

Network Working Group
Internet-Draft
Intended status: Standards Track
Expires: November 30, 2012

H. Alvestrand
Google
May 29, 2012

Cross Session Stream Identification in the Session Description Protocol
draft-alvestrand-rtcweb-msid-02

Abstract

This document specifies a grouping mechanism for RTP media streams that can be used to specify relations between media streams within different RTP sessions.

This mechanism is used to signal the association between the RTP concept of SSRC and the WebRTC concept of "media stream" / "media stream track" using SDP signalling.

This document is an input document for discussion. It should be discussed in the RTCWEB WG list, rtcweb@ietf.org.

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].

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 November 30, 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
1.1. Why A New Mechanism Is Needed	3
1.2. Application to the WEBRTC MediaStream	3
2. The Msid Mechanism	4
3. The Msid-Semantic Attribute	5
4. Applying Msid to WebRTC Media Streams	5
4.1. Handling of non-signalled tracks	6
5. IANA Considerations	7
6. Security Considerations	8
7. Acknowledgements	8
8. References	8
8.1. Normative References	8
8.2. Informative References	8
Appendix A. Design considerations, open questions and and alternatives	9
Appendix B. Change log	10
B.1. Changes from -00 to -01	10
B.2. Changes from -01 to -02	10
Author's Address	10

1. Introduction

1.1. Why A New Mechanism Is Needed

There exist cases where an application using RTP and SDP needs to signal some relationship between RTP media streams (packets carried using a single SSRC) that may be carried in either the same RTP session or different RTP sessions.

When all SSRCs are carried in a single RTP session, the "a=ssrc-group" mechanism [RFC5576] can be used.

When each RTP session carries one and only one SSRC, the SDP grouping framework [RFC5888] can be used.

However, there are use cases (some of which are discussed in [I-D.westerlund-avtcore-multiplex-architecture]) where neither of these approaches is appropriate; for instance, there may be a need to signal a relationship between a video track in one RTP session and an audio track in another RTP session. In those cases, a new mechanism is needed.

(Note: When the bundle mechanism, [I-D.ietf-mmusic-sdp-bundle-negotiation], is used, the extension is still needed to link SSRCs under different m= lines, even when they are in the same RTP session).

In addition, there is sometimes the need for an application to specify some application-level information about the association between the SSRC and the group. This is not possible using either of the frameworks above.

1.2. Application to the WEBRTC MediaStream

The W3C WebRTC API specification [W3C.WD-webrtc-20120209] specifies that communication between WebRTC entities is done via MediaStreams, which contain MediaStreamTracks. A MediaStreamTrack is generally carried using a single SSRC in an RTP session (forming an RTP media stream. The collision of terminology is unfortunate.) There might possibly with additional SSRCs, possibly within additional RTP sessions, in order to support functionality like forward error correction or simulcast. This complication is ignored below.

In the RTP specification, media streams are identified using the SSRC field. Streams are grouped into RTP Sessions, and also carry a CNAME. Neither CNAME nor RTP session correspond to a MediaStream. Therefore, the association of an RTP media stream to MediaStreams need to be explicitly signalled.

The marking needs to be on a per-SSRC basis, since one RTP session can carry media from multiple MediaStreams, and one MediaStream can have media in multiple RTP sessions. This means that the [RFC4574] "label" attribute, which is used to label RTP sessions, is not usable for this purpose.

The marking needs to also carry the unique identifier of the RTP media stream as a MediaStreamTrack within the media stream; this is done using a single letter to identify whether it belongs in the video or audio track list, and the MediaStreamTrack's position within that array.

This usage is described in Section 4.

2. The Msid Mechanism

Grouping of SSRCs is done via an "msid" attribute attached to the SSRC in the SDP description, using the "Source Specific Media Attribute" mechanism [RFC5576]:

```
a=ssrc:1234 msid:examplefoo v1
```

The ID is a randomly-generated string of ASCII characters chosen from 0-9, a-z, A-Z and - (hyphen), consisting of between 1 and 64 characters. It MUST be unique among the ID values used in the same SDP session.

The value "default" (all lower case) has special meaning, and MUST NOT be generated. Values starting with "example" (all lower case) are reserved for documentation, and MUST NOT be generated by an implementation.

Application data is carried on the same line as the ID, separated from the ID by a space.

