3.3.5 and the semantics defined overall in this document.
- o Associated with each registration MUST be the encoding name for the parameter as well as a short descriptive name for it.
- o Each registration MUST specify if it applies to
 - * Potential configurations
 - * Actual configurations
 - * Latent configurations

5.4. SDP Capability Negotiation Configuration Parameter Registrations

The IANA is hereby requested to register the following capability negotiation configuration parameters:

Encoding Name: a
Descriptive Name: Attribute Configuration
Potential Configuration Definition: [RFC5939]
Actual Configuration Definition: [RFC5939]
Latent Configuration Definition: [Note to RFC Editor: This RFC]

Encoding Name: t
Descriptive Name: Transport Protocol Configuration
Potential Configuration Definition: [RFC5939]
Actual Configuration Definition: [RFC5939]

Latent Configuration Definition: [Note to RFC Editor: This RFC]

Encoding Name: m

Descriptive Name: Media Configuration

Potential Configuration Definition: [Note to RFC Editor: This RFC]

Actual Configuration Definition: [Note to RFC Editor: This RFC]

Latent Configuration Definition: [Note to RFC Editor: This RFC]

Encoding Name: pt

Descriptive Name: Payload Type Number Mapping

Potential Configuration Definition: [Note to RFC Editor: This RFC]

Actual Configuration Definition: [Note to RFC Editor: This RFC]

Latent Configuration Definition: [Note to RFC Editor: This RFC]

Encoding Name: mt

Descriptive Name: Media Type

Potential Configuration Definition: N/A

Actual Configuration Definition: N/A

Latent Configuration Definition: [Note to RFC Editor: This RFC]

6. Security Considerations

The security considerations of RFC 5939 [RFC5939] apply for this document.

In RFC 5939 [RFC5939], it was noted that negotiation of transport protocols (e.g. secure and non-secure) and negotiation of keying methods and material are potential security issues that warrant integrity protection to remedy. Latent configuration support provides hints to the other side about capabilities supported for further offer/answer exchanges, including transport protocols and attribute capabilities, e.g. for keying methods. If an attacker can remove or alter latent configuration information to suggest that only insecure or less secure alternatives are supported, then he may be able to force negotiation of a less secure session than would otherwise have occurred. While the specific attack as described here differs from those described in RFC 5939 [RFC5939], the considerations and mitigation strategies are similar to those described in RFC 5939 [RFC5939].

Another variation on the above attack involves the Session Capability ("sescap") attribute defined in this document. The "sescap" enables a preference order to be specified for all the potential configurations, and that preference will take precedence over any preference indication provided in individual potential configuration attributes. Consequently, an attacker that can insert or modify a "sescap" attribute may be able to force negotiation of an insecure or less secure alternative than would otherwise have occurred. Again, the considerations and mitigation strategies are similar to those described in RFC 5939 [RFC5939].

The addition of negotiable media formats and their associated parameters, defined in this specification can cause problems for middleboxes which attempt to control bandwidth utilization, media flows, and/or processing resource consumption as part of network policy, but which do not understand the media capability negotiation feature. As for the initial SDP Capability Negotiation work [RFC5939], the SDP answer is formulated in such a way that it always carries the selected media encoding for every media stream selected. Pending an understanding of capabilities negotiation, the middlebox should examine the answer SDP to obtain the best picture of the media streams being established. As always, middleboxes can best do their job if they fully understand media capabilities negotiation.

7. Changes from previous versions

7.1. Changes from version 16

- o Changed grammar covering numeric values to forbid use of leading zeroes, and added text to that effect as well.
- o Clarified that all numerical ranges must be strictly increasing (from leftmost number to rightmost number).
- o Now mandating against (rather than recommending against) using real keying material in a crypto attribute when it is used for a latent configuration only.
- o Reworded Offer section description of config-number use in potential configurations to make it clear that all config-numbers must be unique.

7.2. Changes from version 15

- o Fixed style of RFC references to be consistent in body of document.
- o Minor updates to address Gen-ART review comments.

7.3. Changes from version 14

- o Updated IANA Considerations to fix configuration parameter registry. Document now updates RFC 5939 [RFC5939] (IANA considerations only)
- o Minor ABNF updates to fix errors.
- o Editorial nit fixes to address protocol write-up review.

7.4. Changes from version 13

- o Various editorial clarifications and updates to address review comments.

7.5. Changes from version 12

- o Removed "dummy" form in the pcfg payload-type-number, since the functionality is redundant with the non-RTP media capability (omcap) and it was inconsistent with other RTP payload type operation.

- o Clarified that latent configuration attribute (lcfg) can only be used at the media level and hence (technically) as part of a media description
- o Rewrote offer/answer sections and expanded significantly on offer/answer operation.
- o Updated security considerations
- o Various minor editorial clarifications and changes.

7.6. Changes from version 11

- o Corrected several statements implying lcfg was a session-level attribute.
- o Added non-RTP based media format capabilities ("a=omcap") and renamed "mcap" to "rmcap"

7.7. Changes from version 10

- o Defined the latent configuration attribute as a media-level attribute because it specifies a possible future media stream. Added text to clarify how to specify alternative configurations of a single latent stream and/or multiple streams.
- o Improved the definition of the session capability attribute to permit both required configurations and optional configurations - latent configurations cannot be required because they have not yet been offered.
- o Removed the special-case treatment of conflicts between base-level fmtp attributes and fmtp attributes generated for a configuration via invoked mcap and mfcap attributes.
- o Removed reference to bandwidth capability (bcap) attribute.
- o Changed various "must", etc., terms to normative terms ("MUST", etc.) as appropriate, in Section 3.3.5Section 3.3.6.1 Section 3.3.6.3 and Section 3.3.8
- o Attempted to clarify the substitution mechanism in Section 3.3.7 and improve its uniqueness.
- o Made various editorial changes, including changing the title in the header, and removing numbering from some SDP examples.

7.8. Changes from version 09

- o Additional corrections to latent media stream example in Section 4.3
- o Fixed up attribute formatting examples and corresponding ABNF.
- o Removed preference rule for latent configurations.
- o Various spelling and other editorial changes were made.
- o updated cross-references.

7.9. Changes from version 08

The major change is in Section 4.3, Latent Media Streams, fixing the syntax of the answer. All the other changes are editorial.

7.10. Changes from version 04

- o The definitions for bcap, ccap, icap, and kcap attributes have been removed, and are to be defined in another document.
- o Corrected formatting of m= and p= configuration parameters to conform to extension-config-list form defined in RFC 5939 [RFC5939]
- o Reorganized definitions of new parameters to make them easier to find in document.
- o Added ability to renegotiate capabilities when modifying the session (Section 3.4.4).
- o Made various editorial changes, clarifications, and typo corrections.

7.11. Changes from version 03

- o A new session capability attribute (sescap) has been added to permit specification of acceptable media stream combinations.
- o Capability attribute definitions corresponding to the i, c, b, and k SDP line types have been added for completeness.
- o Use of the pcfg: attribute in SDP answers has been included in order to conveniently return information in the answer about acceptable configurations in the media stream offer.

- o The use of the `lcfg:` attribute(s) in SDP answers has been restricted to indicate just which latent configuration offers would be acceptable to the answerer.
- o A suggestion for "naive" middleboxes has been added to the Security Considerations.
- o Various editorial changes have been made.
- o Several errors/omissions have been corrected.
- o The description of the `mscap` attribute has been modified to make it clear that it should not be used to generate undefined SDP attributes, or to "extend" existing attributes.
- o `<ms-parameters>` are made optional in the `mscap` attribute definition.
- o "AMR" changed to "AMR-WB" in cases in which the sample rate is 16000.

7.12. Changes from version 02

This version contains several detail changes intended to simplify capability processing and mapping into conventional SDP media blocks.

- o The "mcap" attribute is enhanced to include the role of the "ecap" attribute; the latter is eliminated.
- o The "fcap" attribute has been renamed "mfcap". New replacement rules vis-a-vis `fntp` attributes in the base media specification have been added.
- o A new "mscap" attribute is defined to handle the problem of attributes (other than `rtpmap` and `fntp`) that are specific to a particular payload type.
- o New rules for processing the `mcap`, `mfcap`, and `mscap` attributes, and overriding standard `rtpmap`, `fntp`, or other media-specific attributes, are put forward to reduce the need to use the deletion option in the `a=` parameter of the potential configuration (`pcfg`) attribute.
- o A new parameter, "mt=" is added to the latent configuration attribute (`lcfg`) to specify the media stream type (audio, video, etc.) when the `lcfg` is declared at the session level.

- o The examples are expanded.
- o Numerous typos and misspellings have been corrected.

7.13. Changes from version 01

The documents adds a new attribute for specifying bandwidth capability and a parameter to list in the potential configuration. Other changes are to align the document with the terminology and attribute names from draft-ietf-mmusic-sdp-capability-negotiation-07. The document also clarifies some previous open issues.

7.14. Changes from version 00

The major changes include taking out the "mcap" and "cptmap" parameter. The mapping of payload type is now in the "pt" parameter of "pcfg". Media subtype need to explicitly defined in the "cmed" attribute if referenced in the "pcfg"

8. Acknowledgements

This document is heavily influenced by the discussions and work done by the SDP Capability Negotiation Design team. The following people in particular provided useful comments and suggestions to either the document itself or the overall direction of the solution defined herein: Cullen Jennings, Matt Lepinski, Joerg Ott, Colin Perkins, and Thomas Stach.

We thank Ingemar Johansson and Magnus Westerlund for examples that stimulated this work, and for critical reading of the document. We also thank Cullen Jennings, Christer Holmberg, and Miguel Garcia for their review of the document.

9. References

9.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, May 2008.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.
- [RFC5939] Andreassen, F., "Session Description Protocol (SDP) Capability Negotiation", RFC 5939, September 2010.

9.2. Informative References

- [RFC2198] Perkins, C., Kouvelas, I., Hodson, O., Hardman, V., Handley, M., Bolot, J., Vega-Garcia, A., and S. Fosse-Parisis, "RTP Payload for Redundant Audio Data", RFC 2198, September 1997.
- [RFC4568] Andreassen, F., Baugher, M., and D. Wing, "Session Description Protocol (SDP) Security Descriptions for Media Streams", RFC 4568, July 2006.
- [RFC4585] Ott, J., Wenger, S., Sato, N., Burmeister, C., and J. Rey, "Extended RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/AVPF)", RFC 4585, July 2006.
- [RFC4733] Schulzrinne, H. and T. Taylor, "RTP Payload for DTMF Digits, Telephony Tones, and Telephony Signals", RFC 4733, December 2006.
- [RFC4867] Sjöberg, J., Westerlund, M., Lankaniemi, A., and Q. Xie, "RTP Payload Format and File Storage Format for the Adaptive Multi-Rate (AMR) and Adaptive Multi-Rate Wideband (AMR-WB) Audio Codecs", RFC 4867, April 2007.

- [RFC5104] Wenger, S., Chandra, U., Westerlund, M., and B. Burman,
"Codec Control Messages in the RTP Audio-Visual Profile
with Feedback (AVPF)", RFC 5104, February 2008.

Authors' Addresses

Robert R Gilman
Independent
3243 W. 11th Ave. Dr.
Broomfield, CO 80020
USA

Email: bob_gilman@comcast.net

Roni Even
Gesher Erova Ltd
14 David Hamelech
Tel Aviv 64953
Israel

Email: ron.even.tlv@gmail.com

Flemming Andreassen
Cisco Systems
Iselin, NJ
USA

Email: fandreas@cisco.com

MMUSIC Working Group
Internet-Draft
Intended status: Standards Track
Expires: April 26, 2012

T. Frankkila
M. Westerlund
B. Burman
Ericsson
October 24, 2011

Extensible Bandwidth Attribute for SDP
draft-westerlund-mmusic-sdp-bw-attribute-00

Abstract

Knowledge of what bandwidths the end-points intend to use is important both for the other end-point and for resource allocation in various types of networks. This is especially important for wireless access networks which typically have quite limited resources. The bandwidth attribute in Session Description Protocol (SDP), 'b=AS', is today quite widely used to define the bandwidth that the end-points intends to use, in various types of sessions. This document will show that the existing bandwidth attribute, such as 'b=AS', although widely used in todays scenarios, has limitations that make it hard or even impossible for the end-points to express their intentions accurately when it comes to bandwidth usage. To solve the identified problems, this document defines a new extensible SDP bandwidth attribute 'a=bw' which enables more detailed control over the bandwidth declarations, request, and allocations. With the new bandwidth attribute it is possible to define different scopes in the session setup and then negotiate the bandwidth individually for each scope.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 26, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	4
2. Definitions	4
2.1. Requirements Language	4
2.2. Terminology	5
3. Use Cases and Design Rationale	5
3.1. Existing Bandwidth Attribute	6
3.1.1. Attribute Definition	6
3.1.2. Offer/answer Procedure for the Existing Bandwidth Attribute	7
3.1.3. End-point Behavior when Generating Traffic	7
3.2. Point-to-point Sessions using SDP offer/answer	8
3.2.1. Symmetric Point-to-point Sessions, Fixed-rate Codecs	8
3.2.2. Symmetric Point-to-Point Sessions with Rate-Adaptive Codec	10
3.2.3. Symmetric Point-to-Point Sessions with Several Rate-Adaptive Codecs	11
3.2.4. Asymmetric Point-to-Point Sessions	12
3.3. Sessions with Multiple Streams	14
3.3.1. Multiple Streams	14
3.4. User Experience and Bandwidth Negotiation	15
3.5. Summary of Findings	15
4. Attribute Specification	17
4.1. SDP Grammar	17
4.2. Declarative Use	21
4.3. Usage in Offer/Answer	21
4.4. Bucket Size Estimation	23
4.4.1. Sender Specified Token Bucket	24
4.4.2. Receiver Specified Token Bucket	25
4.4.3. Bucket Adjustment in Middle Nodes	25
4.4.4. Network Policing	25
4.4.5. Utilizing Network Feedback	26

4.5.	SDP Examples for Point-to-point Sessions	26
4.5.1.	Symmetric Fixed-rate Codecs	26
4.5.2.	Symmetric Rate-Adaptive Codec	26
4.5.3.	Symmetric Several Rate-Adaptive Codecs	27
4.5.4.	Asymmetric Session	28
4.5.5.	Session with Retransmission	29
4.6.	SDP Examples with Sessions with Multiple Streams	29
4.6.1.	Multiple Streams	29
4.6.2.	Declarative Example with Stream Asymmetry	30
4.7.	Interoperability Issues	30
4.7.1.	Interoperability with Existing Bandwidth Attribute	31
4.7.2.	Interoperability with Existing Directional Attribute	33
5.	Rules and Recommendations for Extensions	34
5.1.	Directionality	34
5.2.	Scope	34
5.3.	Semantics	34
5.4.	Values	34
6.	Open Issues	35
7.	IANA Considerations	35
8.	Security Considerations	35
9.	Acknowledgements	36
10.	References	36
10.1.	Normative References	36
10.2.	Informative References	37
	Authors' Addresses	37

1. Introduction

This document looks at the issues of non-basic usage of RTP [RFC3550] and analyzes how well the existing SDP [RFC4566] attribute 'b=AS' for bandwidth negotiation performs in different scenarios.

This analysis is done by defining a number of use cases, containing sessions with:

- o single and multiple media types;
- o symmetric and asymmetric media streams;
- o single and multiple media sources, including multiple sources from the same end-point;
- o multiple end-points each having one or more media sources, including applications that use multiple encodings of a particular media.

It is shown that the existing bandwidth attributes 'b=AS' [RFC4566] and 'b=TIAS' [RFC3890] has limitations which make it unclear or even impossible for end-points and for resource allocation functions in the network to determine how much bandwidth the service will use. The analysis also provides the design rationale for the new bandwidth attribute.

This document then proposes a general and extensible mechanism for bandwidth negotiation that can be used for any type of session. Interoperability with the existing mechanisms for bandwidth negotiation is especially important since the existing bandwidth attribute has a wide-spread usage.

This document also presents several examples for how the new bandwidth attribute can be used in the session setup phase for various types of sessions. The examples are derived for IP/UDP/RTP transport although nothing should prevent using the new bandwidth attribute also for other transport protocols.

2. Definitions

2.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2.2. Terminology

The following terms and abbreviations are used in this document:

Bandwidth: In this document, the bandwidth is defined as the IP level bandwidth, i.e. including the network protocol (IPv4 or IPv6) and transport protocol (TCP, UDP, RTP, etc) overhead. When RTP is used then the RTCP bandwidth is handled separately from the bandwidth used for RTP packets. Bandwidth in this context is in the unit bits per second, not Hz.

Encoding: A particular encoding is the choice of the media encoder (codec) that has been used to compress the media. Different encodings result in the fidelity of that encoding through the choice of sampling, bit-rate and other configuration parameters.

End-point: A single entity sending and/or receiving RTP packets. It may be decomposed into several functional blocks, but as long as it behaves a single RTP stack entity it is classified as a single end-point.

Media stream: A sequence of RTP packets using a single SSRC that together carry part or all of the content of a specific Media Type from a specific sender source within a given RTP session.

RTP session: An RTP session consists of one or more media streams that have the same purpose. The typical example is to have one RTP session per media type, i.e. that voice and video use different RTP sessions (different ports) since they have different purpose. It is however possible to have multiple streams in an RTP session, for example when having both a stream for non-redundant audio and another stream for re-transmissions of audio packets. The fundamental definition of an RTP session is a single SSRC space.

3. Use Cases and Design Rationale

This section describes a number of use cases where the existing bandwidth attribute 'b=AS' is used for bandwidth definition. It also discusses why the limitations of the existing bandwidth attribute makes it hard for other end-points and resource allocation functions to know or estimate how much bandwidth that will be used in the ongoing session.

The analysis is made by defining a set of use cases. The first use cases include fairly simple session types, i.e. point-to-point

sessions with or without asymmetry. A few more complex use cases are then analyzed. The last set up use cases reflect fairly advanced session types, e.g. various variants of multiplexing and usage of multiple media streams.

The discussion is then summarized and the design rationales for the new bandwidth attribute are outlined.

3.1. Existing Bandwidth Attribute

The existing bandwidth modifier 'b=' defined in RFC 4566 [RFC4566] is reviewed in this section.

3.1.1. Attribute Definition

The existing bandwidth attribute 'b=' is defined in Section 5.8 of RFC 4566 [RFC4566]. The syntax is:

b=<bwtype>:<bandwidth>

where:

<bwtype> is either:

'AS' ("Application Specific"), which is the maximum bandwidth as estimated by the application; or:

'CT' ("Conference Total"), which is the total bandwidth for all media at all sites.

<bandwidth> is the bandwidth value in kilobits per second.

Bandwidth types have been defined for the negotiation of the RTCP bandwidth using 'b=RS' and 'b=RR', RFC 3556 [RFC3556].

There is also a bandwidth type for negotiating the transport independent application specific maximum bandwidth, 'b=TIAS', RFC 3890 [RFC3890]. This bandwidth type is similar to the 'b=AS' bandwidth type, except that the overhead caused by the transport protocol headers is not included.

One issue with the existing bandwidth attribute is that the syntax is very limited since it only allows for defining new bandwidth types (<bwtype>) and their respective single numerical value. This limitation needs to be considered in the discussion below.

3.1.2. Offer/answer Procedure for the Existing Bandwidth Attribute

"An Offer/Answer Model with the Session Description Protocol (SDP)" [RFC3264] describes the offer/answer procedures for the existing bandwidth attribute. For the SDP offer, it describes that the bandwidth attribute indicates the desired bandwidth that the offerer would like to receive. For the SDP answer, it describes that the bandwidth attribute indicates the bandwidth that the answerer would like the offerer to use when sending media. Thus, for offer/answer negotiations, the bandwidth attribute indicates the bandwidth for the receive direction of each end-point.

The solution presented in this document focuses primarily on clarifying and assisting the Application Specific (AS) bandwidth.

[It is an open question to decide if and how to handle the RTCP bandwidth negotiation, e.g. corresponding to b=RS and b=RR.]

[It is an open question to develop semantics for the transport independent bandwidth negotiation, e.g. corresponding to b=TIAS.]

3.1.3. End-point Behavior when Generating Traffic

When an end-point is sending media then this can be done in many different ways, depending on the choices the implementers have made.

Some end-points may send it's data in a fairly "nice and smooth" media stream, which means that both the packet sizes and the packet rates are more or less constant all the time. An example of a smooth stream is when the end-point is encoding speech and is sending one packet every 20 ms and when the packets are of equal size.

Other end-points may generate bursty streams, which have a large peak-to-average ratio. An example of a bursty stream is when an end-point is encoding video. Most of the time, the end-point is sending packets with almost the same size and with constant packet rate. However, it happens occasionally that the encoder generates much more data for a frame, which may give a very large packet size. It may even happen that the sender has to segment the data into several packets, which may be transmitted in a burst, thereby causing a very high peak rate.

Whether the stream is smooth or bursty makes a big difference for the network and the policy control that usually applies in QoS controlled networks. If the stream is too bursty, then a policy control function may decide to drop packets that exceed the granted rate. This will lead to degraded quality and reduced user satisfaction.

The existing bandwidth attribute offers no mechanism to negotiate what temporal variations that can be allowed for a stream. The only available mechanism is to negotiate the maximum bandwidth, but there is nothing that defines any kind of averaging window (or something similar) that can be used to control the bandwidth variations from the transmitted stream.

It is therefore proposed to use a Token Bucket model to describe the bandwidth with two parameters, the token bucket rate and the bucket size, see RFC 2212 [RFC2212].