ABNF[RFC5234] grammar:

```
msidattribute = "msid:" identifier [ " " appdata ]
identifier = 1*64 ("0".."9" / "a".."z" / "-")
appdata = 1*64 ("0".."9" / "a".."z" / "-")
```

(Note: one possible generation algorithm is to generate 6 random bytes, base64 encode them (giving 8 bytes), and prefixing with a letter that is neither "d" nor "e". Another possibility is using some form of UUID.)

The ID uniquely identifies a group within the scope of an SDP description.

There may be multiple msid attributes on a single SSRC.

3. The Msid-Semantic Attribute

In order to fully reproduce the semantics of the SDP and SSRC grouping frameworks, a session-level attribute is defined for signalling the semantics associated with an msid grouping.

This OPTIONAL attribute gives the message ID and its group semantic.
a=msid-semantic: examplefoo LS

The ABNF of msid-semantic is:

```
msid-semantic-attr = "msid-semantic:" " " msid token  
token = <as defined in RFC 4566>
```

The semantic field may hold values from the IANA registries "Semantics for the "ssrc-group" SDP Attribute" and "Semantics for the "group" SDP Attribute".

4. Applying Msid to WebRTC Media Streams

The semantic for WebRTC Media Streams is "WMS".

The value of the msid corresponds to the "id" attribute of a MediaStream. (note: as of Jan 11, 2012, this is called "label". The word "label" means many other things, so the same word should not be used.)

In a WebRTC-compatible SDP description, all SSRCs intending to be sent from one peer will be identified in the SDP generated by that entity.

The appdata for a WebRTC MediaStreamTrack consists of the track type and the track number; the track type is encoded as the single letter "a" (audio) or "v" (video), and the track number is encoded as a decimal integer with no leading zeroes. The first track is track zero, and is identified as "a0" for audio, and "v0" for video.

When an SDP description is updated, a specific msid continues to refer to the same media stream; an msid value MUST NOT be reused for another media stream within a PeerConnection's lifetime.

The following are the rules for handling updates of the list of SSRCs and their msid values.

- o When a new msid value occurs in the description, the recipient can signal to its application that a new media stream has been added.
- o When a description is updated to have more SSRCs with the same msid value, the recipient can signal to its application that new media stream tracks have been added to the media stream.
- o When a description is updated to no longer list the msid value on a specific ssrc, the recipient can signal to its application that the corresponding media stream track has been closed.
- o When a description is updated to no longer list the msid value on any ssrc, the recipient can signal to its application that the media stream has been closed.

OPEN ISSUE: Exactly when should the recipient signal that the track is closed? When the msid value disappears from the description, when the SSRC disappears by the rules of RFC 3550 section 6.3.4 (BYE packet received) and 6.3.5 (timeout), any of the above, or some combination of the above?

4.1. Handling of non-signalled tracks

Pre-WebRTC entities will not send msid. This means that there will be some incoming RTP packets with SSRCs where the recipient does not know about a corresponding MediaStream id.

Handling will depend on whether or not any SSRCs are signalled in the relevant RTP session. There are two cases:

- o No SSRC is signalled with an msid attribute. The SDP session is assumed to be a backwards-compatible session. All incoming SSRCs, on all RTP sessions that are part of the SDP session, are assumed to belong to a single media stream. The ID of this media stream is "default".
- o Some SSRCs are signalled with an msid attribute. In this case, the session is WebRTC compatible, and the newly arrived SSRCs are either caused by a bug or by timing skew between the arrival of the media packets and the SDP description. These packets MAY be discarded, or they MAY be buffered for a while in order to allow immediate startup of the media stream when the SDP description is updated. The arrival of media packets MUST NOT cause a new MediaStreamTrack to be created.

Note: This means that it is wise to include at least one `a=ssrc:` line with an `msid` attribute, even when no media streams are yet attached to the session. (Alternative: Mark the RTP session explicitly as "I will signal the media stream tracks explicitly").

It follows from the above that media stream tracks in the "default" media stream cannot be closed by signalling; the application must instead signal these as closed when either an RTCP BYE packet or the absence of media for a defined interval <what interval?> indicates that the stream is gone.