3.2. Point-to-point Sessions using SDP offer/answer

The existing modifier for the application specific bandwidth 'b=AS' is frequently used in the SDP offer/answer negotiation RFC 3264 [RFC3264] for setting up point-to-point sessions, for example for bi-directional point-to-point VoIP or video telephony sessions. In this section, the use of the legacy bandwidth modifier is reviewed for the use in point-to-point sessions using SDP offer/answer.

3.2.1. Symmetric Point-to-point Sessions, Fixed-rate Codecs

This example below shows the SDP offer from end-point A for several fixed-rate codecs, mu-law and A-law PCM/G.711 [G.711], AD-PCM/G.726 [G.726] and CS-ACELP/G.729 [G.729]. The codecs have different bit rates. PCM encodes speech at 64 kbps. G.726 can encode speech at four different rates, 64, 32, 24 and 16 kbps, but in this case it is assumed that the 32 kbps variant is used. G.729 encodes speech at 8 kbps. The IP/UDP/RTP overhead with 20 ms packetization and IPv4 becomes 16 kbps in all cases giving 80, 48 and 24 kbps, respectively.

```
m=audio 49200 RTP/AVP 8 0 96 18
b=AS:80
a=rtpmap:96 G726-32/8000/1
a=ptime:20
a=maxptime:80
```

SDP offer for mu-law PCM, A-law PCM, G.726 and G.729 with IPv4

If end-point B accepts to use this codec then a likely SDP answer would be:

```
m=audio 49400 RTP/AVP 8 0 96 18
b=AS:80
a=rtpmap:96 G726-32/8000/1
a=ptime:20
a=maxptime:80
```

SDP answer for mu-law PCM, A-law PCM, G.726 and G.729 with IPv4

In this case, both end-points offer to receive 80 kbps. A resource allocation function would thereby allocate 80 kbps in each direction.

However, if end-point B accepts to use one of the lower rate codecs, for example G.729, but not the PCM codecs, then a likely SDP answer would be:

```
m=audio 49400 RTP/AVP 18
b=AS:24
a=ptime:20
a=maxptime:80
```

SDP answer for G.729 with IPv4

This means that the offerer has offered to receive 80 kbps while the answerer has offered to receive 24 kbps. In the direction A to B it is clear that a resource allocation function should allocate 24 kbps. However, in the direction B to A it is a little more unclear. On one hand, end-point A has offered to receive 80 kbps. But, on the other hand, end-point B has only indicated support for the G.729 codec and its unknown if B can send with something in addition to G.729 from A's offered set.

A resource allocation may also (incorrectly) conclude that end-point B will also send maximum 24 kbps, since b=AS indicates 24 kbps. But, since maxptime is 80 ms, this means that end-point B could very well use application layer redundancy and encapsulate redundant frames together with non-redundant frames, which would result in a bandwidth exceeding 24 kbps. Even if maxptime would be 20 ms, end-point B could still use application layer redundancy, if the non-redundant and redundant frames are transmitted in different packets. This is possible since end-point A has indicated that it is capable of receiving 80 kbps. Hence, if the resource allocation function uses the codec information and assumes that end-point B will send with only 24 kbps, then this may cause packet losses and/or long delays.

It should be clear with this example that the current bandwidth attribute, b=AS, can create ambiguities related to what bandwidth that will be used in each direction. If the end-points and the resource allocation functions make different interpretations then there is a risk for either poor quality or wasted resources.

To solve this, a new bandwidth negotiation method should enable negotiating different bandwidths for different codecs. If a codec can be configured in several different ways, e.g. G.726 offers the possibility to use four different static bit rates then this would typically be negotiated using different RTP Payload Types. This means that the solution needs to be capable of negotiating different bandwidths for different Payload Types.

3.2.2. Symmetric Point-to-Point Sessions with Rate-Adaptive Codec

This use case describes what might happen when using rate-adaptive codecs in a session, for example AMR [AMR]. The rate adaptation should adapt to a high bitrate when the operating conditions are good, but should adapt to a low bitrate when the operating conditions are degraded, e.g. due to congestion or bad coverage.

One example of the SDP offer-answer negotiation for rate-adaptive codec is shown below.

```
m=audio 49200 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
aptime:20
a=maxptime:100
```

SDP offer from end-point A for AMR and IPv4

The bandwidth attribute in the SDP indicates the bandwidth that the offerer would like to receive, RFC 3264 [RFC3264].

```
m=audio 49100 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
aptime:20
a=maxptime:100
```

SDP answer from end-point B also for AMR and IPv4

The bandwidth attribute in the SDP answer indicates the maximum bandwidth that the answerer would like the offerer to use when sending media, RFC 3264 [RFC3264].

In this case, it is clear that both end-points are prepared to receive up to 29 kbps of media. Since AMR can adapt the rate for the encoding, this means that the bandwidth can be reduced, e.g. to the 5.9 kbps mode, if congestion is detected. The existing bandwidth attribute 'b=AS' is however only used to negotiate the maximum rate. This means that there is nothing in the SDPs that describes how the rate will be adapted. In some cases, usually for speech codec, it might be possible to derive the lowest rate from the codec information. However, there is no guarantee that the end-points will adapt to this rate or whether it will stay at some higher rate. For video codecs, there is usually no codec information at all that could be used to determine how low rate the end-points will use. The lowest usable rate for a video codec is generally not a video codec limitation, but rather some end-user or service consideration on what

is the lowest video quality that is still useful or acceptable in the actual scenario.

This means that a resource allocation function has no information which could be used to determine how the end-points will adapt during periods of congestion. Hence the network does not know what to assume from the end-points.

To solve this, a new bandwidth negotiation method should allow for negotiating not only the highest rate but also the minimum rate that is still useful.

3.2.3. Symmetric Point-to-Point Sessions with Several Rate-Adaptive Codecs

Another example is when the originating end-point offers several rate-adaptive codecs, with different bandwidths, and when the answerer only support one or several of the lower-rate configurations but not the configuration that uses the highest bandwidth. With the legacy bandwidth modifier 'b=AS' it is only possible to indicate one bandwidth for the whole RTP session, which means that the end-point needs to indicate the highest bandwidth since this is the worst-case scenario. An offer/answer for this case is shown below. The offerer supports both AMR and AMR-WB AMR-WB [AMR-WB] and therefore indicates the bandwidth needed for the AMR-WB configuration since it is higher than for AMR. If the answerer does not support the AMR-WB codec then it will have to remove this configuration from the SDP when creating the SDP answer. This means that the answerer calculates the bandwidth required for AMR instead of AMR-WB.

```
m=audio 49200 RTP/AVP 96 97
b=AS:41
a=rtpmap:96 AMR-WB/16000/1
a=fmtp:96 mode-change-capability=2; max-red=80
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=ptime:20
a=maxptime:100
```

SDP offer from end-point A for AMR-WB, AMR and IPv4

```
m=audio 49100 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=ptime:20
a=maxptime:100
```

SDP answer from end-point B for AMR and IPv4 (AMR-WB is removed)

Since the indicated bandwidth is for the receiving direction in this example this means that:

- o A must send media with a bandwidth not exceeding 29 kbps; and:
- o B must send media with a bandwidth not exceeding 41 kbps.

This gives the same problem with ambiguous maximum rate as shown in Section 3.2.1. In addition, since both AMR and AMR-WB are rate-adaptive codecs, with different bit rates, they also have different minimum rates. This means that a resource allocation would be unaware about both the maximum bandwidth and the minimum (required) bandwidth.

To solve this, a new bandwidth attribute should allow for negotiating both maximum and minimum bitrates individually for each payload type.

For speech codecs, it is usually possible to derive the minimum rate from the codec information. However, this is typically not possible for video codecs since they only indicate the maximum encoding level. For example, if end-point A offers to use H.264 level 3.0 H.264 [H.264] but end-point B is only capable of using level 1.2, then this only limits the maximum bandwidth in the direction from A to B. In the other direction, end-point A is still capable of receiving level 3.0.

3.2.4. Asymmetric Point-to-Point Sessions

The session setup for asymmetric streams is not always straight forward. Lets say that one want to set up a session with 600 kbps in the sending direction and 200 kbps in the receiving direction.

```
m=video 49200 RTP/AVP 96
b=AS:200
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c00c
a=sendrecv
```

SDP offer to receive 200 kbps video

From this SDP, it can be determined that the end-point wants to receive 200 kbps. There is some implicit information in the level part of the profile-level-id for the H.264 example above, indicating that the end-point can send using a higher bandwidth (up to 768 kbps), but it requires codec-specific knowledge to be able to extract that implicit information. In this example, lets assume that the sender does not even want to utilize the maximum allowed bandwidth for the signaled level, but a slightly lower one, say 600 kbps. So how is the answerer supposed to know that the offerer really wants to

send up to 600 kbps, especially since not even the implicit level-related can be used? There could be many reasons to use a lower video bandwidth than the one defined as level maximum; limited terminal performance in the send direction, a known network bandwidth limitation, a bandwidth charging model that makes the user prefer a lower bandwidth, etc.

One way to express the asymmetry is to set up different RTP sessions for sending and receiving directions. An SDP offer for this might be:

```
m=video 49200 RTP/AVP 96
b=AS:600
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c00d
a=sendonly
m=video 49202 RTP/AVP 97
b=AS:200
a=rtpmap:97 H264/90000
a=fmtp:97 profile-level-id=42c00c
a=recvonly
```

SDP offer with separate sessions for send and receive

If the answerer decides to accept this then the SDP answer might be:

```
m=video 49200 RTP/AVP 96
b=AS:600
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c00d
a=recvonly
m=video 49202 RTP/AVP 97
b=AS:200
a=rtpmap:97 H264/90000
a=fmtp:97 profile-level-id=42c00c
a=sendonly
```

SDP answer with separate sessions for send and receive

In this example, it is clear that the offerer can send video with 600 kpbs and receive video with up to 200 kbps. However, if the offer is for different codecs, using different bandwidths, then one have the same problem as described in Section 3.2.3.

Specifically for video, but possibly also for other media, it may happen that different implementations send the media in different ways. Some implementations may try to provide a fairly "smooth" stream in terms of bandwidth variation over time, while other implementations may give a very "bursty" stream.

There also exist cases where opening additional RTP sessions just for expressing asymmetric transmission bandwidths are not desirable.

3.3. Sessions with Multiple Streams

In this part of the analysis, it is assumed that an RTP session is set up for multiple streams. This can be done in several ways and for several reasons, as discussed in RTP Multiplexing Architecture [I-D.westerlund-avtcore-multiplex-architecture].

3.3.1. Multiple Streams

The assumed usage here is a multi-party session, for example a video conference using an RTP mixer. Some of the attendees are active and their audio and video is distributed to the other users. Some attendees are inactive and thus only receive media. In this example, each end-point sends one video stream, but can receive up to four simultaneous video streams, multiplexed as different SSRC in the same RTP session. One or more central nodes (RTP Mixer) are used to help facilitate the media transport between the participants, and are involved in choosing the streams to be forwarded. Assume that there is an aggregate bandwidth limit of 3 Mbps in the receive direction, and that each received video stream should be limited to max 1 Mbps.

An SDP offer for the setting up a session with one video stream for the sending direction and four video streams for the receiving direction is shown below when using [I-D.westerlund-avtcore-max-ssrc] to explicitly declare capability to handle multiple streams. In this case, only the legacy 'b=AS' bandwidth attribute is used, valid only for the aggregate.

```
m=video 49300 RTP/AVP 96
b=AS:3000
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c016
a=max-recv-ssrc:* 4
```

SDP offer to receive multiple video streams

This example again highlights the asymmetry problem with the existing bandwidth attribute, but it also highlights the lack of per-stream bandwidth specification. This means that it is not possible to declare the 1 Mbps bandwidth limit that should be used for each one of the for streams in the receiving direction, which thus is a desirable property of the new bandwidth attribute. Note also that in this example, the 1 Mbps limit per stream cannot be fully utilized if all four streams are used simultaneously.

3.4. User Experience and Bandwidth Negotiation

Resource allocation is typically a compromise between perceived quality and network utilization. From an end-user perspective, the bandwidth for a service should be as high rate as possible, since this should give the best user experience. However, from a network perspective, one would like to minimize the rate, since this should maximize the number of sessions that can be supported.

For some services, like conversational voice- and/or video-telephony, one needs to ensure that the network is capable of delivering a certain at least required rate, even when the network load is high. This is needed to ensure user satisfaction, both in terms of quality and end-to-end delay. This means that the end-points and the network need to agree on what maximum bandwidth that can be used for the session as well as some lowest useful "least required" bandwidth.

The current bandwidth modifier, 'b=AS', is used to negotiate the maximum bandwidth. However, since it only allows for negotiate one bandwidth it cannot be used to also negotiate a lower bandwidth limit.

To solve this, a new bandwidth negotiation method should allow for negotiating not only the highest rate but also the "at least required" rate. To enable a negotiation between the end-point and the network, a reasonable approach is that the end-point requests a lower bandwidth limit and then the network indicate what "least required" rate that was granted.

3.5. Summary of Findings

It should be clear from the above discussion that the current bandwidth attribute is too limited to be used for all use cases and that some extensions are needed.

The current bandwidth attribute, 'b=AS', is sufficient for simple sessions but gives ambiguities when negotiating more advanced session types. One of the drawbacks is that 'b=AS' only indicates the desired bandwidth for the receiving direction but, if the answering end-point wants to use a lower rate than what is offered, then there is often no way for the resource allocation function to know what bandwidth that will be used in the offerer's sending direction.

Implementers of end-points and resource allocation functions may try to resolve this ambiguity by using other information available in the SDP, e.g. codec-specific information. However, such information is not always easily available, e.g. for video codecs.

End-points may have to perform a second offer/answer negotiation to resolve the ambiguity. This, obviously, has the drawbacks that the SIP traffic is increased and that this takes some extra time. It is also not guaranteed that the end-points will actually initiate a second offer/answer negotiation.

The analysis above has also shown that the current bandwidth attribute is insufficient to properly describe the session for multi-stream scenarios.

The analysis above has also shown that the current bandwidth modifier can be used to negotiate the maximum bit rate in bearers allocated in some wireless networks, but it is insufficient for also negotiating a lower, "least required", bandwidth limit.

Another problem with the existing bandwidth attribute is that the syntax is very limited and does not allow for introducing extensions, only additional identifiers with a single value each.

It is therefore proposed to define a new bandwidth attribute, including a new syntax. The new bandwidth attribute should support:

Directionality: One need to be able to have different sets of attributes values depending on direction.

Payload specific: With the new bandwidth attribute it should be possible to specify different bandwidth values for different RTP Payload types. This is because some codecs have different characteristics and one may want to limit a specific codec and payload configuration to a particular bandwidth. Especially combined with codec negotiation there is a need to express intentions and limitations on usage for that particular codec. In addition, payload agnostic information is also needed.

Multiple streams: The new bandwidth attribute should support bandwidth negotiation both for single streams and for multiple streams. When multiple streams are used, the new bandwidth attribute should allow for declaring both the bandwidth per stream and the aggregated bandwidth.

Bandwidth specification method: To have a clear specification of what any bit-rate values mean we propose that Token bucket parameters should be used, i.e. bucket depth and bucket fill rate, where appropriate for the semantics. If single values are to be specified, a clear definition on how to derive that value must be specified, including averaging intervals etc.

Bandwidth semantics: It should be possible to negotiate different types of bandwidths for each scope, including several bandwidth properties in the same negotiation. It should, at least, be possible to negotiate the highest bandwidth and a lower bandwidth limit that indicates the lowest useful bandwidth to use the related media. The least required bandwidth limit should ideally, but need not necessarily, be guaranteed by the network and the remote end-point(s).

Extensibility: The semantics need to be extensible, so that new semantics can be defined in the future.

The existing bandwidth modifier, 'b=AS', is widely used today. The existing SDP attributes for directionality, 'a=sendrecv', 'a=recvonly', 'a=sendonly' and 'a=inactive', are also widely used. It is therefore important to ensure interworking between the new bandwidth attribute and the mechanisms already existing in SDP.

4. Attribute Specification

This section proposes a new bandwidth attribute 'a=bw' that can be used either as an extension to the already existing bandwidth attribute 'b=AS' or replacing the existing bandwidth attribute. The new bandwidth attribute includes semantics that allows for also replacing the existing bandwidth attribute.

The syntax for the new bandwidth attribute is:
a=bw:<direction> <scope> <semantic>:<value>

where:

<direction> is the direction in which the scope and semantics applies,

<scope> describes for what scope the definitions applies,

<semantic> is the actual bandwidth specification,

<value> in the form defined by the semantic used.

The new attribute is designed to allow for future extendability.

4.1. SDP Grammar

The ABNF RFC 5234 [RFC5234] for this attribute is the following:

```

bw-attrib      = "a=bw:" direction SP [req] scope SP
                  [req] semantics ":" values
direction      = "send" / "recv" / "sendrecv" / direction-ext
scope          = payloadType / scope-ext
payloadType    = "pt=" ("*" / (PT-spec) *("," PT-spec))
PT-spec        = PT-value / PT-value-range
PT-value       = 1*3DIGIT
PT-value-range = PT-value "-" PT-value
req            = "!"
semantics       = "SMT" / "AMT" / "SLT" / "SLTR" / "ALT" / "ALTR" /
                  semantics-ext
values         = token-bucket / value-ext
token-bucket   = "tb=" br-value ":" bs-value
br-value       = "*" / 1*15DIGIT ; Bucket Rate [bps]
bs-value       = "*" / 1*15DIGIT ; Bucket Size [bytes]

direction-ext  = token ; As defined in RFC 4566
scope-ext      = 1*VCHAR ; As defined in RFC 5234
semantics-ext  = token ; As defined in RFC 4566
value-ext      = 0*(WSP / VCHAR) ; As defined in RFC 5234

```

ABNF for 'bw' attribute

The 'a=bw' attribute defines three possible directionalities for the bandwidth:

send: In the send direction for SDP Offer/Answer agent or in case of declarative use in relation to the device that is being configured by the SDP.

recv: In the receiving direction for the SDP Offer/Answer agent providing the SDP or in case of declarative use in relation to the device that is being configured by the SDP.

sendrecv: The provided bandwidth values apply equally in send and receive directions, i.e. the values configures the directions symmetrically.

The directionality must be specified when the 'a=bw' attribute is used. Only one directionality can be specified on each 'a=bw' line. Special care must be taken to avoid conflicting definitions. For example, if 'sendrecv' has been specified on one 'a=bw' line for a scope, e.g. payload number 96, then the direction cannot be set to 'send' or 'recv' on another 'a=bw' line for the same scope. However, it is allowed to specify directionality 'send' on one 'a=bw' line for a scope and directionality 'recv' on another 'a=bw' line. This is useful when the bandwidth is different in different directions. Using 'sendrecv' as directionality on an 'a=bw' line is a shortcut in

the sense that it is equivalent to using two separate 'a=bw' lines where one uses 'send' and the other 'recv' but that otherwise are semantically identical.

The scope indicates what is being configured by the bandwidth semantics on this attribute line. Two different scopes are defined based on payload type:

Payload Type: The bandwidth configuration applies to the specific payload type value(s).

pt=*: Applies to all payload types being used.

Using pt=* indicates that the definitions apply to all payload types being used. The scope may be a single payload type value, e.g. pt=96. A list of payload type values can be created by using a comma-separated list, e.g. pt=96,98,105. It is also possible to specify a range of payload type values, e.g. pt=96-102, which means that the definitions apply to all the payload type numbers from 96 to 102. It is also possible to combine payload type values, payload type lists and payload type ranges, e.g. pt=96,98-102,104,105,110-113.

The scope parameter is extensible to allow for adding other scope definitions in the future.

This specification defines six related semantics. All semantics represent either the bandwidth consumption of a single stream or the aggregate of streams as a token bucket defining a transmission profile which the media sender must stay within. The token bucket values are the token rate in bits per second and the bucket size in bytes both provided as integers, see RFC 2212 [RFC2212]. The below semantics includes the whole IP packet, for example IP, UDP, RTP headers and RTP payload, as what shall be metered when determining if the send pattern is within the profile. The token bucket definition allows for wild cards enable to specify that one want a value as token bucket, but has no proposed value.

The definitions of the semantics in more detail are:

SMT (Stream Maximum Token bucket): The maximum intended or allowed bandwidth usage, including protocol overhead, for each individual source (each SSRC) in an RTP session at the sender side specified by a token bucket. The token bucket wild cards ("*") should not be used for the SMT semantics since it should always be possible to estimate the maximum bandwidth. This semantics is possible to use with the scope for any payload type (pt=*) where it applies independent of encoding and packetization, or for a specific or a

set of payload type(s).

AMT (Aggregate Maximum Token bucket): The maximum intended or allowed bandwidth usage for the sum of all sources (SSRCs) in an RTP session according to the specified directionality at the media sender specified by a token bucket. The 'sendrecv' directionality parameter indicates equal token buckets in both directions, i.e. the aggregate of streams sent to an end-point shall be within the token bucket defined transmission profile, and the aggregate of streams sent from that end-point shall also be within the same token bucket profile at the sender. It can be used either to express the maximum for one particular payload type, for a set of payload types or for any payload type (pt=*). The token bucket wild card ("*") should not be used for the AMT semantics since it should always be possible to estimate the maximum bandwidth.

SLT (Stream Least required Token bucket): The least required bandwidth, including IP protocol overhead, needed for the stream for each individual source (each SSRC) in an RTP session as specified by a token bucket at the sender. When using the SLT semantic, the SMT semantic SHOULD also be specified for the same direction and scope. If the SLT semantics is not defined then this means that the least required bandwidth limit is zero. The least required bandwidth is the minimum bandwidth that is necessary for the service to work with usable quality.