5. IANA Considerations

This document requests IANA to register the "msid" attribute in the "att-field (source level)" registry within the SDP parameters registry, according to the procedures of [RFC5576]

The required information is:

- o Contact name, email: IETF, contacted via rtcweb@ietf.org, or a successor address designated by IESG
- o Attribute name: `msid`
- o Long-form attribute name: Media stream group Identifier
- o The attribute value contains only ASCII characters, and is therefore not subject to the charset attribute.
- o The attribute gives an association over a set of SSRCs, potentially in different RTP sessions. It can be used to signal the relationship between a WebRTC `MediaStream` and a set of SSRCs.
- o The details of appropriate values are given in RFC XXXX.

This document requests IANA to register the "WMS" semantic within the "Semantics for the "ssrc-group" SDP Attribute" registry within the SDP parameters registry.

The required information is:

- o Description: WebRTC Media Stream, as given in RFC XXXX.
- o Token: WMS
- o Standards track reference: RFC XXXX

IANA is requested to replace "RFC XXXX" with the RFC number of this document upon publication.

6. Security Considerations

An adversary with the ability to modify SDP descriptions has the ability to switch around tracks between media streams. This is a special case of the general security consideration that modification of SDP descriptions needs to be confined to entities trusted by the application.

No attacks that are relevant to the browser's security have been identified that depend on this mechanism.

7. Acknowledgements

This note is based on sketches from, among others, Justin Uberti and Cullen Jennings.

8. References

8.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.
- [RFC5576] Lennox, J., Ott, J., and T. Schierl, "Source-Specific Media Attributes in the Session Description Protocol (SDP)", RFC 5576, June 2009.
- [W3C.WD-webrtc-20120209]
Bergkvist, A., Burnett, D., Narayanan, A., and C. Jennings, "WebRTC 1.0: Real-time Communication Between Browsers", World Wide Web Consortium WD WD-webrtc-20120209, February 2012,
<<http://www.w3.org/TR/2012/WD-webrtc-20120209>>.

8.2. Informative References

- [I-D.ietf-mmusic-sdp-bundle-negotiation]
Holmberg, C. and H. Alvestrand, "Multiplexing Negotiation Using Session Description Protocol (SDP) Port Numbers",

draft-ietf-mmusic-sdp-bundle-negotiation-00 (work in progress), February 2012.

- [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.
- [RFC4574] Levin, O. and G. Camarillo, "The Session Description Protocol (SDP) Label Attribute", RFC 4574, August 2006.
- [RFC5888] Camarillo, G. and H. Schulzrinne, "The Session Description Protocol (SDP) Grouping Framework", RFC 5888, June 2010.

Appendix A. Design considerations, open questions and alternatives

This appendix should be deleted before publication as an RFC.

One suggested mechanism has been to use CNAME instead of a new attribute. This was abandoned because CNAME identifies a synchronization context; one can imagine both wanting to have tracks from the same synchronization context in multiple media streams and wanting to have tracks from multiple synchronization contexts within one media stream.

Another suggestion has been to put the msid value within an attribute of RTCP SR (sender report) packets. This doesn't offer the ability to know that you have seen all the tracks currently configured for a media stream.

There has been a suggestion that this mechanism could be used to mute tracks too. This is not done at the moment.

The special value "default" and the reservation of "example*" seems bothersome; apart from that, it's a random string. It's uncertain whether "example" has any benefit.

An alternative to the "default" media stream is to let each new media stream track without a msid attribute create its own media stream. Input on this question is sought.

Discarding of incoming data when the SDP description isn't updated yet (section 3) may cause clipping. However, the same issue exists when crypto keys aren't available. Input sought.

There's been a suggestion that acceptable SSRCs should be signalled

in a response, giving a recipient the ability to say "no" to certain SSRCS. This is not supported in the current version of this document.

This specification reuses the ssrc-group semantics registry for this semantic, on the argument that the WMS purpose is more similar to an SSRC grouping than a session-level grouping, and allows values from both registries, on the argument that some semantics (like LS) are well defined for MSID. Input sought.

Appendix B. Change log

This appendix should be deleted before publication as an RFC.

B.1. Changes from -00 to -01

Added track identifier.

Added inclusion-by-reference of draft-lennox-mmusic-source-selection for track muting.

Some rewording.

B.2. Changes from -01 to -02

Split document into sections describing a generic grouping mechanism and sections describing the application of this grouping mechanism to the WebRTC MediaStream concept.

Removed the mechanism for muting tracks, since this is not central to the MSID mechanism.

Author's Address

Harald Alvestrand
Google
Kungsbron 2
Stockholm, 11122
Sweden

Email: harald@alvestrand.no

MMUSIC
Internet-Draft
Intended status: Standards Track
Expires: September 12, 2012

A. Begen
Y. Cai
H. Ou
Cisco
March 11, 2012

Duplication Grouping Semantics in the Session Description Protocol
draft-begen-mmusic-redundancy-grouping-03

Abstract

Packet loss is undesirable for real-time multimedia sessions, but can occur due to congestion, or other unplanned network outages. This is especially true for IP multicast networks, where packet loss patterns can vary greatly between receivers. One technique that can be used to recover from packet loss without incurring unbounded delay for all the receivers is to duplicate the packets and send them in separate redundant streams. This document defines the semantics for grouping redundant streams in the Session Description Protocol (SDP). The semantics defined in this document are to be used with the SDP Grouping Framework [RFC5888]. SSRC-level (Synchronization Source) grouping semantics are also defined in this document for RTP streams using SSRC multiplexing.

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 12, 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. Requirements Notation	3
3. Duplication Grouping	3
3.1. "DUP" Grouping Semantics	3
3.2. DUP Grouping for SSRC-Multiplexed RTP Streams	4
3.3. SDP Offer/Answer Model Considerations	4
4. SDP Examples	5
4.1. Separate Source Addresses	5
4.2. Separate Destination Addresses	5
4.3. Temporal Redundancy	6
5. Security Considerations	7
6. IANA Considerations	7
7. Acknowledgments	8
8. References	8
8.1. Normative References	8
8.2. Informative References	8
Authors' Addresses	9

1. Introduction

The Real-time Transport Protocol (RTP) [RFC3550] is widely used today for delivering IPTV traffic, and other real-time multimedia sessions. Many of these applications support very large numbers of receivers, and rely on intra-domain UDP/IP multicast for efficient distribution of traffic within the network.

While this combination has proved successful, there does exist a weakness. As [RFC2354] noted, packet loss is not avoidable, even in a carefully managed network. This loss might be due to congestion, it might also be a result of an unplanned outage caused by a flapping link, link or interface failure, a software bug, or a maintenance person accidentally cutting the wrong fiber. Since UDP/IP flows do not provide any means for detecting loss and retransmitting packets, it leaves up to the RTP layer and the applications to detect, and recover from, packet loss.

One technique to recover from packet loss without incurring unbounded delay for all the receivers is to duplicate the packets and send them in separate redundant streams. Variations on this idea have been implemented and deployed today [IC2011].

[I-D.begen-avtcore-rtp-duplication] explains how duplication can be achieved for RTP streams without breaking the RTP and RTCP functionality. In this document, we describe the semantics needed in the Session Description Protocol (SDP) [RFC4566] to support this technique.

2. Requirements Notation

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].

3. Duplication Grouping

3.1. "DUP" Grouping Semantics

Each "a=group" line is used to indicate an association relationship between the redundant streams. The streams included in one "a=group" line are called a Duplication Group.

Using the framework in [RFC5888], this document defines "DUP" as the grouping semantics for redundant streams.

The "a=group:DUP" semantics MUST be used to group the redundant streams except when the streams are specified in the same media description, i.e., in the same "m" line (See Section 3.2).

When the redundant streams are described in separate "m" lines and the 'group' attribute is used to describe the redundancy relation, the SSRCS for each redundant stream MUST be announced in the SDP description using the 'ssrc' attribute [RFC5576]. According to [I-D.begen-avtcore-rtp-duplication], the sender must also use the same RTCP CNAME for both the main and redundant streams, and must include an "a=ssrc:... srcname:..." attribute to correlate the flows.

3.2. DUP Grouping for SSRC-Multiplexed RTP Streams

[RFC5576] defines an SDP media-level attribute, called 'ssrc-group', for grouping the RTP streams that are SSRC multiplexed and carried in the same RTP session. The grouping is based on the SSRC identifiers. Since SSRC-multiplexed RTP streams are defined in the same "m" line, the 'group' attribute cannot be used.

This section specifies how duplication is used with SSRC-multiplexed streams using the 'ssrc-group' attribute [RFC5576].

The semantics of "DUP" for the 'ssrc-group' attribute are the same as the one defined for the 'group' attribute except that the SSRC identifiers are used to designate the duplication grouping associations: a=ssrc-group:DUP *(SP ssrc-id) [RFC5576].