SLTR (Stream Least required Token bucket Request): The request for establishing the least required bandwidth, including protocol overhead, needed for the stream for each individual source (each SSRC) in an RTP session, as specified by a token bucket at the stream sender. An end-point may use the SLTR semantics to request to establish a least required bandwidth. An end-point using the SLTR semantics may set the token bucket rate and/or the token bucket size to "*" to indicate that the end-point has no preference, but that it expects some network node or the answering end-point to define the value(s). A network node answering to the SLTR SHALL replace this with the SLT semantics to indicate the least required bandwidth it sees necessary and which it has attempted to guarantee. If the request is for certain specified payload types, a network node that cannot grant bandwidth based on payload types MAY replace those requested payload types with "*" in the SLT response to indicate a payload type agnostic grant. An end-point receiving an SDP with SLTR, i.e. where the network has not replaced the SLTR semantics with any SLT semantics, SHOULD NOT assume that the requested bandwidth is guaranteed.

ALT (Aggregated Least required Token bucket): The least required bandwidth, including protocol overhead, needed for the sum of all sources (all SSRCs) in an RTP session as specified by a token bucket at the stream sender. When using the ALT semantic the AMT semantic SHOULD also be specified for the same direction and scope. The directionality and payload type considerations for ALT are the same as for AMT. If the ALT semantics is not defined then this means that the least required bandwidth is zero.

ALTR (Aggregated Least required Token bucket Request): The request for establishing a least required bandwidth, including protocol overhead, needed for the sum of all sources (all SSRCs) in an RTP session as specified by a token bucket at the media sender side. The directionality and payload type considerations for ALTR are the same as for SLTR. The ALTR semantics MUST only be used together with AMT.

The SMT and AMT semantics, with or without SLT and ALT respectively, may be used both symmetrically and in a particular direction. They can be used either to express the maximum (and minimum) for one particular payload type, for a set of payload types or for any payload type (pt=*).

The required prefix ("!") is used when the direction, scope and semantics is required be supported and understood by the SDP consuming end-point.

4.2. Declarative Use

In declarative usage the SDP attribute is interpreted from the perspective of the end-point being configured by the particular SDP. An interpreter MAY ignore 'a=bw' attribute lines that contains unknown scope or semantics that does not start with the required ("!") prefix. If a "required" prefix is present at an unknown scope or semantics, the interpreter SHALL NOT use this SDP to configure the end-point.

4.3. Usage in Offer/Answer

The offer/answer negotiation is performed for each 'a=bw' attribute line individually with the scope and semantics immutable.

An offerer may use the 'a=bw' attribute(s) for some or all of the offered media types. An answerer may remove the 'a=bw' attribute(s) for the media types where it was used in the SDP offer.

The SDP may include an offer for an Aggregated Maximum Token bucket (AMT) without specifying any Stream Token Buckets (SMTs) for any

individual streams.

When using the 'a=bw' attribute to define the token bucket for a certain scope then the offerer should define token buckets for all scopes of the same type. For example, if the SDP offer includes three payload types, e.g. 96, 97 and 98, and if a token bucket is defined for payload type 96, then the offerer should also define token buckets for the other payload types. This can be done either by defining one token bucket each for payload type 97 and 98 or by defining a common token bucket for payload type 97 and 98.

When the token bucket rate and size are declared in an offer for directionality 'sendrecv' then this indicates the token bucket rate and the token bucket sizes are the same in both directions. For example, if the offered bandwidth is 1 Mbps, then the end-point declares that it is capable of sending with a bandwidth up to 1 Mbps and that it is capable of receiving with a bandwidth up to 1 Mbps.

If either the token bucket rate(s) or the token bucket sizes are different in sending and receiving direction then 'sendrecv' cannot be used. One should instead include two or more 'a=bw' lines with the respective directionality, bandwidths and sizes.

When the token bucket parameters are declared in an SDP offer for directionality 'send' then this indicates the token bucket parameters the sender intends to use. The answerer may change this value, both to increase it and to reduce it, see below.

When the token bucket parameters are declared in an SDP offer for directionality 'recv' then this indicates that the largest envelope for the token bucket parameters that the offerer thinks the media sender shall use.

An agent understanding the 'a=bw' attribute and answering to an offer including the 'a=bw' attribute SHOULD include the attribute in the answer for all media types for which it was offered.

An answerer SHOULD ignore 'a=bw' attribute lines that contains unknown scope or semantics that does not contain the required ("!") prefix. If a "required" prefix is present at an unknown scope or semantics, then the answerer SHALL reject the media description by setting the port to 0 and copy the 'a=bw' attributes not understood in the answer. In this case, 'a=bw' attributes that are understood SHALL NOT be included in the answer.

If an answerer would like to add additional bandwidth configurations using other directionality, scope, and semantics combination, then it MAY do so by adding such definitions in the SDP answer.

An agent may also divide an 'a=bw' offer into several 'a=bw' offers. One example is when the SDP offer included an 'a=bw' offer with directionality 'sendrecv', which indicates that the token bucket parameters are the same in sending and receiving direction. If the answerer would like to change the parameters for one or both directions, so that the parameters are no longer the same for both directions, then the answerer can include two 'a=bw' lines in the SDP answer, one for sending direction and another for receiving direction. In case an offered sendrecv media becomes a single direction media then the sendrecv can be modified to that single direction.

An agent responding to an offer will need to consider the directionality and reverse them in the answer when responding to media streams using unicast.

For media stream offers over unicast with directionality send, the answerer SHALL reverse the directionality and indicate its reception bandwidth capability, which may be lower or higher than what the sender has indicated as its intended maximum.

For media stream offers over unicast with directionality receive, the token bucket parameters indicate the upper limits. The answerer SHALL reverse the directionality and may reduce the bandwidth when producing the answer indicating the answerer intended maximum transmission rate.

If the answerer removes one or several RTP Payload Types from the SDP when creating the SDP answer then the corresponding 'a=bw' lines SHOULD be removed as well. The answerer MAY however keep an 'a=bw' line when the removed RTP Payload Type number is included within an identified range or list of Payload Type numbers.

4.4. Bucket Size Estimation

In SDP bandwidth terms, the bucket size is a new parameter and what value to use for it may be hard to understand for implementers of this specification. This section therefore gives some guidelines on how to set bucket size values.

A token bucket specifies an envelope for a transmission profile where individual measurements have some impact if the media stream or aggregate should be considered within the specified profile. The semantics defined in this document only require that the media stream is within the token bucket specification at the point emitting it into the network. The network may add jitter causing the media stream/aggregate to no longer be within the specified token bucket profile.

4.4.1. Sender Specified Token Bucket

A sender SHOULD base the choice of token bucket size on how it plans to send data. That can in turn be decided from e.g. codec configuration, intended number of encoded frames per packet (ptime), network interface, maximum transmission unit (MTU), etc. In practice, for the simplified case where the sender is designed to send all packets with precisely even time spacing, the token bucket size can be set to the maximum packet size and the bit-rate to the long term highest bit-rate intended to be used.

However, for media streams that are more variable the bucket parameters should be chosen so that the emitted traffic is not too bursty measured over a shorter interval. Until the bucket is drained, the media sender will be able to emit packets at or close to the interface's maximum bit-rate. Long burst of packets at interface speed becomes more sensitive to loss due to cross-traffic in switching fabrics with small buffers. Due to this, a sender can consider transmission scheduling to a rate lower than the interface rate but higher than the token bucket average rate.

Let's consider the example of a large video intra frame consisting of 10 full MTU (let's assume 1500 bytes) packets which is 5 times the size of the median frame size of two full MTU packets. The average bit-rate may be 1 Mbps. If the token bucket was to be configured to (1 Mbps, 1500) then that would imply that a new full MTU packet could be emitted no more often than one packet every 12 ms. That would require 120 ms to transmit the intra frame, which for a 25 frames per second video is 3 frame intervals. Thus potentially inducing significant playout jitter at a receiver. A token buffer specification of (1 Mbps, 15000) would allow all 10 packets be sent up to line speed. This could result in them being emitted every 1.2 ms over a 100 Mbps interface if there is no competing traffic. To ensure that a 10 packet burst should be possible to transmit within one frame interval of 40 ms, then the bucket depth needed is burst size in bits, minus time interval times bucket fill rate, and the resulting value converted back into bytes: $(15000 * 8 - 0.04 * 1M) / 8 = 10000$ bytes. The average bit-rate for this intra frame over a single frame period becomes 4 Mbps. So the question is if bursts up to 4 Mbps should be allowed now and then as long as the average is within 1 Mbps, or if the sender has to transmit the intra using several frame intervals, skipping the next frame(s) and hoping that the receiver doesn't drop the intra frame as being too late. The sender could also consider reducing the quality of the intra frame, resulting in a reduced number of MTU required to transmit it.

A sender SHOULD avoid adding excessive safety margins to the sending bucket size. A sender MAY add bucket size margins if it has

knowledge of internal transmission timing variations, or if it knows about packet handling outside the sender itself that will affect the effective bucket size (as seen from a receiver) that is otherwise not reflected in the conveyed bucket size figure.

4.4.2. Receiver Specified Token Bucket

With the semantics specified in this document, the intended media receiver gets to provide token bucket parameters that specifies how the sender should behave. The traffic received by the receiver (or intermediate nodes) may no longer conform to the token bucket due to jitter introduced by the network path between the sender and the receiver. This document assumes that the receiver will have receiver buffers for de-jittering that are significantly larger than the token bucket parameters. This due to that a media unit like a video frame may be transmitted over time using more data than the bucket depth provides and instead spread it in time, transmitting each fragment when the bucket is refilled enough for the next fragment to be sent.

A receiver's input to the sender's bit-rate limitation should be based on known limitations such as the networks, decoding capabilities etc. The bucket depth will control how bursty the traffic can be beyond the long term average specified by the bucket refill rate.

4.4.3. Bucket Adjustment in Middle Nodes

When there are media aware middle nodes on the media path between the sender and receiver, those middle nodes may have to or want to apply similar considerations as the original media sender and receiver. If those middle nodes are aware of SDP and the new bandwidth attribute from this specification, and have in-path SDP adjustment capabilities, they could benefit from modifying the values to better fit the actually available end-to-end media path capabilities. For example, an RTP Media Translator can express what it actually is going to deliver of the far end-point's media to an end-point instead of that far end-point's provided values.

4.4.4. Network Policing

As the token bucket specified for the semantics in this document is based on what the sender emit into the network, a policer should have some margin allowing for network introduced jitter. The amount will of course be dependent on the policer's location in relation to the media sender.

4.4.5. Utilizing Network Feedback

If the media uses RTP and when the media has been transmitted for some time, the sender should have received a fair amount of RTCP receiver reports from the receiver. The sender can from RTCP estimate the observed network jitter at the receiver and may be able to dynamically adjust the sender behavior such that the aggregate of the sender behavior and the reported network jitter are fulfilling the senders token bucket profile.

4.5. SDP Examples for Point-to-point Sessions

These SDP examples show how the new bandwidth attribute can be used. The benefits, compared to the legacy bandwidth attribute, are also highlighted.

The SDP examples included below are intentionally not complete. Only the parts that are relevant for this description are included.

4.5.1. Symmetric Fixed-rate Codecs

This example shows the SDP offer for several fixed-rate codecs, mu-law and A-law PCM, G.726 and G.728.

```
m=audio 49200 RTP/AVP 8 0 96 18
b=AS:80
a=rtpmap:96 G726-32/8000/1
a=bw:sendrecv pt=0,8 SMT:tb=80000:1000
a=bw:sendrecv pt=96 SMT:tb=48000:1000
a=bw:sendrecv pt=18 SMT:tb=24000:1000
a=ptime:20
a=maxptime:20
```

SDP offer for mu-law and A-law PCM and IPv4

The new bandwidth attribute offers the possibility to negotiate the bandwidth individually for each codec. If the answerer removes a codec when creating the answer then it is still known how much bandwidth the other codecs will use. This means that the ambiguities listed in Section 3.2.1 can be avoided.

4.5.2. Symmetric Rate-Adaptive Codec

This example shows the SDP negotiation for offering using the AMR codec, AMR [AMR].

```
m=audio 49200 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=bw:sendrecv pt=97 SMT:tb=28800:200
a=bw:sendrecv pt=97 SLTR:tb=22400:200
a=ptime:20
a=maxptime:100
```

SDP offer from end-point A for AMR and IPv4

```
m=audio 49100 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=bw:sendrecv pt=97 SMT:tb=28800:200
a=bw:sendrecv pt=97 SLT:tb=22400:200
a=ptime:20
a=maxptime:100
```

SDP answer from end-point B also for AMR and IPv4

Since the new bandwidth attribute offers a possibility to negotiate both the maximum and the at least required bandwidth, it is possible for both the other end-point and any resource allocation function to know how the end-points will adapt when congestion is detected.

4.5.3. Symmetric Several Rate-Adaptive Codecs

This example shows how the new bandwidth attribute, 'a=bw', can be used to negotiate the maximum and the least required bandwidths for multiple rate-adaptive codecs, in this case for AMR and AMR-WB, AMR-WB [AMR-WB]. For AMR, the highest codec mode is 12.2 kbps, giving a maximum bandwidth of 28.8 kbps, and the at least required mode is selected to be 5.9 kbps, giving a least required bandwidth of 22.4 kbps. For AMR-WB, the highest codec mode is 23.85 kbps, giving a maximum bandwidth of 40.4 kbps, and the least required mode is 8.85 kbps, giving a least required bandwidth of 25.6 kbps.


```
m=audio 49200 RTP/AVP 96 97
b=AS:41
a=rtpmap:96 AMR-WB/16000/1
a=fmtp:96 mode-change-capability=2; max-red=80
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=bw:sendrecv pt=96 SMT:tb=40400: 350
a=bw:sendrecv pt=96 SLTR:tb=25600:350
a=bw:sendrecv pt=97 SMT:tb=28800:200
a=bw:sendrecv pt=97 SLTR:tb=22400:200
a=ptime:20
a=maxptime:100
```

SDP offer from end-point A for AMR-WB, AMR and IPv4

```
m=audio 49100 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=bw:sendrecv pt=97 SMT:tb=28800:200
a=bw:sendrecv pt=97 SLT:tb=22400:200
a=ptime:20
a=maxptime:100
```

SDP answer from end-point B for AMR and IPv4 (AMR-WB is removed)

In this case, it is clear when the answer is received that the bandwidth needed for AMR applies to both directions. There is no need for a send offer/answer negotiation to clarify that the bandwidth applies also to end-point A's receiving direction. Thereby, the issues listed in Section 3.2.3 are resolved.

4.5.4. Asymmetric Session

The following SDP example shows how to use the new bandwidth attribute to offer asymmetric streams. In this case, the end-point offers to send H.264 video with 1 Mbps while it is capable of receiving H.264 with up to 3 Mbps. Note that this example does not make use of the codec-specific H.264 level asymmetry signaling as defined in RFC 6184 [RFC6184].

```
m=video 50324 RTP/AVP 96
b=AS:3000
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c016
a=bw:send pt=96 SMT:tb=1000000:8192
a=bw:recv pt=96 SMT:tb=3000000:16384
```

SDP offer with asymmetric video bandwidth

It should be clear from this example that the new bandwidth attribute is useful when negotiating asymmetric sessions since it offers the possibility to define the token bucket parameters for both sending and receiving directions separately.

4.5.5. Session with Retransmission

This SDP example shows how the new bandwidth attribute, 'a=bw', can be used for negotiating the bandwidth when the RTP Retransmission Payload Format RFC 4588 [RFC4588] is used.

```
m=video 49170 RTP/AVPF 96 97
b=AS:500
a=rtpmap:96 MP4V-ES/90000
a=rtcp-fb:96 nack
a=fmtp:96 profile-level-id=8; config=01010000012000884006682C2090A21F
a=rtpmap:97 rtx/90000
a=fmtp:97 apt=96;rtx-time=3000
a=bw:send pt=* AMT:tb=500000:4096
a=bw:recv pt=* AMT:tb=500000:8192
```

SDP offer with aggregate bandwidth and RTP retransmission

In this case, it is beneficial to use the Aggregate Maximum Token bucket semantics to allow the end-points to adapt the bandwidths used for the original stream and for the retransmission stream during the session. The end-point can send more original packets when the packet loss rate is low. When the packet loss rate is high then the end-point can use less bandwidth for the original packets and instead allow for more retransmissions. It would also be possible to specify separate limits for the original stream and the retransmission stream by using a separate set of 'a=bw'-lines for pt=96 and pt=97.

4.6. SDP Examples with Sessions with Multiple Streams

4.6.1. Multiple Streams

The example below is based on the use case described in Section 3.3.1. Only the negotiation for video is shown here.

```
m=video 49300 RTP/AVP 96
b=AS:3000
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c01f
a=bw:send pt=* SMT:tb=1000000:1000
a=bw:recv pt=* SMT:tb=1000000:2000
a=bw:send pt=* AMT:tb=1000000:1000
a=bw:recv pt=* AMT:tb=3000000:6000
a=max-recv-ssrc:* 4
```

SDP offer with both per-stream and aggregate bandwidth

With the new bandwidth attribute, it is possible to define the bandwidth for each received stream independently from each other. In this case, the SDP shows that the end-point is prepared to send maximum 1 Mbps, and that the end-point is prepared to receive maximum 1 Mbps per stream. The SDP also shows that the end-point is prepared to receive maximum 3 Mbps, aggregated for the up to four streams in the receiving direction. Note that this implies that to receive more than three streams, each stream's bandwidth must be reduced to comply with the maximum aggregate.

4.6.2. Declarative Example with Stream Asymmetry

This example shows a declarative usage of the new bandwidth attribute.

```
m=video 50324 RTP/AVP 96 97 98
a=rtpmap:96 H264/90000
a=rtpmap:97 H263-2000/90000
a=rtpmap:98 MP4V-ES/90000
a=max-recv-ssrc:96 2
a=max-recv-ssrc:* 5
a=bw:send pt=* SMT:tb=1200000:16384
a=bw:recv pt=96 SMT:tb=1500000:16384
a=bw:recv pt=97,98 SMT:tb=2500000:16384
a=bw:recv pt=* AMT:tb=8000000:65535
```

SDP offer with payload-specific per-stream bandwidth

In the above example, the outgoing single stream is limited to bucket rate of 1.2 Mbps and bucket size of 16384 bytes. The up to 5 incoming streams can in total use maximum 8 Mbps bucket rate and with a bucket size of 65535 bytes. However, the individual streams maximum rate is depending on payload type. Payload type 96 (H.264) is limited to 1.5 Mbps with a bucket size of 16384 bytes, while the Payload types 97 (H.263) and 98 (MPEG-4) may use up to 2.5 Mbps with a bucket size of 16384 bytes.

4.7. Interoperability Issues

The proposed new bandwidth attribute obviously has connections to the bandwidth modifier 'b=AS' and the attributes defined for directionality ('a=sendrecv', 'a=sendonly', 'a=recvonly' and 'a=inactive') defined in RFC 4566 [RFC4566]. It is therefore important to properly analyze these relationships so that any interoperability issues can be avoided.

4.7.1. Interoperability with Existing Bandwidth Attribute

If the SDP includes both the 'b=AS' bandwidth modifier and 'a=bw' bandwidth attribute then alignment may be necessary to avoid confusion. This section gives some guidelines for such alignment. It may however happen that some usage needs other alignments than what is discussed below. If so, then those alignments need to be considered on a case-by-case. The discussion below should therefore not be seen as an exhaustive list.

In general, the bandwidths offered with 'b=AS' and 'a=bw' should be aligned for the direction that applies for the 'b=AS' bandwidth modifier. For 'sendrecv' and 'recvonly' sessions, 'b=AS' indicates the bandwidth for the receiving direction. The b=AS is closest in interpretation to the AMT semantic. If the stream maximum semantic (SMT) is used then the sum of the bandwidths in the receive direction may exceed the 'b=AS' bandwidth but the AMT should not exceed the b=AS value.

If the session includes multiple streams, but if not all of the streams will be active simultaneously, then 'b=AS' should indicate the maximum bandwidth that will be used for the combinations of streams that are active simultaneously, the same way AMT could be used in such a session. This also means that the bandwidths offered with 'a=bw' are accumulated for the combination of streams that are active, and this aggregated bandwidth should not exceed the bandwidth defined with 'b=AS'. Note however that it is possible and feasible to specify an aggregate that is less than the sum of the maximum bandwidth for the maximum amount of available streams. It may be possible to use the maximum number of active streams with a lower bandwidth than the maximum, or it may be possible to reduce the active number of streams to stay within the bandwidth limit.

The SDP below gives an example of how this is done. In this example, the intention is to use either the payload type pair (96, 97) or the payload type pair (98, 99). The intention is however to, for example, not pair payload types 96 and 98.

```
m=video 50000 RTP/AVP 96 97 98 99 100
b=AS:1000
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c00d
a=rtpmap:97 H264/90000
a=fmtp:97 profile-level-id=42c00c
a=rtpmap:98 H264/90000
a=fmtp:98 profile-level-id=42c00d
a=rtpmap:99 H264/90000
a=fmtp:99 profile-level-id=42c00c
a=rtpmap:100 H264/90000
a=fmtp:100 profile-level-id=42c00c
a=bw:sendrecv 96 SMT:tb=700000:4000
a=bw:recv 97 SMT:tb=300000:3000
a=bw:sendrecv 98 SMT:tb=500000:3000
a=bw:recv 99 SMT:tb=200000:2000
a=bw:send 100 SMT:tb=300000:1400
a=sendrecv
```

SDP offer with complex bandwidth relations

This session is bi-directional, as shown with the 'a=sendrecv' attribute. The bandwidth offered with 'b=AS' therefore applies to the receive direction. The 'b=AS' is then set based on the combination of streams that gives the highest bandwidth, i.e. the payload type pair (96, 97).