3.3. SDP Offer/Answer Model Considerations

When offering duplication grouping using SDP in an Offer/Answer model [RFC3264], the following considerations apply.

A node that is receiving an offer from a sender may or may not understand line grouping. It is also possible that the node understands line grouping but it does not understand the "DUP" semantics. From the viewpoint of the sender of the offer, these cases are indistinguishable.

When a node is offered a session with the "DUP" grouping semantics but it does not support line grouping or the duplication grouping semantics, as per [RFC5888], the node responds to the offer either (1) with an answer that ignores the grouping attribute or (2) with a refusal to the request (e.g., 488 Not Acceptable Here or 606 Not Acceptable in SIP).

In the first case, the original sender of the offer must send a new offer without any duplication grouping. In the second case, if the

sender of the offer still wishes to establish the session, it should retry the request with an offer without the duplication grouping. This behavior is specified in [RFC5888].

4. SDP Examples

4.1. Separate Source Addresses

In this example, the redundant streams use the same IP destination address (232.252.0.1) but they are sourced from different addresses (198.51.100.1 and 198.51.100.2). Thus, the receiving host needs to join both SSM sessions separately.

```
v=0
o=ali 1122334455 1122334466 IN IP4 dup.example.com
s=DUP Grouping Semantics
t=0 0
m=video 30000 RTP/AVP 100
c=IN IP4 232.252.0.1/127
a=source-filter:incl IN IP4 232.252.0.1 198.51.100.1 198.51.100.2
a=rtpmap:100 MP2T/90000
a=ssrc:1000 cname:chl@example.com
a=ssrc:1010 cname:chl@example.com
a=ssrc-group:DUP 1000 1010
a=mid:Group1
```

Note that in actual use, SSRC values, which are random 32-bit numbers, can be much larger than the ones shown in this example.

4.2. Separate Destination Addresses

In this example, the redundant streams have different IP destination addresses. The example shows the same UDP port number and IP source addresses, but either or both could have been different for the two streams.


```
v=0
o=ali 1122334455 1122334466 IN IP4 dup.example.com
s=DUP Grouping Semantics
t=0 0
a=group:DUP S1a S1b
m=video 30000 RTP/AVP 100
c=IN IP4 233.252.0.1/127
a=source-filter:incl IN IP4 233.252.0.1 198.51.100.1
a=rtpmap:100 MP2T/90000
a=ssrc:1000 cname:chl@example.com
a=ssrc:1000 srcname:45:a8:f4:19:b4:c3
a=mid:S1a
m=video 30000 RTP/AVP 101
c=IN IP4 233.252.0.2/127
a=source-filter:incl IN IP4 233.252.0.2 198.51.100.1
a=rtpmap:101 MP2T/90000
a=ssrc:1010 cname:chl@example.com
a=ssrc:1010 srcname:45:a8:f4:19:b4:c3
a=mid:S1b
```

4.3. Temporal Redundancy

In this example, the redundant streams have the same IP source and destination addresses but different UDP port numbers. Due to the same source and destination addresses, the packets in both streams will be routed over the same path. To provide resiliency against packet loss, the duplicate of an original packet is transmitted 50 ms later as indicated by the 'duplication-delay' attribute (defined in [I-D.begen-mmusic-temporal-interleaving]).

```
v=0
o=ali 1122334455 1122334466 IN IP4 dup.example.com
s=DUP Grouping Semantics
t=0 0
a=group:DUP S1a S1b
a=duplication-delay:50
m=video 30000 RTP/AVP 100
c=IN IP4 233.252.0.1/127
a=source-filter:incl IN IP4 233.252.0.1 198.51.100.1
a=rtpmap:100 MP2T/90000
a=ssrc:1000 cname:chl@example.com
a=ssrc:1000 srcname:45:a8:f4:19:b4:c3
a=mid:S1a
m=video 40000 RTP/AVP 101
c=IN IP4 233.252.0.1/127
a=source-filter:incl IN IP4 233.252.0.1 198.51.100.1
a=rtpmap:101 MP2T/90000
a=ssrc:1010 cname:chl@example.com
a=ssrc:1010 srcname:45:a8:f4:19:b4:c3
a=mid:S1b
```

5. Security Considerations

There is a weak threat for the receiver that the duplication grouping can be modified to indicate relationships that do not exist. Such attacks might result in failure of the duplication mechanisms, and/or mishandling of the media streams by the receivers.