This means that the bandwidths offered with 'a=bw' are aligned with the bandwidth offered with 'b=AS'.

If, on the other hand, the intention would be to use another combination of payload types, for example (96, 98), then this would add up to 1200 kbps, which would mean that the stream bandwidths would not be aligned with the 'b=AS' bandwidth.

This shows that bandwidths for 'sendrecv' and 'recv' directions are added together when determining the bandwidth for the combined streams.

If the offer is "complex", for example offering multiple streams for both speech and video, possibly with many different codecs, (and therefore uses 'a=bw' together with the 'b=AS' bandwidth modifier) and if the answerer wants to change this into a "simple" session (e.g. plain simple VoIP with only one RTP payload type for codec X) then the answerer may remove the 'a=bw' lines when creating the answer. It may therefore happen that the answer includes only 'b=AS' bandwidth modifier in the SDP answer. However, if the offer does not include any 'b=AS' line then it is recommended to maintain the 'a=bw'

lines also in the answer, even for "simple" sessions. This means that the offerer cannot rely on the existence of 'a=bw' in the answer.

4.7.2. Interoperability with Existing Directional Attribute

Since the 'a=bw' attribute includes a parameter for directionality it is important to clarify the relationship to the already existing directional attributes in SDP ('sendrecv', 'sendonly', 'recvonly' and 'inactive'). In general, one can say that:

- o The SDP attribute indicates the directionality for the session.
- o The 'a=bw' attribute defines the directionality for the bandwidth for streams within the session.
- o The SDP attribute for directionality has precedence over the 'a=bw' parameter for directionality when it comes to the media that is actually being transmitted.

At session setup time, it is therefore acceptable to define streams with other directionality than what is shown with the SDP attribute for directionality. However, when media is transmitted, then the SDP attribute for directionality has to be followed. An example of this is shown below.

```
m=video 5000 RTP/AVP 96 97 98
b=AS:1000
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c00d
a=rtpmap:97 H264/90000
a=fmtp:97 profile-level-id=42c00c
a=bw:sendrecv 96 SMT:tb=700000:4000
a=bw:recv 97 SMT:tb=200000:3000
a=bw:send 97 SMT:tb=300000:1400
a=recvonly
```

SDP offer specifying bandwidth in 'inactive' direction

This means that three bandwidths are defined at session setup:

- o one stream (PT=96) for 700 kbps bi-directional video;
- o one stream (PT=97) for 200 kbps receive-only video; and:
- o one stream (PT=97) for 300 kbps send-only video.

However, since 'a=recvonly' is defined then this means that the endpoint is, at the session setup time, only willing to receive media

even though the SDP contains bandwidth declarations also for the sending direction. This allows for setting up streams that are effectively inactive in one or both directions from the beginning of the session and then enabling them later in the session.

This can be compared with the case when one defines one or more codecs, even if the session starts up as 'inactive'.

5. Rules and Recommendations for Extensions

The a=bw attribute is defined to be extensible and this section discusses the extension points that are available.

5.1. Directionality

The current specification defines send, recv and sendrecv. In case some new directionality behavior is needed that doesn't match the existing, a new one could be defined. This should be avoided unless a clear need for a new directionality is found.

5.2. Scope

It is expected that there will be a need to extend the bandwidth scope. This document only defines two scope types, session and payload type, and there is very likely other desirable scopes that will be defined in the future. Possible examples of scopes are those applying to a specific SSRC, a particular end-point, or a class of end-points.

5.3. Semantics

This is the extension point that is expected to be frequently used in the future. A major proliferation of semantics is not good for interoperability, but it is likely that bandwidth shortcomings or missing functionalities will be discovered in the future. Thus defining new semantics gives maximum flexibility to define the meaning of the provided value(s), the format of the values and how to interpret the directionality and scope values.

5.4. Values

This document only defines token buckets as values. In case fewer or more parameters are needed to express a particular semantics, new value formats can be defined. Defining new value formats should be done with some consideration of generality and reuse so that future semantics can also use the new value format, with the target to try to minimize the number of different formats.

6. Open Issues

This document contain a few open issues:

1. Multicast behavior needs to be specified.
2. It is an open question to decide if and how to handle the RTCP bandwidth negotiation, e.g. corresponding to b=RS and b=RR.
3. It is an open question to develop semantics for the transport independent bandwidth negotiation, e.g. corresponding to b=TIAS.
4. It is an open question what rules and recommendations there should be for extensions to this memo.

7. IANA Considerations

Following the guidelines in RFC 4566 [RFC4566] and in RFC 3550 [RFC3550], the IANA is requested to register:

1. The bw attribute as defined in Section 4.1.
2. The bw attribute directionality registry rules
3. The bw attribute scope registry rules.
4. The bw attribute semantics registry rules.
5. The bw attribute values registry rules.

This section will be filled out in future versions of this document.

8. Security Considerations

Excessive bandwidth allocation can consume all the resources, much more than what the end-point(s) intend to use. So, if a session allocates an unnecessarily high bandwidth then this will likely mean that some other users cannot be admitted, or that they cannot get QoS guaranteed resources that they requested and have to use best effort. It can also happen that the session itself is rejected, if the end-points try to allocate resources that are not available. Allocating too little bandwidth is likely to negatively impact the perceived media quality or entirely prevent reception of requested media.

The above shows that the bandwidth attribute is a potential vector for attacks both from malicious end-points or third party attackers

that attempts to modify the attribute to impact the system to allocate unnecessary resources, deny end-points service, reduce quality for end-points or incur cost on users.

To prevent third party attacks the signalling should be source authenticated and integrity protected to prevent any on or off-path attacker from injecting or modifying the SDP. Malicious end-points can't as easily be protected against using crypto, instead behavior analysis and preventing such a malicious end-point from having serious impact on other end-points are needed.

9. Acknowledgements

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2212] Shenker, S., Partridge, C., and R. Guerin, "Specification of Guaranteed Quality of Service", RFC 2212, September 1997.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.
- [RFC3556] Casner, S., "Session Description Protocol (SDP) Bandwidth Modifiers for RTP Control Protocol (RTCP) Bandwidth", RFC 3556, July 2003.
- [RFC3890] Westerlund, M., "A Transport Independent Bandwidth Modifier for the Session Description Protocol (SDP)", RFC 3890, September 2004.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

10.2. Informative References

- [AMR] "3GPP TS 26.090, "Adaptive Multi-Rate (AMR) speech codec; Transcoding functions".", June 1999.
- [AMR-WB] "3GPP TS 26.190, "Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Transcoding functions".", April 2001.
- [G.711] "ITU-T Recommendation G.711, "Pulse Code Modulation (PCM) of Voice Frequencies".", November 1988.
- [G.726] "ITU-T Recommendation G.726, "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)".", December 1990.
- [G.729] "ITU-T Recommendation G.729, "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)".", March 1996.
- [H.264] "ITU-T Recommendation H.264, "Advanced video coding for generic audiovisual services".", May 2003.
- [I-D.westerlund-avtcore-max-ssrc]
Westerlund, M., Burman, B., and F. Jansson, "Multiple Synchronization sources (SSRC) in RTP Session Signaling", draft-westerlund-avtcore-max-ssrc-00 (work in progress), October 2011.
- [I-D.westerlund-avtcore-multiplex-architecture]
Westerlund, M., Burman, B., and C. Perkins, "RTP Multiplexing Architecture", draft-westerlund-avtcore-multiplex-architecture-00 (work in progress), October 2011.
- [RFC4588] Rey, J., Leon, D., Miyazaki, A., Varsa, V., and R. Hakenberg, "RTP Retransmission Payload Format", RFC 4588, July 2006.
- [RFC6184] Wang, Y., Even, R., Kristensen, T., and R. Jesup, "RTP Payload Format for H.264 Video", RFC 6184, May 2011.

Authors' Addresses

Tomas Frankkila
Ericsson
Laboratoriegrand 11
SE-971 28 Lulea
Sweden

Phone: +46 10 714 30 20
Email: tomas.frankkila@ericsson.com

Magnus Westerlund
Ericsson
Farogatan 6
SE-164 80 Kista
Sweden

Phone: +46 10 714 82 87
Email: magnus.westerlund@ericsson.com

Bo Burman
Ericsson
Farogatan 6
SE-164 80 Kista
Sweden

Phone: +46 10 714 13 11
Email: bo.burman@ericsson.com

MMUSIC Working Group
Internet-Draft
Intended status: Standards Track
Expires: January 17, 2013

T. Frankkila
M. Westerlund
B. Burman
Ericsson
July 16, 2012

Extensible Bandwidth Attribute for SDP
draft-westerlund-mmusic-sdp-bw-attribute-02

Abstract

Knowledge of what bandwidths the end-points intend to use is important both for the other end-point and for resource allocation in various types of networks. This is especially important for wireless access networks which typically have quite limited resources. The bandwidth attribute in Session Description Protocol (SDP), 'b=AS', is today quite widely used to define the bandwidth that the end-points intends to use, in various types of sessions. This document will show that the existing bandwidth attribute, such as 'b=AS', although widely used in today's scenarios, has limitations that make it hard or even impossible for the end-points to express their intentions accurately when it comes to bandwidth usage. To solve the identified problems, this document defines a new extensible SDP bandwidth attribute 'a=bw' which enables more detailed control over the bandwidth declarations, request, and allocations. With the new bandwidth attribute it is possible to define different scopes in the session setup and then negotiate the bandwidth individually for each scope.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on January 17, 2013.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction	4
2. Definitions	4
2.1. Requirements Language	4
2.2. Terminology	5
3. Use Cases and Design Rationale	5
3.1. Existing Bandwidth Attribute	6
3.1.1. Attribute Definition	6
3.1.2. Offer/answer Procedure for the Existing Bandwidth Attribute	7
3.1.3. End-point Behavior when Generating Traffic	7
3.2. Point-to-point Sessions using SDP offer/answer	8
3.2.1. Symmetric Point-to-point Sessions, Fixed-rate Codecs	8
3.2.2. Symmetric Point-to-Point Sessions with Rate-Adaptive Codec	10
3.2.3. Symmetric Point-to-Point Sessions with Several Rate-Adaptive Codecs	11
3.2.4. Asymmetric Point-to-Point Sessions	13
3.3. Sessions with Multiple Streams	14
3.3.1. Multiple Streams	15
3.4. User Experience and Bandwidth Negotiation	15
3.5. Summary of Findings	16
4. Attribute Specification	18
4.1. SDP Grammar	18
4.2. Declarative Use	23
4.3. Usage in Offer/Answer	23
4.4. Bucket Size Estimation	25
4.4.1. Sender Specified Token Bucket	25
4.4.2. Receiver Specified Token Bucket	27
4.4.3. Bucket Adjustment in Middle Nodes	27
4.4.4. Network Policing	27
4.4.5. Utilizing Network Feedback	27

4.5.	SDP Examples for Point-to-point Sessions	28
4.5.1.	Symmetric Fixed-rate Codecs	28
4.5.2.	Symmetric Rate-Adaptive Codec	28
4.5.3.	Several Symmetric Rate-Adaptive Codecs	29
4.5.4.	Asymmetric Session	30
4.5.5.	Session with Retransmission	31
4.6.	SDP Examples with Sessions with Multiple Streams	31
4.6.1.	Multiple Streams	32
4.6.2.	Declarative Example with Stream Asymmetry	32
4.7.	Interoperability Issues	33
4.7.1.	Interoperability with Existing Bandwidth Attribute	33
4.7.2.	Interoperability with Existing Directional Attribute	35
5.	Rules and Recommendations for Extensions	36
5.1.	Directionality	36
5.2.	Scope	36
5.3.	Semantics	36
5.4.	Values	37
6.	Open Issues	37
7.	IANA Considerations	37
8.	Security Considerations	37
9.	Acknowledgements	38
10.	References	38
10.1.	Normative References	38
10.2.	Informative References	39
	Authors' Addresses	40

1. Introduction

This document looks at the issues of both basic and non-basic usage of RTP [RFC3550] and analyzes how well the existing SDP [RFC4566] attribute 'b=AS' for bandwidth negotiation performs in different scenarios.

This analysis is done by defining a number of use cases, containing sessions with:

- o single and multiple media types;
- o symmetric and asymmetric media streams;
- o single and multiple media sources, including multiple sources from the same end-point;
- o multiple end-points each having one or more media sources, including applications that use multiple encodings of a particular media.

It is shown that the existing bandwidth attributes 'b=AS' [RFC4566] and 'b=TIAS' [RFC3890] has limitations which make it unclear or even impossible for end-points and for resource allocation functions in the network to determine how much bandwidth the service will use. The analysis also provides the design rationale for the new bandwidth attribute.

This document then proposes a general and extensible mechanism for bandwidth negotiation that can be used for any type of session. Interoperability with the existing mechanisms for bandwidth negotiation is especially important since the existing bandwidth attribute has a wide-spread usage.

This document also presents several examples for how the new bandwidth attribute can be used in the session setup phase for various types of sessions. The examples are derived for IP/UDP/RTP transport although nothing should prevent using the new bandwidth attribute also for other transport protocols.

2. Definitions

2.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2.2. Terminology

The following terms and abbreviations are used in this document:

Bandwidth: In this document, the bandwidth is defined as the IP level bandwidth, i.e. including the network protocol (IPv4 or IPv6) and transport protocol (TCP, UDP, RTP, etc) overhead. When RTP is used then the RTCP bandwidth is handled separately from the bandwidth used for RTP packets. Bandwidth in this context is in the unit bits per second, not Hz.

Encoding: A particular encoding is the choice of the media encoder (codec) that has been used to compress the media. Different encodings result in the fidelity of that encoding through the choice of sampling, bit-rate and other configuration parameters.

End-point: A single entity sending and/or receiving RTP packets. It may be decomposed into several functional blocks, but as long as it behaves a single RTP stack entity it is classified as a single end-point.

Media stream: A sequence of RTP packets using a single SSRC that together carry part or all of the content of a specific Media Type from a specific sender source within a given RTP session.

RTP session: An RTP session consists of one or more media streams that have the same purpose. The typical example is to have one RTP session per media type, i.e. that voice and video use different RTP sessions (different ports) since they have different purpose. It is however possible to have multiple streams in an RTP session, for example when having both a stream for non-redundant audio and another stream for re-transmissions of audio packets. The fundamental definition of an RTP session is a single SSRC space.

3. Use Cases and Design Rationale

This section describes a number of use cases where the existing bandwidth attribute 'b=AS' is used for bandwidth definition. It also discusses why the limitations of the existing bandwidth attribute makes it hard for other end-points and resource allocation functions to know or estimate how much bandwidth that will be used in the ongoing session.

The analysis is made by defining a set of use cases. The first use cases include fairly simple session types, i.e. point-to-point

sessions with or without asymmetry. A few more complex use cases are then analyzed. The last set up use cases reflect fairly advanced session types, e.g. various variants of multiplexing and usage of multiple media streams.

The discussion is then summarized and the design rationales for the new bandwidth attribute are outlined.

3.1. Existing Bandwidth Attribute

The existing bandwidth modifier 'b=' defined in RFC 4566 [RFC4566] is reviewed in this section.

3.1.1. Attribute Definition

The existing bandwidth attribute 'b=' is defined in Section 5.8 of RFC 4566 [RFC4566]. The syntax is:

`b=<bwtype>:<bandwidth>`

where:

`<bwtype>` is either:

'AS' ("Application Specific"), which is the maximum bandwidth as estimated by the application; or:

'CT' ("Conference Total"), which is the total bandwidth for all media at all sites.

`<bandwidth>` is the bandwidth value in kilobits per second.

Bandwidth types have been defined for the negotiation of the RTCP bandwidth using 'b=RS' and 'b=RR', RFC 3556 [RFC3556].

There is also a bandwidth type for negotiating the transport independent application specific maximum bandwidth, 'b=TIAS', RFC 3890 [RFC3890]. This bandwidth type is similar to the 'b=AS' bandwidth type, except that the overhead caused by the transport protocol headers is not included.

One issue with the existing bandwidth attribute is that the syntax is very limited since it only allows for defining new bandwidth types (`<bwtype>`) and their respective single numerical value. This limitation needs to be considered in the discussion below.

3.1.2. Offer/answer Procedure for the Existing Bandwidth Attribute

"An Offer/Answer Model with the Session Description Protocol (SDP)" [RFC3264] describes the offer/answer procedures for the existing bandwidth attribute. For the SDP offer, for sendrecv and recvonly streams, it describes that the bandwidth attribute indicates the desired bandwidth that the offerer would like to receive. For the SDP answer, for sendrecv and recvonly streams, it describes that the bandwidth attribute indicates the bandwidth that the answerer would like the offerer to use when sending media. Thus, for offer/answer negotiations, the bandwidth attribute indicates the bandwidth for the receive direction of each end-point.

The solution presented in this document focuses primarily on clarifying and assisting the Application Specific (AS) bandwidth.

[It is an open question to decide if and how to handle the RTCP bandwidth negotiation, e.g. corresponding to b=RS and b=RR.]

[It is an open question to develop semantics for the transport independent bandwidth negotiation, e.g. corresponding to b=TIAS.]

3.1.3. End-point Behavior when Generating Traffic

When an end-point is sending media then this can be done in many different ways, depending on the choices the implementers have made.

Some end-points may send their data in a fairly "nice and smooth" media stream, which means that both the packet sizes and the packet rates are more or less constant all the time. An example of a smooth stream is when the end-point is encoding speech and is sending one packet every 20 ms and when the packets are of equal size.

Other end-points may generate bursty streams, which have a large peak-to-average ratio. An example of a bursty stream is when an end-point is encoding video. Most of the time, the end-point is sending packets with almost the same size and with constant packet rate. However, it happens occasionally that the encoder generates much more data for a frame, which may give a very large packet size. It may even happen that the sender has to segment the data into several packets, which may be transmitted in a burst, thereby causing a very high peak rate.

Whether the stream is smooth or bursty makes a big difference for the network and the policy control that usually applies in QoS controlled networks. If the stream is too bursty, then a policy control function may decide to drop packets that exceed the granted rate. This will lead to degraded quality and reduced user satisfaction.

The existing bandwidth attribute offers no mechanism to negotiate what temporal variations that can be allowed for a stream. The only available mechanism is to negotiate the maximum bandwidth, but there is nothing that defines any kind of averaging window (or something similar) that can be used to control the bandwidth variations for the transmitted stream.

It is therefore proposed to use a Token Bucket model to describe the bandwidth with two parameters, the token bucket rate and the bucket size, see RFC 2212 [RFC2212].

3.2. Point-to-point Sessions using SDP offer/answer

The existing modifier for the application specific bandwidth 'b=AS' is frequently used in the SDP offer/answer negotiation RFC 3264 [RFC3264] for setting up point-to-point sessions, for example for bi-directional point-to-point VoIP or video telephony sessions. In this section, the use of the legacy bandwidth modifier is reviewed for the use in point-to-point sessions using SDP offer/answer.

3.2.1. Symmetric Point-to-point Sessions, Fixed-rate Codecs

This example below shows the SDP offer from end-point A for several fixed-rate codecs, mu-law and A-law PCM/G.711 [G.711], AD-PCM/G.726 [G.726] and CS-ACELP/G.729 [G.729]. The codecs have different bit rates. PCM encodes speech at 64 kbps. G.726 can encode speech at four different rates, 64, 32, 24 and 16 kbps, but in this case it is assumed that the 32 kbps variant is used. G.729 encodes speech at 8 kbps. The IP/UDP/RTP overhead with 20 ms packetization and IPv4 becomes 16 kbps in all cases giving 80, 48 and 24 kbps, respectively. The media bandwidth, negotiated with b=AS, needs to be set to the highest of the bandwidths required for each respective codec.

```
m=audio 49200 RTP/AVP 8 0 96 18
b=AS:80
a=rtpmap:96 G726-32/8000/1
a=ptime:20
a=maxptime:80
```

SDP offer for mu-law PCM, A-law PCM, G.726 and G.729 with IPv4

If end-point B accepts to use this codec then a likely SDP answer would be:

```
m=audio 49400 RTP/AVP 8 0 96 18
b=AS:80
a=rtpmap:96 G726-32/8000/1
a=ptime:20
a=maxptime:80
```

SDP answer for mu-law PCM, A-law PCM, G.726 and G.729 with IPv4

In this case, both end-points offer to receive 80 kbps. A resource allocation function would thereby allocate 80 kbps in each direction.

However, if end-point B accepts to use one of the lower rate codecs, for example G.729, but rejects the PCM and G.726 codecs, then a likely SDP answer would be:

```
m=audio 49400 RTP/AVP 18
b=AS:24
a=ptime:20
a=maxptime:80
```

SDP answer for G.729 with IPv4

This means that the offerer has offered to receive 80 kbps while the answerer has offered to receive only 24 kbps. In the direction A to B it is clear that a resource allocation function should allocate 24 kbps. However, in the direction B to A it is a more unclear. On one hand, end-point A has offered to receive 80 kbps, but on the other hand, end-point B has only indicated support for the G.729 codec which only supports up to 8 kbps encoding (excluding IP overhead). It is therefore unclear if end-point B will send with only 24 kbps or if the remaining bandwidth will be used, for example for application layer redundancy.

A resource allocation may also (incorrectly) conclude that end-point B will also send maximum 24 kbps, since b=AS indicates 24 kbps. But, since maxptime is 80 ms, this means that end-point B could very well use application layer redundancy and encapsulate redundant frames together with non-redundant frames, which would result in a bandwidth exceeding 24 kbps. Even if maxptime would be 20 ms, end-point B could still use application layer redundancy, if the non-redundant frames and the redundant frames are transmitted in different packets. This is possible since end-point A has indicated that it is capable of receiving 80 kbps. Hence, if the resource allocation function uses the codec information and assumes that end-point B will send with only 24 kbps, then this may cause packet losses and/or long delays.

It should be clear with this example that the current bandwidth

attribute, b=AS, can create ambiguities related to what bandwidth that will be used in each direction. If the end-points and the resource allocation functions make different interpretations then there is a risk for either poor quality or wasted resources.

To solve this, a new bandwidth negotiation method should enable negotiating different bandwidths for different codecs and/or different codec configurations. If a codec can be configured to give several different bandwidths, e.g. G.726 offers the possibility to use four different static bit rates then this would typically be negotiated using different RTP Payload Types. This means that the solution needs to be capable of negotiating different bandwidths for different RTP Payload Types.

3.2.2. Symmetric Point-to-Point Sessions with Rate-Adaptive Codec

This use case describes what might happen when using a rate-adaptive codec in a session, for example AMR [AMR]. The rate adaptation should adapt to a high bitrate when the operating conditions are good, but should adapt to a low bitrate when the operating conditions are degraded, e.g. due to congestion or bad coverage.

One example of the SDP offer-answer negotiation for rate-adaptive codec is shown below.

```
m=audio 49200 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=ptime:20
a=maxptime:100
```

SDP offer from end-point A for AMR and IPv4

The bandwidth attribute in the SDP indicates the bandwidth that the offerer would like to receive, RFC 3264 [RFC3264].

```
m=audio 49100 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=ptime:20
a=maxptime:100
```

SDP answer from end-point B also for AMR and IPv4

The bandwidth attribute in the SDP answer indicates the maximum bandwidth that the answerer would like the offerer to use when

sending media, RFC 3264 [RFC3264].

In this case, it is clear that both end-points are prepared to receive up to 29 kbps of media. Since AMR can adapt the rate for the encoding, this means that the bandwidth can be reduced, e.g. to the 5.9 kbps mode, if congestion is detected. The existing bandwidth attribute 'b=AS' is however only used to negotiate the maximum rate. This means that there is nothing in the SDPs that describes how the rate will be adapted during congestion. In some cases, usually for speech codecs, it might be possible to derive the lowest possible rate from the codec information. However, there is no guarantee that the end-points will adapt to this rate or whether it will stay at some higher rate. For video codecs, there is usually no codec information at all that could be used to determine how low rate the end-points will use. The lowest usable rate for a video codec is generally not a video codec limitation, but rather some end-user or service consideration on what is the lowest video quality that is still useful or acceptable in the actual scenario.

This means that a resource allocation function has no information which could be used to determine how the end-points will adapt during periods of congestion. Hence the network does not know what to assume from the end-points.

To solve this, a new bandwidth negotiation method should allow for negotiating not only the highest rate but also the minimum rate that is still useful.

3.2.3. Symmetric Point-to-Point Sessions with Several Rate-Adaptive Codecs

Another example is when the originating end-point offers several rate-adaptive codecs, with different bandwidths, and when the answerer only support one or several of the lower-rate configurations but not the configuration that uses the highest bandwidth. With the legacy bandwidth modifier 'b=AS' it is only possible to indicate one bandwidth for the whole RTP session, which means that the end-point needs to indicate the highest bandwidth since this is the worst-case scenario. An offer/answer for this case is shown below. The offerer supports both AMR and AMR-WB [AMR-WB] and therefore indicates the bandwidth needed for the AMR-WB configuration since it is higher than for AMR. If the answerer does not support the AMR-WB codec then it will have to remove this configuration from the SDP when creating the SDP answer. This means that the answerer calculates the bandwidth required for AMR instead of AMR-WB.

```
m=audio 49200 RTP/AVP 96 97
b=AS:41
a=rtpmap:96 AMR-WB/16000/1
a=fmtp:96 mode-change-capability=2; max-red=80
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=ptime:20
a=maxptime:100
```

SDP offer from end-point A for AMR-WB, AMR and IPv4

```
m=audio 49100 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=ptime:20
a=maxptime:100
```

SDP answer from end-point B for AMR and IPv4 (AMR-WB is removed)

Since the indicated bandwidth is for the receiving direction in this example this means that:

- o A must send media with a bandwidth not exceeding 29 kbps; and:
- o B must send media with a bandwidth not exceeding 41 kbps.

This gives the same problem with ambiguous maximum rate as shown in Section 3.2.1. In addition, since both AMR and AMR-WB are rate-adaptive codecs, with different bit rates, they also have different minimum rates. This means that a resource allocation would be unaware about both the maximum bandwidth and the minimum (required) bandwidth.

To solve this, a new bandwidth attribute should allow for negotiating both maximum and minimum bitrates individually for each payload type.

For speech codecs, it is usually possible to derive the minimum rate from the codec information. However, this is typically not possible for video codecs since they only indicate the maximum encoding level. For example, if end-point A offers to use H.264 level 3.0 H.264 [H.264] but end-point B is only capable of using level 1.2, then this only limits the maximum bandwidth in the direction from A to B. In the other direction, end-point A is still capable of receiving level 3.0.

3.2.4. Asymmetric Point-to-Point Sessions

The session setup for asymmetric streams is not always straight forward. Lets say that one want to set up a session with 600 kbps in the sending direction and 200 kbps in the receiving direction.

```
m=video 49200 RTP/AVP 96
b=AS:200
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c00c
a=sendrecv
```

SDP offer to receive 200 kbps video

From this SDP, it can be determined that the end-point wants to receive 200 kbps. There is some implicit information in the level part of the profile-level-id for the H.264 example above, indicating that the end-point can send using a higher bandwidth (up to 768 kbps), but it requires codec-specific knowledge to be able to extract that implicit information. In this example, lets assume that the sender does not even want to utilize the maximum allowed bandwidth for the signaled codec level, but a slightly lower one, say 600 kbps. There could be many reasons to use a lower video bandwidth than the one defined by the level maximum, for example: limited terminal performance in the send direction, a known network bandwidth limitation, a bandwidth charging model that makes the user prefer a lower bandwidth, etc. There is however nothing in the SDP that indicates that the offerer in this case only wants to send video with a bandwidth up to 600 kbps, especially since it does not want to use the bandwidth implicitly indicated with the codec level. Hence, the answerer cannot know what bandwidth the offerer intends to use in the sending direction.

One way to express the asymmetry is to set up different RTP sessions for sending and receiving directions. An SDP offer for this might be:

```
m=video 49200 RTP/AVP 96
b=AS:600
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c00d
a=sendonly
m=video 49202 RTP/AVP 97
b=AS:200
a=rtpmap:97 H264/90000
a=fmtp:97 profile-level-id=42c00c
a=recvonly
```

SDP offer with separate sessions for send and receive

If the answerer decides to accept this then the SDP answer might be:

```
m=video 49200 RTP/AVP 96
b=AS:600
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c00d
a=recvonly
m=video 49202 RTP/AVP 97
b=AS:200
a=rtpmap:97 H264/90000
a=fmtp:97 profile-level-id=42c00c
a=sendonly
```

SDP answer with separate sessions for send and receive

In this example, it is clear that the offerer can send video with 600 kbps and receive video with up to 200 kbps. However, if the offer is for different codecs, using different bandwidths, then one have the same problem as described in Section 3.2.3.

Specifically for video, but possibly also for other media, it may happen that different implementations send the media in different ways. Some implementations may try to provide a fairly "smooth" stream in terms of bandwidth variation over time, while other implementations may give a very "bursty" stream.

There also exist cases where opening additional RTP sessions just for expressing asymmetric transmission bandwidths are not desirable.

3.3. Sessions with Multiple Streams

In this part of the analysis, it is assumed that an RTP session is set up for multiple streams. This can be done in several ways and for several reasons, as discussed in RTP Multiplexing Architecture [I-D.westerlund-avtcore-multiplex-architecture].

3.3.1. Multiple Streams

The assumed usage here is a multi-party session, for example a video conference using an RTP mixer. Some of the attendees are active and their audio and video is distributed to the other users. Some attendees are inactive and thus only receive media. In this example, each end-point sends one video stream, but can receive up to four simultaneous video streams, multiplexed as different SSRC in the same RTP session. One or more central nodes (RTP Mixer) are used to help facilitate the media transport between the participants, and are involved in choosing the streams to be forwarded. In this example it is assumed that there is an aggregate bandwidth limit of 3 Mbps in the receive direction, and that each received video stream should be limited to max 1 Mbps.

An SDP offer for the setting up a session with one video stream for the sending direction and four video streams for the receiving direction is shown below when using [I-D.westerlund-avtcore-max-ssrc] to explicitly declare capability to handle multiple streams. In this case, only the legacy 'b=AS' bandwidth attribute is used, valid only for the aggregate.

```
m=video 49300 RTP/AVP 96
b=AS:3000
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c016
a=max-recv-ssrc:* 4
```

SDP offer to receive multiple video streams

This example again highlights the asymmetry problem with the existing bandwidth attribute, but it also highlights the lack of per-stream bandwidth specification. This means that it is not possible to declare the 1 Mbps bandwidth limit that should be used for each one of the for streams in the receiving direction, which thus is a desirable property of the new bandwidth attribute. Note also that in this example, the 1 Mbps limit per stream cannot be fully utilized if all four streams are used simultaneously.

3.4. User Experience and Bandwidth Negotiation

Resource allocation is typically a compromise between perceived quality and network utilization. From an end-user perspective, the bandwidth for a service should be as high rate as possible, since this should give the best user experience. However, from a network perspective, one would like to minimize the rate, since this should maximize the number of sessions that can be supported.

For some services, like conversational voice- and/or video-telephony, there is a need to ensure that the network is capable of delivering a certain minimum rate, even when the network load is high. This is needed to ensure user satisfaction, both in terms of quality and end-to-end delay. If the minimum rate cannot be delivered then there is no meaning in setting up the session. This means that the end-points and the network need to agree on what maximum bandwidth that can be used for the session as well as some lowest useful "least required" bandwidth.

The current bandwidth modifier, 'b=AS', is used to negotiate the maximum bandwidth. However, since it only allows for negotiate one bandwidth it cannot be used to also negotiate a lower bandwidth limit.

To solve this, a new bandwidth negotiation method should allow for negotiating not only the highest rate but also the "at least required" rate. To enable a negotiation between the end-point and the network, a reasonable approach is that the end-point requests a lower bandwidth limit and then the network indicate what "least required" rate that was granted.

3.5. Summary of Findings

It should be clear from the above discussion that the current bandwidth attribute is too limited to be used for all use cases and that some extensions are needed.

The current bandwidth attribute, 'b=AS', is sufficient for simple sessions but gives ambiguities when negotiating more advanced session types. One of the drawbacks is that 'b=AS' only indicates the desired bandwidth for the receiving direction but. If the answering end-point wants to use a lower rate than what is offered, then there is often no way for the resource allocation function to know what bandwidth that will be used in the offerer's sending direction.

Implementers of end-points and resource allocation functions may try to resolve this ambiguity by using other information available in the SDP, e.g. codec-specific information. However, such information is not always easily available, e.g. for video codecs.

End-points may have to perform a second offer/answer negotiation to resolve the ambiguity. This, obviously, has the drawbacks that the SIP traffic is increased and that this takes some extra time. It is also not guaranteed that the end-points will actually initiate a second offer/answer negotiation.

The analysis above has also shown that the current bandwidth

attribute is insufficient to properly describe the session for multi-stream scenarios.

Furthermore, the analysis above has also shown that the current bandwidth modifier can be used to negotiate the maximum bit rate in bearers allocated in some wireless networks, but it is insufficient for also negotiating a lower, "least required", bandwidth limit.

Another problem with the existing bandwidth attribute is that the syntax is very limited and does not allow for introducing extensions. Only additional identifiers with a single value each can be added.

It is therefore proposed to define a new bandwidth attribute, including a new syntax. The new bandwidth attribute should support:

Directionality: One need to be able to have different sets of attributes values depending on the direction.

Payload specific: With the new bandwidth attribute it should be possible to specify different bandwidth values for different RTP Payload types. This is because some codecs have different characteristics and one may want to limit a specific codec and payload configuration to a particular bandwidth. Especially combined with codec negotiation there is a need to express intentions and limitations on usage for that particular codec. In addition, payload agnostic information is also needed.

Multiple streams: The new bandwidth attribute should support bandwidth negotiation both for single streams and for multiple streams. When multiple streams are used, the new bandwidth attribute should allow for declaring both the bandwidth per stream and the aggregated bandwidth.

Bandwidth specification method: To have a clear specification of what any bit-rate values mean we propose that Token bucket parameters should be used, i.e. bucket depth and bucket fill rate, where appropriate for the semantics. If single values are to be specified, a clear definition on how to derive that value must be specified, including averaging intervals etc.

Bandwidth semantics: It should be possible to negotiate different types of bandwidths for each scope, including several bandwidth properties in the same negotiation. It should, at least, be possible to negotiate the highest bandwidth. It should also be possible to negotiate a lower bandwidth limit that indicates the lowest useful bandwidth to use for the related media. The least required bandwidth limit should ideally, but need not necessarily, be guaranteed by the network and the remote end-point(s).

Extensibility: The semantics need to be extensible, so that new semantics can be defined in the future.

The existing bandwidth modifier, 'b=AS', is widely used today. The existing SDP attributes for directionality, 'a=sendrecv', 'a=recvonly', 'a=sendonly' and 'a=inactive', are also widely used. It is therefore important to ensure interworking between the new bandwidth attribute and the mechanisms already existing in SDP.

4. Attribute Specification

This section proposes a new bandwidth attribute 'a=bw' that can be used either as an extension to the already existing bandwidth attribute 'b=AS' or replacing the existing bandwidth attribute. The new bandwidth attribute includes semantics that allows for also replacing the existing bandwidth attribute.

The syntax for the new bandwidth attribute is:

a=bw:<direction> <scope> <semantic>:<value>

where:

<direction> is the direction in which the scope and semantics applies,

<scope> describes for what scope the definitions applies,

<semantic> is the actual bandwidth specification,

<value> in the form defined by the semantic used.

The new attribute is designed to allow for future extendability.

4.1. SDP Grammar

The ABNF RFC 5234 [RFC5234] for this attribute is the following:

```

bw-attrib      = "a=bw:" direction SP [req] scope SP
                  [req] semantics ":" values
direction      = "send" / "recv" / "sendrecv" / direction-ext
scope          = payloadType / scope-ext
payloadType    = "pt=" ("*" / (PT-spec) *("," PT-spec))
PT-spec        = PT-value / PT-value-range
PT-value       = 1*3DIGIT
PT-value-range = PT-value "-" PT-value
req            = "!"
semantics       = "SMT" / "AMT" / "SLT" / "SLTR" / "ALT" / "ALTR" /
                  semantics-ext
values         = token-bucket / value-ext
token-bucket   = "tb=" br-value ":" bs-value
br-value       = "*" / 1*15DIGIT ; Bucket Rate [bps]
bs-value       = "*" / 1*15DIGIT ; Bucket Size [bytes]

direction-ext  = token ; As defined in RFC 4566
scope-ext      = 1*VCHAR ; As defined in RFC 5234
semantics-ext  = token ; As defined in RFC 4566
value-ext      = 0*(WSP / VCHAR) ; As defined in RFC 5234

```

ABNF for 'bw' attribute

The 'a=bw' attribute defines three possible directionalities for the bandwidth:

send: In the send direction for SDP Offer/Answer agent or in case of declarative use in relation to the device that is being configured by the SDP.

recv: In the receiving direction for the SDP Offer/Answer agent providing the SDP or in case of declarative use in relation to the device that is being configured by the SDP.

sendrecv: The provided bandwidth values apply equally in send and receive directions, i.e. the values configures the directions symmetrically.

The directionality must be specified when the 'a=bw' attribute is used. Only one directionality can be specified on each 'a=bw' line. Special care must be taken to avoid conflicting definitions. For example, if 'sendrecv' has been specified on one 'a=bw' line for a scope, e.g. payload number 96, then the direction cannot be set to 'send' or 'recv' on another 'a=bw' line for the same scope. However, it is allowed to specify directionality 'send' on one 'a=bw' line for a scope and directionality 'recv' on another 'a=bw' line. This is useful when the bandwidth is different in different directions. Using 'sendrecv' as directionality on an 'a=bw' line is a shortcut in

the sense that it is equivalent to using two separate 'a=bw' lines where one uses 'send' and the other 'recv' but that otherwise are semantically identical.

The scope indicates what is being configured by the bandwidth semantics on this attribute line. Two different scopes are defined based on payload type:

Payload Type: The bandwidth configuration applies to the specific payload type value(s).

pt=*: Applies to all payload types being used.

Using pt=* indicates that the definitions apply to all payload types being used. The scope may be a single payload type value, e.g. pt=96. A list of payload type values can be created by using a comma-separated list, e.g. pt=96,98,105. It is also possible to specify a range of payload type values, e.g. pt=96-102, which means that the definitions apply to all the payload type numbers from 96 to 102. It is also possible to combine payload type values, payload type lists and payload type ranges, e.g. pt=96,98-102,104,105,110-113.

The scope parameter is extensible to allow for adding other scope definitions in the future.

This specification defines six related semantics. All semantics represent either the bandwidth consumption of a single stream or the aggregate of streams as a token bucket defining a transmission profile which the media sender must stay within. The token bucket values are the token rate in bits per second and the bucket size in bytes both provided as integers, see RFC 2212 [RFC2212]. The below semantics includes the whole IP packet, for example IP, UDP, RTP headers and RTP payload, as what shall be metered when determining if the send pattern is within the profile. The token bucket definition allows for wild cards to specify that one want values to be specified for the token bucket, but that one has no values to propose.

The definitions of the semantics in more detail are:

SMT (Stream Maximum Token bucket): The maximum intended or allowed bandwidth, including protocol overhead, for each individual source (each SSRC) in an RTP session, as specified by a token bucket at the sender. The token bucket wild cards ("*") should not be used for the SMT semantics since it should always be possible to estimate the maximum bandwidth. This semantics is possible to use with the scope for any payload type (pt=*) where it applies independent of encoding and packetization, or for a specific or a

set of payload type(s).

AMT (Aggregate Maximum Token bucket): The maximum intended or allowed aggregate bandwidth, including protocol overhead, for the sum of all sources (all SSRCs) for the given scope in an RTP session, as specified by a token bucket at the sender. The token bucket wild cards ("*") should not be used for the AMT semantics since it should always be possible to estimate the maximum bandwidth. The 'sendrecv' directionality parameter indicates equal token buckets in both directions, i.e. the aggregate of streams sent to an end-point shall be within the token bucket defined transmission profile, and the aggregate of streams sent from that end-point shall also be within the same token bucket profile at the sender. It can be used either to express the maximum for one particular payload type, for a set of payload types or for any payload type (pt=*).

SLT (Stream Least required Token bucket): The least required bandwidth, including protocol overhead, needed for the stream for each individual source (each SSRC) in an RTP session, as specified by a token bucket at the sender. The least required bandwidth is the minimum bandwidth that is necessary for the service to work with usable quality. When using the SLT semantic, the SMT semantic SHOULD also be specified for the same direction and scope. If the SLT semantics is not defined then this means that the least required bandwidth limit is zero. This semantics can be used with the scope for any payload type (pt=*) where it applies independent of encoding and packetization, or for a specific or a set of payload type(s).

SLTR (Stream Least required Token bucket Request): The request for establishing the least required bandwidth, including protocol overhead, needed for the stream for each individual source (each SSRC) in an RTP session, as specified by a token bucket at the stream sender. An end-point may use the SLTR semantics to request to establish a least required bandwidth for a stream. An end-point using the SLTR semantics may set the token bucket rate and/or the token bucket size to "*" to indicate that the end-point has no preference, but that it expects some network node or the answering end-point to define the value(s). A network node answering to the SLTR SHALL replace this with the SLT semantics to indicate the least required bandwidth it sees necessary and which it has attempted to guarantee. If the request is for certain specified payload types, a network node that cannot grant bandwidth based on payload types MAY replace those requested payload types with "*" in the SLT response to indicate a payload type agnostic grant. An end-point receiving an SDP with SLTR, i.e. where the network has not replaced the SLTR semantics with

any SLT semantics, SHOULD NOT assume that the requested bandwidth is guaranteed. This semantics can be used with the scope for any payload type (pt=*) where it applies independent of encoding and packetization, or for a specific or a set of payload type(s).

ALT (Aggregated Least required Token bucket): The least required aggregate bandwidth, including protocol overhead, needed for the sum of all sources (all SSRCs) for the given scope in an RTP session, as specified by a token bucket at the sender. The least required aggregate bandwidth is the minimum bandwidth that is necessary for the service to work with usable quality. If the ALT semantics is not defined then this means that the least required aggregate bandwidth is zero. It may still happen that the least required bandwidth is defined for some or all individual streams using the SLT semantics. When using the ALT semantic the AMT semantic SHOULD also be specified for the same direction and scope. The directionality and payload type considerations for ALT are the same as for AMT. This semantics can be used with the scope for any payload type (pt=*) where it applies independent of encoding and packetization, or for a specific or a set of payload type(s).

ALTR (Aggregated Least required Token bucket Request): The request for establishing the least required aggregate bandwidth, including protocol overhead, needed for the sum of all sources (all SSRCs) for the given scope in an RTP session, as specified by a token bucket at the sender. An end-point may use the ALTR semantics to request to establish a least required bandwidth for aggregated streams. The directionality and payload type considerations for ALTR are the same as for SLTR. When using the ALTR semantics, the AMT semantics SHOULD also be specified for the same direction and scope. A network node answering to the ALTR SHALL replace this with the ALT semantics to indicate the least required bandwidth it sees necessary and which it has attempted to guarantee. If the request is for certain specified payload types, a network node that cannot grant bandwidth based on payload types MAY replace those requested payload types with "*" in the ALT response to indicate a payload type agnostic grant. An end-point receiving an SDP with ALTR, i.e. where the network has not replaced the ALTR semantics with any ALT semantics, SHOULD NOT assume that the requested bandwidth is guaranteed. This semantics can be used with the scope for any payload type (pt=*) where it applies independent of encoding and packetization, or for a specific or a set of payload type(s).

The SMT and AMT semantics, with or without SLT and ALT respectively, may be used both symmetrically and in a particular direction. They can be used either to express the maximum (and

minimum) for one particular payload type, for a set of payload types or for any payload type (pt=*).

The required prefix ("!") is used when the direction, scope and semantics is required be supported and understood by the SDP consuming end-point.

4.2. Declarative Use

In declarative usage the SDP attribute is interpreted from the perspective of the end-point being configured by the particular SDP. An interpreter MAY ignore 'a=bw' attribute lines that contains unknown scope or semantics that does not start with the required ("!") prefix. If a "required" prefix is present at an unknown scope or semantics, the interpreter SHALL NOT use this SDP to configure the end-point.

4.3. Usage in Offer/Answer

The offer/answer negotiation is performed for each 'a=bw' attribute line individually with the scope and semantics immutable.

An offerer may use the 'a=bw' attribute(s) for some or all of the offered media types. An answerer may remove the 'a=bw' attribute(s) for the media types where it was used in the SDP offer.

The SDP may include an offer for an Aggregated Maximum Token bucket (AMT) without specifying any Stream Token Buckets (SMTs) for any individual streams. Correspondingly, the SDP may include an offer for an Aggregated Least required Token bucket (ALT) or Aggregated Least required Token bucket Request (ALTR) without specifying any Stream Least required Token bucket (SLT) or Stream Least required Token bucket Request (STLR), respectively.

When using the 'a=bw' attribute to define the token bucket for a certain scope then the offerer should define token buckets for all scopes of the same type. For example, if the SDP offer includes three payload types, e.g. 96, 97 and 98, and if a token bucket is defined for payload type 96, then the offerer should also define token buckets for the other payload types. This can be done either by defining one token bucket each for payload type 97 and 98 or by defining a common token bucket for payload type 97 and 98.

When the token bucket rate and size are declared in an offer for directionality 'sendrecv' then this indicates the token bucket rate and the token bucket sizes are the same in both directions. For example, if the offered bandwidth is 1 Mbps, then the end-point declares that it is capable of sending with a bandwidth up to 1 Mbps

and that it is capable of receiving with a bandwidth up to 1 Mbps.

If either the token bucket rate(s) or the token bucket sizes are different in sending and receiving direction then 'sendrecv' cannot be used. One should instead include two or more 'a=bw' lines with the respective directionality, bandwidths and sizes.

When the token bucket parameters are declared in an SDP offer for directionality 'send' then this indicates the token bucket parameters the sender intends to use. The answerer may change this value, both to increase it and to reduce it, see below.

When the token bucket parameters are declared in an SDP offer for directionality 'recv' then this indicates that the largest envelope for the token bucket parameters that the offerer thinks the media sender shall use.

When the token bucket parameters are declared in an SDP offer for directionality 'send' and 'sendrecv' then the payload type number is only used to reference the configuration for which the token bucket parameters apply. Normal offer/answer rules are then used to determine what payload type number that will be used when sending RTP media to the receivers.

An agent understanding the 'a=bw' attribute and answering to an offer including the 'a=bw' attribute SHOULD include the attribute in the answer for all media types for which it was offered.

An answerer SHOULD ignore 'a=bw' attribute lines that contains unknown scope or semantics that does not contain the required ("!") prefix. If a "required" prefix is present at an unknown scope or semantics, then the answerer SHALL reject the media description by setting the port to 0 and copy the 'a=bw' attributes not understood in the answer. In this case, 'a=bw' attributes that are understood SHALL NOT be included in the answer.

If an answerer would like to add additional bandwidth configurations using other directionality, scope, and semantics combination, then it MAY do so by adding such definitions in the SDP answer.

An agent may also divide an 'a=bw' offer into several 'a=bw' offers. One example is when the SDP offer included an 'a=bw' offer with directionality 'sendrecv', which indicates that the token bucket parameters are the same in sending and receiving direction. If the answerer would like to change the parameters for one or both directions, so that the parameters are no longer the same for both directions, then the answerer can include two 'a=bw' lines in the SDP answer, one for sending direction and another for receiving

direction. In case an offered sendrecv media becomes a single direction media then the sendrecv can be modified to that single direction.

An agent responding to an offer will need to consider the directionality and reverse them in the answer when responding to media streams using unicast.

For media stream offers over unicast with directionality send, the answerer SHALL reverse the directionality and indicate its reception bandwidth capability, which may be lower or higher than what the sender has indicated as its intended maximum.

For media stream offers over unicast with directionality receive, the token bucket parameters indicate the upper limits. The answerer SHALL reverse the directionality and may reduce the bandwidth when producing the answer indicating the answerer intended maximum transmission rate.

If the answerer removes one or several RTP Payload Types from the SDP when creating the SDP answer then the corresponding 'a=bw' lines SHOULD be removed as well. The answerer MAY however keep an 'a=bw' line when the removed RTP Payload Type number is included within an identified range or list of Payload Type numbers.

4.4. Bucket Size Estimation

In SDP bandwidth terms, the bucket size is a new parameter and what value to use for it may be hard to understand for implementers of this specification. This section therefore gives some guidelines on how to set bucket size values.

A token bucket specifies an envelope for a transmission profile where individual measurements have some impact if the media stream or aggregate should be considered within the specified profile. The semantics defined in this document only require that the media stream is within the token bucket specification at the point of emitting it into the network. The network may add jitter causing the media stream/aggregate to no longer be within the specified token bucket profile.

4.4.1. Sender Specified Token Bucket

A sender SHOULD base the choice of token bucket size on how it plans to send data. That can in turn be decided from e.g. codec configuration, intended number of encoded frames per packet (ptime), network interface, maximum transmission unit (MTU), etc. In practice, for the simplified case where the sender is designed to

send all packets with precisely even time spacing, the token bucket size can be set to the maximum packet size and the bit-rate to the long term highest bit-rate intended to be used.

However, for media streams that are more variable the bucket parameters should be chosen so that the emitted traffic is not too bursty measured over a shorter interval. Until the bucket is drained, the media sender will be able to emit packets at or close to the interface's maximum bit-rate. Long burst of packets at interface speed becomes more sensitive to loss due to cross-traffic in switching fabrics with small buffers. Due to this, a sender can consider transmission scheduling to a rate lower than the interface rate but higher than the token bucket average rate.

Let's consider the example of a large video intra frame consisting of 10 full MTU (let's assume 1500 bytes) packets which is 5 times the size of the median frame size of two full MTU packets. The average bit-rate may be 1 Mbps. If the token bucket was to be configured to (1 Mbps, 1500) then that would imply that a new full MTU packet could be emitted no more often than one packet every 12 ms. That would require 120 ms to transmit the intra frame, which for a 25 frames per second video is 3 frame intervals. Thus potentially inducing significant playout jitter at a receiver. A token buffer specification of (1 Mbps, 15000 bytes) would allow all 10 packets to be sent with up to line speed. This could result in them being emitted every 1.2 ms over a 100 Mbps interface if there is no competing traffic. To ensure that a 10 packet burst should be possible to transmit within one frame interval of 40 ms, then the bucket depth needed is the burst size in bits, minus the time interval, times the bucket fill rate, and the resulting value converted back into bytes: $(15000 * 8 - 0.04 * 1M) / 8 = 10000$ bytes. The average bit-rate for this intra frame over a single frame period becomes 4 Mbps. So the question is if bursts up to 4 Mbps should be allowed now and then as long as the average is within 1 Mbps, or if the sender has to transmit the intra using several frame intervals, skipping the next frame(s) and hoping that the receiver doesn't drop the intra frame as being too late. The sender could also consider reducing the quality of the intra frame, resulting in a reduced number of MTU required to transmit it.

A sender SHOULD avoid adding excessive safety margins to the sending bucket size. A sender MAY add bucket size margins if it has knowledge of internal transmission timing variations, or if it knows about packet handling outside the sender itself that will affect the effective bucket size (as seen from a receiver) that is otherwise not reflected in the conveyed bucket size figure.

4.4.2. Receiver Specified Token Bucket

With the semantics specified in this document, the intended media receiver gets to provide token bucket parameters that specifies how the sender should behave. The traffic received by the receiver (or intermediate nodes) may no longer conform to the token bucket due to jitter introduced by the network path between the sender and the receiver. This document assumes that the receiver will have receiver buffers for de-jittering that are significantly larger than the token bucket parameters. This due to that a media unit like a video frame may be transmitted over time using more data than the bucket depth provides and instead spread it in time, transmitting each fragment when the bucket is refilled enough for the next fragment to be sent.

A receiver's input to the sender's bit-rate limitation should be based on known limitations such as the networks, decoding capabilities etc. The bucket depth will control how bursty the traffic can be beyond the long term average specified by the bucket refill rate.

4.4.3. Bucket Adjustment in Middle Nodes

When there are media aware middle nodes on the media path between the sender and receiver, those middle nodes may have to or want to apply similar considerations as the original media sender and receiver. If those middle nodes are aware of the SDP and the new bandwidth attribute from this specification, and have in-path SDP adjustment capabilities, they could benefit from modifying the values to better fit the actually available end-to-end media path capabilities. For example, an RTP Media Translator can express what it actually is going to deliver of the far end-point's media to an end-point instead of that far end-point's provided values.

4.4.4. Network Policing

As the token bucket specified for the semantics in this document is based on what the sender emit into the network, a policer should have some margin allowing for network introduced jitter. The amount will of course be dependent on the policer's location in relation to the media sender.

4.4.5. Utilizing Network Feedback

If the media uses RTP and when the media has been transmitted for some time, the sender should have received a fair amount of RTCP receiver reports from the receiver. The sender can from RTCP estimate the observed network jitter at the receiver and may be able to dynamically adjust the sender behavior such that the aggregate of

the sender behavior and the reported network jitter are fulfilling the senders token bucket profile.

4.5. SDP Examples for Point-to-point Sessions

These SDP examples show how the new bandwidth attribute can be used. The benefits, compared to the legacy bandwidth attribute, are also highlighted.

The SDP examples included below are intentionally not complete. Only the parts that are relevant for this description are included.

The SDP examples assume that IPv4 is used.

4.5.1. Symmetric Fixed-rate Codecs

This example shows the SDP offer for several fixed-rate codecs, mu-law and A-law PCM, G.726 and G.728. The token bucket size is allocated to allow for storing up to 5 packets when PCM coding is used and 20 ms of speech is encapsulated in each packet.

```
m=audio 49200 RTP/AVP 8 0 96 18
b=AS:80
a=rtpmap:96 G726-32/8000/1
a=bw:sendrecv pt=0,8 SMT:tb=80000:1000
a=bw:sendrecv pt=96 SMT:tb=48000:1000
a=bw:sendrecv pt=18 SMT:tb=24000:1000
a=ptime:20
a=maxptime:20
```

SDP offer for mu-law and A-law PCM and IPv4

The new bandwidth attribute offers the possibility to negotiate the bandwidth individually for each codec. If the answerer removes a codec when creating the answer then it is still known how much bandwidth the other codecs will use. This means that the ambiguities listed in Section 3.2.1 can be avoided.

4.5.2. Symmetric Rate-Adaptive Codec

This example shows the SDP negotiation for offering using the AMR codec [AMR]. It is assumed that the bandwidth-efficient payload format is used. The SMT bandwidth is based on the AMR 12.2 kbps mode and the SLTR bandwidth is based on the AMR 5.9 kbps mode. The token bucket size is allocated to allow for storing up to 1 packet when the AMR 12.2 kbps mode is used and 5 frames are encapsulated in the packet, which corresponds to 197 bytes per packet.


```
m=audio 49200 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=bw:sendrecv pt=97 SMT:tb=28800:200
a=bw:sendrecv pt=97 SLTR:tb=22400:200
a=ptime:20
a=maxptime:100
```

SDP offer from end-point A for AMR and IPv4

```
m=audio 49100 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=bw:sendrecv pt=97 SMT:tb=28800:200
a=bw:sendrecv pt=97 SLT:tb=22400:200
a=ptime:20
a=maxptime:100
```

SDP answer from end-point B also for AMR and IPv4

Since the new bandwidth attribute offers a possibility to negotiate both the maximum and the at least required bandwidth, it is possible for both the other end-point and any resource allocation function to know how the end-points will adapt when congestion is detected.

4.5.3. Several Symmetric Rate-Adaptive Codecs

This example shows how the new bandwidth attribute, 'a=bw', can be used to negotiate the maximum and the least required bandwidths for several rate-adaptive codecs, in this case for AMR and AMR-WB [AMR-WB]. For AMR, the highest codec mode is 12.2 kbps, giving a maximum bandwidth of 28.8 kbps, and the at least required mode is chosen to be 5.9 kbps, giving a least required bandwidth of 22.4 kbps. For AMR-WB, the highest codec mode is 23.85 kbps, giving a maximum bandwidth of 40.4 kbps, and the least required mode is chosen to be 8.85 kbps, giving a least required bandwidth of 25.6 kbps.

```
m=audio 49200 RTP/AVP 96 97
b=AS:41
a=rtpmap:96 AMR-WB/16000/1
a=fmtp:96 mode-change-capability=2; max-red=80
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=bw:sendrecv pt=96 SMT:tb=40400: 350
a=bw:sendrecv pt=96 SLTR:tb=25600:350
a=bw:sendrecv pt=97 SMT:tb=28800:200
a=bw:sendrecv pt=97 SLTR:tb=22400:200
a=ptime:20
a=maxptime:100
```

SDP offer from end-point A for AMR-WB, AMR and IPv4

```
m=audio 49100 RTP/AVP 97
b=AS:29
a=rtpmap:97 AMR/8000/1
a=fmtp:97 mode-change-capability=2; max-red=80
a=bw:sendrecv pt=97 SMT:tb=28800:200
a=bw:sendrecv pt=97 SLT:tb=22400:200
a=ptime:20
a=maxptime:100
```

SDP answer from end-point B for AMR and IPv4 (AMR-WB is removed)

In this case, it is clear when the answer is received that the bandwidth needed for AMR applies to both directions. There is no need for a second offer/answer negotiation to clarify that the bandwidth applies also to end-point A's receiving direction. Thereby, the issues listed in Section 3.2.3 are resolved.

4.5.4. Asymmetric Session

The following SDP example shows how to use the new bandwidth attribute to offer asymmetric streams. In this case, the end-point offers to send H.264 video with 1 Mbps while it is capable of receiving H.264 with up to 3 Mbps. Note that this example does not make use of the codec-specific H.264 level asymmetry signaling as defined in RFC 6184 [RFC6184].

```
m=video 50324 RTP/AVP 96
b=AS:3000
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c016
a=bw:send pt=96 SMT:tb=1000000:8192
a=bw:recv pt=96 SMT:tb=3000000:16384
```

SDP offer with asymmetric video bandwidth

It should be clear from this example that the new bandwidth attribute is useful when negotiating asymmetric sessions since it offers the possibility to define the token bucket parameters for both sending and receiving directions separately.

4.5.5. Session with Retransmission

This SDP example shows how the new bandwidth attribute, 'a=bw', can be used for negotiating the bandwidth when the RTP Retransmission Payload Format RFC 4588 [RFC4588] is used.

```
m=video 49170 RTP/AVPF 96 97
b=AS:500
a=rtpmap:96 MP4V-ES/90000
a=rtcp-fb:96 nack
a=fmtp:96 profile-level-id=8; config=01010000012000884006682C2090A21F
a=rtpmap:97 rtx/90000
a=fmtp:97 apt=96;rtx-time=3000
a=bw:send pt=* AMT:tb=500000:4096
a=bw:recv pt=* AMT:tb=500000:8192
```

SDP offer with aggregate bandwidth and RTP retransmission

In this case, it is beneficial to use the Aggregate Maximum Token bucket semantics to allow the end-points to adapt the bandwidths used for the original stream and for the retransmission stream during the session. The end-point can send more original packets when the packet loss rate is low. When the packet loss rate is high then the end-point can use less bandwidth for the original packets and instead allow for more retransmissions. It would also be possible to specify separate limits for the original stream and the retransmission stream by using a separate set of 'a=bw'-lines for pt=96 and pt=97.

4.6. SDP Examples with Sessions with Multiple Streams

4.6.1. Multiple Streams

The example below is based on the use case described in Section 3.3.1. Only the negotiation for video is shown here.

```
m=video 49300 RTP/AVP 96
b=AS:3000
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c01f
a=bw:send pt=* SMT:tb=1000000:1000
a=bw:recv pt=* SMT:tb=1000000:2000
a=bw:send pt=* AMT:tb=1000000:1000
a=bw:recv pt=* AMT:tb=3000000:6000
a=max-recv-ssrc:* 4
```

SDP offer with both per-stream and aggregate bandwidth

With the new bandwidth attribute, it is possible to define the bandwidth for each received stream independently from each other. In this case, the SDP shows that the end-point is prepared to send maximum 1 Mbps, and that the end-point is prepared to receive maximum 1 Mbps per stream. The SDP also shows that the end-point is prepared to receive maximum 3 Mbps, aggregated for the up to four streams in the receiving direction. Note that this implies that to receive more than three streams, each stream's bandwidth must be reduced to comply with the maximum aggregate.

4.6.2. Declarative Example with Stream Asymmetry

This example shows a declarative usage of the new bandwidth attribute.

```
m=video 50324 RTP/AVP 96 97 98
a=rtpmap:96 H264/90000
a=rtpmap:97 H263-2000/90000
a=rtpmap:98 MP4V-ES/90000
a=max-recv-ssrc:96 2
a=max-recv-ssrc:* 5
a=bw:send pt=* SMT:tb=1200000:16384
a=bw:recv pt=96 SMT:tb=1500000:16384
a=bw:recv pt=97,98 SMT:tb=2500000:16384
a=bw:recv pt=* AMT:tb=8000000:65535
```

SDP offer with payload-specific per-stream bandwidth

In the above example, the outgoing single stream is limited to bucket rate of 1.2 Mbps and bucket size of 16384 bytes. The up to 5 incoming streams can in total use maximum 8 Mbps bucket rate and with

a bucket size of 65535 bytes. However, the individual streams maximum rate is depending on payload type. Payload type 96 (H.264) is limited to 1.5 Mbps with a bucket size of 16384 bytes, while the Payload types 97 (H.263) and 98 (MPEG-4) may use up to 2.5 Mbps with a bucket size of 16384 bytes.

4.7. Interoperability Issues

The proposed new bandwidth attribute is obviously related to both the bandwidth modifier 'b=AS' and the attributes defined for directionality ('a=sendrecv', 'a=sendonly', 'a=recvonly' and 'a=inactive') defined in RFC 4566 [RFC4566]. It is therefore important to properly analyze these relationships so that any interoperability issues can be avoided.

4.7.1. Interoperability with Existing Bandwidth Attribute

If the SDP includes both the 'b=AS' bandwidth modifier and 'a=bw' bandwidth attribute then alignment may be necessary to avoid confusion. This section gives some guidelines for such alignment. It may however happen that some usage needs other alignments than what is discussed below. If so, then those alignments need to be considered on a case-by-case. The discussion below should therefore not be seen as an exhaustive list.

In general, the bandwidths offered with 'b=AS' and 'a=bw' should be aligned for the direction that applies for the 'b=AS' bandwidth modifier. For 'sendrecv' and 'recvonly' sessions, 'b=AS' indicates the bandwidth for the receiving direction. The b=AS is closest in interpretation to the AMT semantic. If the stream maximum semantic (SMT) is used then the sum of the bandwidths in the receive direction may exceed the 'b=AS' bandwidth but the AMT should not exceed the b=AS value.

If the session includes multiple streams, but if not all of the streams will be active simultaneously, then 'b=AS' should indicate the maximum bandwidth that will be used for the combinations of streams that are active simultaneously, the same way AMT would be used in such a session. This also means that the bandwidths offered with 'a=bw' are accumulated for the combination of streams that are active, and this aggregated bandwidth should not exceed the bandwidth defined with 'b=AS'. Note however that it is possible and feasible to specify an aggregate that is less than the sum of the maximum bandwidth for the maximum amount of available streams. It may be possible to use the maximum number of active streams with a lower bandwidth than the maximum, or it may be possible to reduce the active number of streams to stay within the bandwidth limit.

The SDP below gives an example of how this is done. In this example, the intention is to use either the payload type pair (96, 97) or the payload type pair (98, 99). The intention is however to, for example, not pair payload types 96 and 98.

```
m=video 50000 RTP/AVP 96 97 98 99 100
b=AS:1000
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c00d
a=rtpmap:97 H264/90000
a=fmtp:97 profile-level-id=42c00c
a=rtpmap:98 H264/90000
a=fmtp:98 profile-level-id=42c00d
a=rtpmap:99 H264/90000
a=fmtp:99 profile-level-id=42c00c
a=rtpmap:100 H264/90000
a=fmtp:100 profile-level-id=42c00c
a=bw:sendrecv 96 SMT:tb=700000:4000
a=bw:recv 97 SMT:tb=300000:3000
a=bw:sendrecv 98 SMT:tb=500000:3000
a=bw:recv 99 SMT:tb=200000:2000
a=bw:send 100 SMT:tb=300000:1400
a=sendrecv
```

SDP offer with complex bandwidth relations

This session is bi-directional, as shown with the 'a=sendrecv' attribute. The bandwidth offered with 'b=AS' therefore applies to the receive direction. The 'b=AS' is then set based on the combination of streams that gives the highest bandwidth, i.e. the payload type pair (96, 97).

This means that the bandwidths offered with 'a=bw' are aligned with the bandwidth offered with 'b=AS'.

If, on the other hand, the intention would be to use another combination of payload types, for example (96, 98), then this would add up to 1200 kbps, which would mean that the stream bandwidths would not be aligned with the 'b=AS' bandwidth.

This shows that bandwidths for 'sendrecv' and 'recv' directions are added together when determining the bandwidth for the combined streams.

If the offer is "complex", for example offering multiple streams for both speech and video, possibly with many different codecs, (and therefore uses 'a=bw' together with the 'b=AS' bandwidth modifier) and if the answerer wants to change this into a "simple" session

(e.g. plain simple VoIP with only one RTP payload type for codec X) then the answerer may remove the 'a=bw' lines when creating the answer. It may therefore happen that the answer includes only 'b=AS' bandwidth modifier in the SDP answer. However, if the offer does not include any 'b=AS' line then it is recommended to maintain the 'a=bw' lines also in the answer, even for "simple" sessions. This means that the offerer cannot rely on the existence of 'a=bw' in the answer.

4.7.2. Interoperability with Existing Directional Attribute

Since the 'a=bw' attribute includes a parameter for directionality it is important to clarify the relationship to the already existing directional attributes in SDP ('sendrecv', 'sendonly', 'recvonly' and 'inactive'). In general, one can say that:

- o The SDP attribute indicates the directionality for the session.
- o The 'a=bw' attribute defines the directionality for the bandwidth for streams within the session.
- o The SDP attribute for directionality has precedence over the 'a=bw' parameter for directionality when it comes to the media that is actually being transmitted.

At session setup time, it is therefore acceptable to define streams with other directionality than what is shown with the SDP attribute for directionality. However, when media is transmitted, then the SDP attribute for directionality has to be followed. An example of this is shown below.

```
m=video 5000 RTP/AVP 96 97 98
b=AS:1000
a=rtpmap:96 H264/90000
a=fmtp:96 profile-level-id=42c00d
a=rtpmap:97 H264/90000
a=fmtp:97 profile-level-id=42c00c
a=bw:sendrecv 96 SMT:tb=700000:4000
a=bw:recv 97 SMT:tb=200000:3000
a=bw:send 97 SMT:tb=300000:1400
a=recvonly
```

SDP offer specifying bandwidth in 'inactive' direction

This means that three bandwidths are defined at session setup:

- o one stream (PT=96) for 700 kbps bi-directional video;

- o one stream (PT=97) for 200 kbps receive-only video; and:
- o one stream (PT=97) for 300 kbps send-only video.

However, since 'a=recvonly' is defined then this means that the end-point is, at the session setup time, only willing to receive media even though the SDP contains bandwidth declarations also for the sending direction. This allows for setting up streams that are effectively inactive in one or both directions from the beginning of the session and then enabling them later in the session.

This can be compared with the case when one defines one or more codecs, even if the session starts up as 'inactive'.

5. Rules and Recommendations for Extensions

The a=bw attribute is defined to be extensible and this section discusses the extension points that are available.

5.1. Directionality

The current specification defines send, recv and sendrecv. In case some new directionality behavior is needed that doesn't match the existing, a new one could be defined. This should be avoided unless a clear need for a new directionality is found.

5.2. Scope

It is expected that there will be a need to extend the bandwidth scope. This document only defines two scope types, session and payload type, and there is very likely other desirable scopes that will be defined in the future. Possible examples of scopes are those applying to a specific SSRC, a particular end-point, or a class of end-points.

5.3. Semantics

This is the extension point that is expected to be frequently used in the future. A major proliferation of semantics is not good for interoperability, but it is likely that bandwidth shortcomings or missing functionalities will be discovered in the future. Thus defining new semantics gives maximum flexibility to define the meaning of the provided value(s), the format of the values and how to interpret the directionality and scope values.

5.4. Values

This document only defines token buckets as values. In case fewer or more parameters are needed to express a particular semantics, new value formats can be defined. Defining new value formats should be done with some consideration of generality and reuse so that future semantics can also use the new value format, with the target to try to minimize the number of different formats.

6. Open Issues

This document contain a few open issues:

1. Multicast behavior needs to be specified.
2. It is an open question to decide if and how to handle the RTCP bandwidth negotiation, e.g. corresponding to b=RS and b=RR.
3. It is an open question to develop semantics for the transport independent bandwidth negotiation, e.g. corresponding to b=TIAS.
4. It is an open question what rules and recommendations there should be for extensions to this memo.

7. IANA Considerations

Following the guidelines in RFC 4566 [RFC4566] and in RFC 3550 [RFC3550], the IANA is requested to register:

1. The bw attribute as defined in Section 4.1.
2. The bw attribute directionality registry rules
3. The bw attribute scope registry rules.
4. The bw attribute semantics registry rules.
5. The bw attribute values registry rules.

This section will be filled out in future versions of this document.

8. Security Considerations

Excessive bandwidth allocation can consume all the resources, much more than what the end-point(s) intend to use. So, if a session

allocates an unnecessarily high bandwidth then this will likely mean that some other users cannot be admitted, or that they cannot get QoS guaranteed resources that they requested and have to use best effort. It can also happen that the session itself is rejected, if the end-points try to allocate resources that are not available. Allocating too little bandwidth is likely to negatively impact the perceived media quality or entirely prevent reception of requested media.

The above shows that the bandwidth attribute is a potential vector for attacks both from malicious end-points or third party attackers that attempts to modify the attribute to impact the system to allocate unnecessary resources, deny end-points service, reduce quality for end-points or incur cost on users.

To prevent third party attacks the signaling should be source authenticated and integrity protected to prevent any on or off-path attacker from injecting or modifying the SDP. Malicious end-points can't as easily be protected against using crypto, instead behavior analysis and preventing such a malicious end-point from having serious impact on other end-points are needed.

9. Acknowledgements

10. References

10.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC2212] Shenker, S., Partridge, C., and R. Guerin, "Specification of Guaranteed Quality of Service", RFC 2212, September 1997.
- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, June 2002.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.
- [RFC3556] Casner, S., "Session Description Protocol (SDP) Bandwidth Modifiers for RTP Control Protocol (RTCP) Bandwidth", RFC 3556, July 2003.

- [RFC3890] Westerlund, M., "A Transport Independent Bandwidth Modifier for the Session Description Protocol (SDP)", RFC 3890, September 2004.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

10.2. Informative References

- [AMR] "3GPP TS 26.090, "Adaptive Multi-Rate (AMR) speech codec; Transcoding functions".", June 1999.
- [AMR-WB] "3GPP TS 26.190, "Adaptive Multi-Rate - Wideband (AMR-WB) speech codec; Transcoding functions".", April 2001.
- [G.711] "ITU-T Recommendation G.711, "Pulse Code Modulation (PCM) of Voice Frequencies".", November 1988.
- [G.726] "ITU-T Recommendation G.726, "40, 32, 24, 16 kbit/s Adaptive Differential Pulse Code Modulation (ADPCM)".", December 1990.
- [G.729] "ITU-T Recommendation G.729, "Coding of speech at 8 kbit/s using conjugate-structure algebraic-code-excited linear prediction (CS-ACELP)".", March 1996.
- [H.264] "ITU-T Recommendation H.264, "Advanced video coding for generic audiovisual services".", May 2003.
- [I-D.westerlund-avtcore-max-ssrc]
Westerlund, M., Burman, B., and F. Jansson, "Multiple Synchronization sources (SSRC) in RTP Session Signaling", draft-westerlund-avtcore-max-ssrc-01 (work in progress), April 2012.
- [I-D.westerlund-avtcore-multiplex-architecture]
Westerlund, M., Burman, B., and C. Perkins, "RTP Multiplexing Architecture", draft-westerlund-avtcore-multiplex-architecture-01 (work in progress), March 2012.
- [RFC4588] Rey, J., Leon, D., Miyazaki, A., Varsa, V., and R. Hakenberg, "RTP Retransmission Payload Format", RFC 4588, July 2006.

[RFC6184] Wang, Y., Even, R., Kristensen, T., and R. Jesup, "RTP Payload Format for H.264 Video", RFC 6184, May 2011.

Authors' Addresses

Tomas Frankkila
Ericsson
Laboratoriegrand 11
SE-971 28 Lulea
Sweden

Phone: +46 10 714 30 20
Email: tomas.frankkila@ericsson.com

Magnus Westerlund
Ericsson
Farogatan 6
SE-164 80 Kista
Sweden

Phone: +46 10 714 82 87
Email: magnus.westerlund@ericsson.com

Bo Burman
Ericsson
Farogatan 6
SE-164 80 Kista
Sweden

Phone: +46 10 714 13 11
Email: bo.burman@ericsson.com

mmusic
Internet-Draft
Intended status: Standards Track
Expires: April 23, 2012

P. Yue
Huawei Technologies
October 21, 2011

RTSP Extension for Substream Control
draft-yue-mmusic-rtsp-substream-control-extension-00

Abstract

This document defines extensions to RTSP 2.0 protocol, including header "Substream", feature tag "Play.substream" , and related new status codes. These extensions enables the playback control of a media stream on substream basis.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on April 23, 2012.

Copyright Notice

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	3
2.	Terminology	3
3.	Definitions and Abbreviations	3
3.1.	Definitions	3
3.2.	Abbreviations	4
4.	Protocol Overview	4
4.1.	Substream Annotation	4
4.2.	Capability Negotiation	5
4.3.	Substream Playback Control	5
4.3.1.	Play a Substream	5
4.3.2.	Pause a Substream	7
5.	RTSP Extensions	7
5.1.	Substream Header	7
5.2.	Play.substream Feature Tag	8
5.3.	Status Code Extension	8
5.3.1.	552 Substream Type Not Recognized	8
5.3.2.	553 Substream Control Not Allowed	8
5.3.3.	554 Substream Id Not Valid	9
6.	Security Considerations	9
7.	IANA Considerations	9
7.1.	RTSP Feature-tag Extensions	9
7.2.	RTSP Status Code Extension	9
7.3.	RTSP Header Extensions	10
7.4.	Hold of Substream Type Registration	10
7.4.1.	Guidance of Substream Type Registration	10
7.4.2.	Registration of Substream Types	10
8.	References	11
8.1.	Normative References	11
8.2.	Informative References	11
	Author's Address	11

1. Introduction

Single Session Transmission (SST) is recommended for the SVC (Scalable Video Codec) video transport [RFC6190] and MVC (Multiview Video Codec) video transport [I-D.ietf-payload-rtp-mvc] . In SST, a single RTP session conveys all the related media data, namely all the bitstream components. A bitstream component here is part of a media stream that has a common property which could be:

- o Layer: in SVC media stream; In this case, a bitstream component is media data of a specific layer.
- o View: in MVC media stream; In this case, a bitstream component is media data of a specific view.

In such a SST session, one or several bitstream components together may be decodable by themselves and therefore make sense to the receiver. These bitstream component(s) combinations are here called Substream.

In SVC and MVC RTP sessions, such a Substream is identified by an Operation Point.

There are cases that a client wants to retrieve a specific substream in such a SST session. However, in the current RTSP 2.0 protocol, the control of the media is on media stream basis, e.g. as an RTP session when RTP is used. This prevents a client to receive only certain substream(s) from the media.

This memo extends the RTSP 2.0[I-D.ietf-mmusic-rfc2326bis] to establish a substream playback control framework, which enables a client to play a part of a media stream.

This memo also defines the substream usage for SVC and MVC.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Definitions and Abbreviations

3.1. Definitions

The following terms are used in this document and have specific meaning within the context of this document.

- o Substream: a part of a media stream containing one or more bitstream components, which can be independently decoded. Typically a Substream is identified by one operation point.
- o Substream type: the compound mode of the bitstream components in a substream, e.g. SVC, MVC, etc

3.2. Abbreviations

SVC: Scalable Video Coding
MVC: Multiview Video Coding
SST: Single Session Transmission

4. Protocol Overview

The whole framework includes substream annotation, capability negotiation and substream playback control. This section provides an overview of substream control.

4.1. Substream Annotation

Substream annotation is used to announce identifiers for all substreams available in a media stream. which can be retrieved selectively. It is done through a presentation description, typically using an SDP description. Substream annotation based on SDP is described in this section. Substream annotation based on other formats is out of scope of this document.

In SVC, a substream is described as an operation point, which consists of the bitstream components required to be able to decode a particular dependency_id, quality_id, and temporal_id combination. SVC operation points are described through the sprop-operation-point-info attribute in SDP as defined in [RFC6190].

According to [RFC6190], the value of sprop-operation-point-info consists of a comma-separated list of operation-point-description vectors. Each operation-point-description vector represents a substream. Each operation-point-description vector has ten elements, where the first element layer-ID is the identifier of the operation point. This layer-ID can be taken as the identifier for the SVC substream. Since the layer-ID is optional, it may not be included. In this case, a [dependency-ID, temporal_ID, quality-ID] combination is taken as the identifier instead.

Annotation of MVC substreams is specified in [I-D.ietf-payload-rtp-mvc].

editor notes: pending on the complement of

[I-D.ietf-payload-rtp-mvc].

4.2. Capability Negotiation

Capability negotiation can be used to negotiate support of substream control between RTSP client and server. It makes use of the RTSP feature tag mechanism defined in RTSP 2.0[I-D.ietf-mmusic-rfc2326bis]. A new feature tag (play.substream) is defined in Section 5.2.

Whenever it has received a media presentation with substream annotation, a substream control capable RTSP client SHALL include play.substream in the SUPPORTED header of the RTSP SETUP request. This indicates that the client is able to control the playback of the media on substream basis.

A 2xx response implies that the server is capable of substream control. A server MAY refuse the request with a 551 "Option not supported" response, with an UNSUPPORT header including play.substream.

The receiving client may try to setup a session without this feature later. This means that the client will not perform substream control and play the media as a whole. In the SVC case, the client will play all the layers in the session, whereas in the MVC case it will play all the views in the session.

4.3. Substream Playback Control

In case the session is setup correctly, the client may control the play back on substream basis, including:

- o start to play a substream
- o pause a substream. Note that this means it will pause the play back of the whole session.

4.3.1. Play a Substream

After successful set up of an RTSP session, the client may perform substream playback control. The playback of a substream is similar with the normal playback of a session, except that the PLAY request shall contain a "SubstreamCtrl" header with substream type and the substream id corresponding to the substream that the client intends to play. Here the substream id identifies a specific operation point. The actual definition of the substream id is defined by each substream type.

When more than one media stream are controlled, the header shall further contain Request-URIs for each of the the substreams.

Following is the specification of substream types for SVC and MVC media stream. Specification for other substream type is out of scope of current document and should be registered in IANA, as described in Section 7.4.

When the media stream is an SVC stream as defined in [RFC6190], the substream type shall be "SVC"; The substream id could be either a layer-id or a [dependency-ID, temporal_ID, quality-ID] combination, which comes from an operation-point-description vector.

When the media stream is an MVC stream as defined in [I-D.ietf-payload-rtp-mvc], the substream type shall be "MVC".

For MVC, substream id is...

editor's note: substream id for MVC is pending on the complement of draft-ietf-payload-rtp-mvc-00.

A client should not perform substream play if the server has not indicated support of substream control in an earlier message exchange.

Upon receiving a PLAY request with a "SubstreamCtrl" header, the server SHALL identify the substream(s) according to the id(s) and Request-URI(s) in the request, and then provides the indicated substream(s) after sending a 200 OK response.

If the server doesn't support substream control, it should respond a 551 "Option not supported" response as defined in RTSP 2.0[I-D.ietf-mmusic-rfc2326bis].

If the server supports substream control but the substream type indicated in the "SubstreamCtrl" header is not recognized, it shall response with a 552 "Substream Type Not Recognizable" response, see Section 5.3.1.

If the server supports substream control but the substream type indicated in the "SubstreamCtrl" header is not recognized, it shall response with 552 "Substream Type Not Recognizable", see Section 5.3.1.

If a requested media is not allowed to be played on substream basis as requested, the server SHALL respond with a 553 "Substream Control Not Allowed" response, see Section 5.3.2.

If the requested substream id is not valid, the server shall response with 554 "Substream id not valid", see Section 5.3.3.

4.3.2. Pause a Substream

The pause operation of a substream is identical to the pause operation of a normal session. It is not necessary for the client to include a "SubstreamCtrl" header in the pause request message. However, if the request includes a "SubstreamCtrl" header, it shall list all the substreams are currently played.

If aggregated control is used, it is not allowed to pause only a part of a session. It is also not allowed to pause only a specific substream from a media stream.

5. RTSP Extensions

This section documents the extension to the RTSP 2.0 specification. Specifically Section 5.1 specifies the SubstreamCtrl header, Section 5.2 specifies the substream control feature tag, Section 5.3 specifies the status codes extensions for the substream control feature.

5.1. Substream Header

SubstreamCtrl header is used to indicate the substream(s) of a media stream that the client intends to play. It contains one or more [stream uri, substream type, substream id] triple.

The syntax is:

```
substream = "substream:" substream-id * (";" substream-id)
substream-id = stream-uri "," substream-type "," substream_id
stream-uri = RTSP-URI
substream-type = "SVC" / "MVC" / token
```

When the substream-type is SVC, the syntax of substream_id is:

```
substream_id      = layer-id
                    /dependency-id "," temporal-id "," quality-id
layer-id          = "layer_id=" layer_id_value
layer_id_value    = 1*4DIGIT; 0~2047
dependency-id     = "dependency_id=" dependency_id_value
dependency_id_value = DIGIT ; 0~7
temporal-id       = "temporal_id=" temporal_id_value
temporal_id_value = DIGIT ; 0~7
quality-id        = "quality_id=" quality_id_value
quality_id_value  = 1*2DIGIT; 0~15
DIGIT             = %x30-39 ; any US-ASCII digit "0".."9"
```

An example of Substream header is: substream:
rtsp://example.com/svc.mp4, SVC, lay-id=1

For the MVC case, the syntax of substream_id is:

editor's notes: pending on the complement of mvc-operation-point-id

```
HCOLON = *( SP / HT ) ":" SWS
SEMI    = SWS ";" SVS ; semicolon
SWS     = [LWS] ; Separating White Space
DIGIT   = %x30-39 ; any US-ASCII digit "0".."9"
COMMA   = SWS "," SWS ; comma
EQUAL   = SWS "=" SWS ; equal
```

SubstreamCtrl header can be used in PLAY and PAUSE request. Proxies shall not modify this header and pass through to the server.

5.2. Play.substream Feature Tag

The following feature-tag is defined in this specification and hereby registered. The change control belongs to the IETF. play.substream: Support of substream control operations for media playback. Applies only for servers.

5.3. Status Code Extension

This clause defines the status code extended for the substream control feature. They are:

- o 552 Substream Type Not Recognized
- o 553 Substream Control Not Allowed
- o 554 Substream Id Not Valid

5.3.1. 552 Substream Type Not Recognized

The server can not understand the substream type.

5.3.2. 553 Substream Control Not Allowed

The substream specified in the request is not allowed for the media identified by the Request-URI. The response shall include a xxxx header to indicate the substream information.

Editor notes: whether there is a need to return substream information is pending for discussion.

5.3.3. 554 Substream Id Not Valid

The substream id specified in the request is not valid for the media identified by the Request-URI. The response shall include a xxxx header to indicate the substream information.

Editor notes: whether there is a need to return substream information is pending for discussion.

6. Security Considerations

Editor notes: pending for more discussion.

7. IANA Considerations

Registration is requested for the newly defined RTSP header extensions, RTSP status code extensions and RTSP feature tag extensions, according to the instructions in section 22 of the base specification [I-D.ietf-mmusic-rfc2326bis].

This section also sets up a registry for substream type that should be maintained by IANA.

7.1. RTSP Feature-tag Extensions

The following feature-tag is defined in this specification and hereby registered according to the section 22.1.2 of the base specification [I-D.ietf-mmusic-rfc2326bis]. The change control belongs to the IETF.

- o Feature tag: substream.control
- o Description: Support of control the media playback on substream basis. Applies for both clients, servers and proxies.
- o Contact person: Peiyu YUE
- o Change control: IETF
- o Reference specification: present document.

7.2. RTSP Status Code Extension

The following RTSP status codes are in defined this specification and hereby registered according to section 22.3.2 of the base specification [I-D.ietf-mmusic-rfc2326bis]. The change control belongs to the IETF.

- o Request number: 552
- o Description: as described in Section 5.3.1
- o Request number: 553
- o Description: as described in Section 5.3.2
- o Request number: 554
- o Description: as described in Section 5.3.3

7.3. RTSP Header Extensions

The following RTSP header is defined in this specification and hereby registered according to section 22.4.2 of the base specification [I-D.ietf-mmusic-rfc2326bis]. The change control belongs to the IETF.

- o The name of the header: SubstreamCtrl
- o Description:
- o The syntax of the header: as described in Section 5.1.
- o Where to use: as described in Section 5.1.
- o Handle by proxy: as described in Section 5.1.

7.4. Hold of Substream Type Registration

7.4.1. Guidance of Substream Type Registration

IANA should take the responsibility for the registration of the substream type. A new substream type MUST be registered through an IETF Standards Action. A specification for a new substream type MUST consist of the following items:

- o A substream type;
- o A description of the substream type;
- o A substream id definition

7.4.2. Registration of Substream Types

This specification registers two substream types: SVC and MVC:

- o Substream Type: SVC
- o Description: Scalable Video Codec stream as defined in [RFC6190].
- o Substream id definition: could be layer-id or [dependency-ID, temporal_ID, quality-ID] combination, as described in Section 4.1.
- o Substream Type: MVC
- o Description: Multiview Video Codec stream as defined in [I-D.ietf-payload-rtp-mvc].

- o Substream id definition: as described in Section 4.1.

8. References

8.1. Normative References

- [I-D.ietf-mmusic-rfc2326bis]
Schulzrinne, H., Rao, A., Lanphier, R., Westerlund, M.,
and M. Stiemerling, "Real Time Streaming Protocol 2.0
(RTSP)", draft-ietf-mmusic-rfc2326bis-27 (work in
progress), March 2011.
- [I-D.ietf-payload-rtp-mvc]
Wang, Y. and T. Schierl, "RTP Payload Format for MVC
Video", draft-ietf-payload-rtp-mvc-00 (work in progress),
March 2011.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC6190] Wenger, S., Wang, Y., Schierl, T., and A. Eleftheriadis,
"RTP Payload Format for Scalable Video Coding", RFC 6190,
May 2011.

8.2. Informative References

- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session
Description Protocol", RFC 4566, July 2006.

Author's Address

Peiyu Yue
Huawei Technologies
Huawei Base
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Phone: +86-25-56620258
Email: yuepeiyu@huawei.com

mmusic
Internet-Draft
Intended status: Standards Track
Expires: September 3, 2012

P. Yue
Huawei Technologies
March 2, 2012

RTSP Extension for Substream Control
draft-yue-mmusic-rtsp-substream-control-extension-01

Abstract

This document defines extensions to RTSP 2.0 protocol, including header "SubstreamCtrl", feature tag "Play.substream", and related new status codes. These extensions enables the playback control of a media stream on substream basis.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at <http://datatracker.ietf.org/drafts/current/>.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 3, 2012.

Copyright Notice

Copyright (c) 2012 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust's Legal Provisions Relating to IETF Documents (<http://trustee.ietf.org/license-info>) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1.	Introduction	3
2.	Terminology	3
3.	Definitions and Abbreviations	3
3.1.	Definitions	3
3.2.	Abbreviations	4
4.	Protocol Overview	4
4.1.	Substream Annotation	4
4.2.	Capability Negotiation	5
4.3.	Substream Playback Control	5
4.3.1.	Substream Play/Resume	5
4.3.2.	Substream Pause	6
5.	RTSP Extensions	6
5.1.	Play.substream Feature Tag	7
5.2.	Substream Header	7
5.3.	Status Code Extension	7
5.3.1.	552 Substream Type Not Recognized	7
5.3.2.	553 Substream Control Not Allowed	8
5.3.3.	554 Substream Id Not Valid	8
6.	SVC and MVC Substream Type	8
6.1.	SVC Substream Type	8
6.2.	MVC Substream Type	9
7.	Examples	9
8.	Security Considerations	12
9.	IANA Considerations	12
9.1.	RTSP Feature-tag Extensions	12
9.2.	RTSP Header Extensions	13
9.3.	RTSP Status Code Extension	13
9.4.	Hold of Substream Type Registration	13
9.4.1.	Guidance of Substream Type Registration	13
9.4.2.	Registration of Substream Types	13
10.	References	14
10.1.	Normative References	14
10.2.	Informative References	14
	Author's Address	14

1. Introduction

Single Session Transmission (SST) is recommended for the SVC (Scalable Video Codec) video transport [RFC6190] and MVC (Multiview Video Codec) video transport [I-D.ietf-payload-rtp-mvc] . In SST, a single RTP session conveys all the related media data, namely all the bitstream components. A bitstream component here is part of a media stream that has a common property which could be:

- o Layer: in SVC media stream; In this case, a bitstream component is media data of a specific layer.
- o View: in MVC media stream; In this case, a bitstream component is media data of a specific view.

In such a SST session, one or several bitstream components together may be decodable by themselves and therefore make sense to the receiver. Here these bitstream component(s) combinations are here Substream.

In SVC and MVC RTP sessions, such a Substream is identified by an Operation Point.

There are cases that a client wants to retrieve a specific substream in a SST session. However, in the current RTSP 2.0 protocol, the control of the media is on basis of media stream i.e. an RTP session when RTP is used as transport protocol. This prevents a client to receive only certain substream(s) of the media.

This memo extends the RTSP 2.0[I-D.ietf-mmusic-rfc2326bis] to establish a substream playback control framework, which enables a client to play a part of a media stream.

This memo also defines the substream usage for SVC and MVC.

2. Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

3. Definitions and Abbreviations

3.1. Definitions

The following terms are used in this document and have specific meaning within the context of this document.

- o Substream: a part of a media stream containing one or more bitstream components, which can be independently decoded. Typically a Substream is identified by a single operation point.
- o Substream type: the compound mode of the bitstream components in a media stream, e.g. SVC, MVC, etc

3.2. Abbreviations

SVC: Scalable Video Coding
 MVC: Multiview Video Coding
 SST: Single Session Transmission

4. Protocol Overview

The whole framework includes three parts: substream annotation, capability negotiation and substream playback control. This section provides an overview of substream control framework.

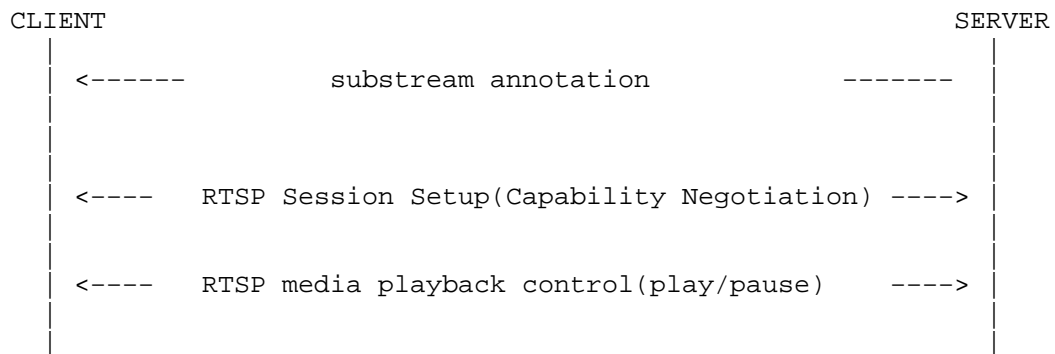


Figure 1: Substream Control Framework Overview

4.1. Substream Annotation

Substream annotation is used to announce all substreams available in a media stream, which can be retrieved selectively. This is done through a presentation description, typically using SDP description.

Substream annotation SHALL signal the substream type and substream id for each substream.

Substream type is the compound mode of the bitstream components in a media stream, e.g. SVC, MVC, etc. Here a media stream with SVC substream type, means conforming to [RFC6190]. A media stream with MVC substream type, means conforming to [I-D.ietf-payload-rtp-mvc]. More specification of SVC and MVC substream type is defined at

Section 6.

Substream id is the identifier for a substream. The actual definition of the substream id is defined by each substream type.

4.2. Capability Negotiation

Capability negotiation is used to negotiate support of substream control between RTSP client and server. It makes use of the RTSP feature tag mechanism defined in RTSP 2.0[I-D.ietf-mmusic-rfc2326bis]. A new feature tag (play.substream) is defined in Section 5.1 of this document.

If received a media presentation with substream annotation, a substream control capable RTSP client SHALL include play.substream in the REQUIRE header of the RTSP SETUP request. This indicates that the client is able to control the playback of the media on substream basis.

A 2xx response implies that the server is capable of substream control. A server MAY refuse the request with a 551 "Option not supported" response, with an UNSUPPORT header including play.substream.

The receiving client may try to setup a session without this feature later. This means that the client will not perform substream control and will retrieve the media as a whole.

4.3. Substream Playback Control

In case the session is setup correctly, the client may control the play back on substream basis, including:

- o start to play/resume a substream
- o pause a substream. Note that this means pause the play back of the whole session.

4.3.1. Substream Play/Resume

After successful set up of an RTSP session, the client may perform substream playback control. The playback of a substream is similar with the normal playback of a session, except that the PLAY request SHALL contain a "SubstreamCtrl" header to signal the substream(s) that the client intends to play.

The SubstreamCtrl header, as defined in Section 5.2, SHALL contain at least the substream type and the substream id of each substream. In the case of aggregate control, the header SHALL further contain Request-URIs for each of the the substreams.

A client SHOULD NOT perform substream play if the server has not indicated support of substream control during session setup.

Upon receiving a PLAY request with a "SubstreamCtrl" header, the server SHALL identify the substream(s) according to the substream type, substream id and optional Request-URI combination(s) in the request, and then provides the indicated substream(s) after sending a 200 OK response.

If the server doesn't support substream control, it SHALL respond with a 551 "Option not supported" response as defined in RTSP 2.0[I-D.ietf-mmusic-rfc2326bis].

If the server supports substream control but the substream type indicated in the "SubstreamCtrl" header is not recognized, it SHALL respond with a 552 "Substream Type Not Recognizable" response, see Section 5.3.1.

If a requested media is not allowed to be played on substream basis as requested, the server SHALL respond with a 553 "Substream Control Not Allowed" response, see Section 5.3.2.

If the requested substream id is not valid, the server SHALL respond with a 554 "Substream id not valid", see Section 5.3.3.

4.3.2. Substream Pause

The pause operation of a substream is identical to the pause operation of a normal RTSP session. It is not necessary for the client to include a "SubstreamCtrl" header in the PAUSE request message. However, if the request includes a "SubstreamCtrl" header, it SHALL list all the substreams are currently played.

Upon receiving a PAUSE request without a SubstreamCtrl header, the server SHALL pause the stream(s) according to RTSP 2.0[I-D.ietf-mmusic-rfc2326bis].

Upon receiving a PAUSE request with a SubstreamCtrl header, the server SHALL check the substream list contained. If the substream list is not identical to the ones that is in the play status, the server SHALL respond with a 553 "Substream Control Not Allowed" response, see Section 5.3.2. Otherwise, the server SHALL respond with 2xx message and pause the stream(s) properly.

5. RTSP Extensions

This section documents the extension to the RTSP 2.0 specification.

Specifically Section 5.2 specifies the SubstreamCtrl header, Section 5.1 specifies the substream control feature tag, Section 5.3 specifies the status codes extensions for the substream control feature.

5.1. Play.substream Feature Tag

The following feature-tag is defined in this specification and hereby registered. The change control belongs to the IETF.

play.substream: Support of substream control operations for media playback. Applies for both clients and servers.

Notes that this feature is based the play.basic therefore support of this feature inherently means support of play.basic.

5.2. Substream Header

SubstreamCtrl header is used to indicate the substream(s) of media stream(s) that the client intends to play.

SubstreamCtrl header can be used in PLAY and PAUSE request. Proxies shall not modify this header and pass through to the server.

SubstreamCtrl header contains one or more [stream uri, substream type, substream id] triple. The syntax is:

```
SubstreamCtrl = "SubstreamCtrl:"  
                substream-id * (";" substream-id)  
substream-id  = [stream-uri ","]  
                substream-type ", " substream_id  
stream-uri    = RTSP-URI  
substream-type= "SVC" / "MVC" / token
```

5.3. Status Code Extension

This clause defines the status codes extended for the substream control feature. They are:

- o 552 Substream Type Not Recognized
- o 553 Substream Control Not Allowed
- o 554 Substream Id Not Valid

5.3.1. 552 Substream Type Not Recognized

The server can not recognize one or more substream type(s) in the SubstreamCtrl header of the request.

5.3.2. 553 Substream Control Not Allowed

One or more media streams identified by the Request-URI in the request is not allowed for the substream control.

5.3.3. 554 Substream Id Not Valid

One or more substream id(s) specified in the request are not valid for the media identified by the Request-URI.

6. SVC and MVC Substream Type

6.1. SVC Substream Type

This section provides the specification of substream type for SVC Stream: Substream Type: "SVC" Substream annotation:

In SVC, a substream is identical to an operation point, which consists of the bitstream components required to be decoded independently.

Operation points are signalled in SDP, by either of the following two parameters :

- o sprop-operation-point-info: According to [RFC6190], this parameter consists of a comma-separated list of operation-point-description vectors, including layer-ID, dependency-ID, temporal_ID, quality-ID. Among them, the value of layer-ID specifies the layer identifier of the operation point, which is identical to the layer_id that would be indicated (for the same values of dependency_id, quality_id, and temporal_id) in the scalability information SEI message.
- o sprop-scalability-info: According to [RFC6190], This parameter is used to convey the NAL unit containing the scalability information SEI message as specified in Annex G of [H.264]. Within the scalability information SEI message, layer information of the SVC stream is signalled, including layer_id, dependency_id, quality_id, and temporal_id.

Therefore, substreams in a SVC stream is identified by either layer-ID or a [dependency-ID, temporal_ID, quality-ID] combination.

The syntax of substream_id is:

```

substream_id      = layer-id
                    /(dependency-id ","temporal-id "," quality-id)
layer-id          = "layer_id=" layer_id_value
layer_id_value    = 1*4DIGIT                               ;0~2047
dependency-id     = "dependency_id=" dependency_id_value
dependency_id_value = DIGIT                                 ;0~7
temporal-id       = "temporal_id=" temporal_id_value
temporal_id_value = DIGIT                                   ;0~7
quality-id        = "quality_id=" quality_id_value
quality_id_value  = 1*2DIGIT                                ;0~15
DIGIT             = %x30-39                                ;any US-ASCII digit "0".."9"

```

An example of Substream header is:

```
SubstreamCtrl: rtsp://example.com/svc.mp4, SVC, lay-id=1
```

6.2. MVC Substream Type

This section provides the specification of substream type for SVC Stream: Substream Type: "SVC" Substream annotation: Annotation of MVC substreams is specified in [I-D.ietf-payload-rtp-mvc].

editor notes: pending on the complement of [I-D.ietf-payload-rtp-mvc].

Substream_id:

For the MVC case, the syntax of substream_id is:

editor's notes: pending on the complement of mvc-operation-point-id

7. Examples

The following takes substream control of a SVC stream as an example. Step 1: Client C requests a presentation from media server S. The media is a video encoded by SVC and transported by SST.

```
C->S: DESCRIBE rtsp://example.com/BreakingDawn.3gp RTSP/2.0
      CSeq: 1
      User-Agent: PhonyClient/1.2

S->C: RTSP/2.0 200 OK
      CSeq: 1
      Server: PhonyServer/1.0
      Date: Wed, 22 Feb 2012 15:20:29 GMT
      Content-Type: application/sdp
      Content-Length: 407
      Content-Base: rtsp://example.com/BreakingDawn.3gp/
      Expires: 22 Feb 2012 15:20:29 GMT

v=0
o=- 2890844256 2890842807 IN IP4 198.51.100.5
s=RTSP Session
i=An Example of substream control usage
e=adm@example.com
c=IN IP4 0.0.0.0
a=control: *
a=range: npt=0-0:10:34.10
t=0 0
m=video 20000 RTP/AVP 97
a=rtpmap:97 H264-SVC/90000
a=fmtp:97 profile-level-id=53000c; packetization-mode=1;
  sprop-parameter-sets={sps0},{sps1},{pps0},{pps1};
  sprop-operation-point-info=<1,0,0,0,4de00a,3200,176,144,128,
256>,<2,1,1,0,53000c,6400,352,288,256,512>;
```

Step 2: The client setup the RTSP session with play.substream feature tag.

```
C->S: SETUP rtsp://example.com/BreakingDawn.3gp RTSP/2.0
      CSeq: 2
      User-Agent: PhonyClient/1.2
      Require: play.substream
      Transport: RTP/AVP;unicast;dest_addr=":8000"/":8001"
      Accept-Ranges: NPT, SMPTE, UTC

S->C: RTSP/2.0 200 OK
      CSeq: 3
      Server: PhonyServer/1.0
      Transport: RTP/AVP;unicast; ssrc=AABBCCDD;
                dest_addr="192.0.2.53:8002"/"192.0.2.53:8003";
                src_addr="198.51.100.5:9002"/"198.51.100.5:9003";

      Session: 12345678
      Expires: 22 Feb 2012 15:22:09 GMT
      Date: 22 Feb 2012 15:22:09 GMT
      Accept-Range: NPT
      Media-Properties: Random-Access=0.8, Immutable, Unlimited
```

Step 3: The client starts the playout of the component layer 1.

```
C->S: PLAY rtsp://example.com/BreakingDawn.3gp/ RTSP/2.0
      CSeq: 4
      User-Agent: PhonyClient/1.2
      Range: npt=0-
      Seek-Style: RAP
      Session: 12345678
      SubstreamCtrl: SVC, layer_id=1

S->C: RTSP/2.0 200 OK
      CSeq: 4
      Server: PhonyServer/1.0
      Date: 22 Feb 2012 15:22:52 GMT
      Session: 12345678
      Range: npt=0-634.10
      Seek-Style: RAP
      RTP-Info: url="rtsp://example.com/BreakingDawn.3gp"
                ssrc=0D12F123;seq=12345;rtptime=3450012,
```

Step 4: The client requests to pause the session.

```
C->S: PAUSE rtsp://example.com/BreakingDawn.3gp/ RTSP/2.0
      CSeq: 5
      User-Agent: PhonyClient/1.2
      Session: 12345678
      SubstreamCtrl: SVC, layer_id=1

S->C: RTSP/2.0 200 OK
      CSeq: 5
      Server: PhonyServer/1.0
      Date: 22 Feb 2012 15:22:58 GMT
      Session: 12345678
      Range: npt=5.66-634.10
```

8. Security Considerations

Considerations outlined in RTSP 2.0 [I-D.ietf-mmusic-rfc2326bis] apply here as well. It is believed that no special security risk is led by this document.

9. IANA Considerations

Registration is requested for the newly defined RTSP feature tag extension, RTSP header extensions and RTSP status code extensions, according to the instructions in section 22 of the base specification [I-D.ietf-mmusic-rfc2326bis].

This section also sets up a registry for substream type that should be maintained by IANA.

9.1. RTSP Feature-tag Extensions

The following feature-tag is defined in this specification and hereby registered according to the section 22.1.2 of the base specification [I-D.ietf-mmusic-rfc2326bis]. The change control belongs to the IETF.

- o Feature tag: Play.substream
- o Description: Support of control the media playback on substream basis. Applies for clients, servers and proxies.
- o Contact person: Peiyu YUE
- o Change control: IETF
- o Reference specification: present document.

9.2. RTSP Header Extensions

The following RTSP header is defined in this specification and hereby registered according to section 22.4.2 of the base specification [I-D.ietf-mmusic-rfc2326bis]. The change control belongs to the IETF.

- o The name of the header: SubstreamCtrl
- o Description: SubstreamCtrl is used to indicate the substreams RTSP client wants to control(i.e. play or pause).
- o The syntax of the header: as described in Section 5.2.
- o Where to use: as described in Section 5.2.
- o Handle by proxy: default.

9.3. RTSP Status Code Extension

The following RTSP status codes are in defined this specification and hereby registered according to section 22.3.2 of the base specification [I-D.ietf-mmusic-rfc2326bis]. The change control belongs to the IETF.

- o Request number: 552
- o Description: as described in Section 5.3.1
- o Request number: 553
- o Description: as described in Section 5.3.2
- o Request number: 554
- o Description: as described in Section 5.3.3

9.4. Hold of Substream Type Registration

9.4.1. Guidance of Substream Type Registration

IANA should take the responsibility for the registration of the substream type. A new substream type MUST be registered through an IETF Standards Action. A specification for a new substream type MUST consist of the following items:

- o A substream type;
- o A description of the substream type;
- o A substream id definition

9.4.2. Registration of Substream Types

This specification registers two substream types: SVC and MVC:

- o Substream Type: SVC
- o Description: Scalable Video Codec stream as defined in [RFC6190].
- o Substream id definition: could be layer-id or [dependency-ID, temporal_ID, quality-ID] combination, as described in Section 4.1.

- o Substream Type: MVC
- o Description: Multiview Video Codec stream as defined in [I-D.ietf-payload-rtp-mvc].
- o Substream id definition: as described in Section 4.1.

10. References

10.1. Normative References

- [I-D.ietf-mmusic-rfc2326bis]
Schulzrinne, H., Rao, A., Lanphier, R., Westerlund, M.,
and M. Stiemerling, "Real Time Streaming Protocol 2.0
(RTSP)", draft-ietf-mmusic-rfc2326bis-27 (work in
progress), March 2011.
- [I-D.ietf-payload-rtp-mvc]
Wang, Y. and T. Schierl, "RTP Payload Format for MVC
Video", draft-ietf-payload-rtp-mvc-00 (work in progress),
March 2011.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC6190] Wenger, S., Wang, Y., Schierl, T., and A. Eleftheriadis,
"RTP Payload Format for Scalable Video Coding", RFC 6190,
May 2011.

10.2. Informative References

- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session
Description Protocol", RFC 4566, July 2006.

Author's Address

Peiyu Yue
Huawei Technologies
Huawei Base
101 Software Avenue, Yuhua District
Nanjing, Jiangsu 210012
China

Phone: +86-25-56620258
Email: yuepeiyu@huawei.com