In order to avoid attacks of this sort, the SDP description needs to be integrity protected and provided with source authentication. This can, for example, be achieved on an end-to-end basis using S/MIME [RFC5652] [RFC5751] when the SDP is used in a signaling packet using MIME types (application/sdp). Alternatively, HTTPS [RFC2818] or the authentication method in the Session Announcement Protocol (SAP) [RFC2974] could be used as well.

6. IANA Considerations

This document registers the following semantics with IANA in Semantics for the 'group' SDP Attribute under SDP Parameters:

Note to the RFC Editor: In the following registrations, please replace "XXXX" with the number of this document prior to publication as an RFC.

Semantics	Token	Reference
-----	-----	-----
Duplication	DUP	[RFCXXXX]

This document also registers the following semantics with IANA in Semantics for the 'ssrc-group' SDP Attribute under SDP Parameters:

Token	Semantics	Reference
-----	-----	-----
DUP	Duplication	[RFCXXXX]

7. Acknowledgments

The authors would like to thank Colin Perkins, Bill Ver Steeg, Dave Oran and Toerless Eckert for their inputs and suggestions.

8. References

8.1. 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.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC5576] Lennox, J., Ott, J., and T. Schierl, "Source-Specific Media Attributes in the Session Description Protocol (SDP)", RFC 5576, June 2009.
- [RFC5888] Camarillo, G. and H. Schulzrinne, "The Session Description Protocol (SDP) Grouping Framework", RFC 5888, June 2010.

8.2. Informative References

- [I-D.begen-avtcore-rtp-duplication]
Begen, A. and C. Perkins, "Duplicating RTP Streams",

draft-begen-avtcore-rtp-duplication-01 (work in progress),
March 2012.

- [I-D.begen-mmusic-temporal-interleaving]
Begen, A., Cai, Y., and H. Ou, "Delayed Duplication
Attribute in the Session Description Protocol",
draft-begen-mmusic-temporal-interleaving-04 (work in
progress), March 2012.
- [IC2011] Evans, J., Begen, A., Greengrass, J., and C. Filsfils,
"Toward Lossless Video Transport (to appear in IEEE
Internet Computing)", November 2011.
- [RFC2354] Perkins, C. and O. Hodson, "Options for Repair of
Streaming Media", RFC 2354, June 1998.
- [RFC2818] Rescorla, E., "HTTP Over TLS", RFC 2818, May 2000.
- [RFC2974] Handley, M., Perkins, C., and E. Whelan, "Session
Announcement Protocol", RFC 2974, October 2000.
- [RFC5652] Housley, R., "Cryptographic Message Syntax (CMS)", STD 70,
RFC 5652, September 2009.
- [RFC5751] Ramsdell, B. and S. Turner, "Secure/Multipurpose Internet
Mail Extensions (S/MIME) Version 3.2 Message
Specification", RFC 5751, January 2010.

Authors' Addresses

Ali Begen
Cisco
181 Bay Street
Toronto, ON M5J 2T3
Canada

Email: abegen@cisco.com

Yiqun Cai
Cisco
170 W. Tasman Dr.
San Jose, CA 95134
USA

Email: ycai@cisco.com

Heidi Ou
Cisco
170 W. Tasman Dr.
San Jose, CA 95134
USA

Email: hou@cisco.com

MMUSIC
Internet-Draft
Intended status: Standards Track
Expires: September 12, 2012

A. Begen
Y. Cai
H. Ou
Cisco
March 11, 2012

Delayed Duplication Attribute in the Session Description Protocol
draft-begen-mmusic-temporal-interleaving-04

Abstract

A straightforward approach to provide protection against packet losses due to network outages with a longest duration of T time units is to simply duplicate the original packets and send each copy separated in time by at least T time units. This approach is commonly referred to as Time-shifted Redundancy, Temporal Redundancy or simply Delayed Duplication. This document defines an attribute to indicate the presence of temporally redundant media streams and the duplication delay in the Session Description Protocol.

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 12, 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. Requirements Notation	4
3. The 'duplication-delay' Attribute	4
4. SDP Examples	4
5. Security Considerations	6
6. IANA Considerations	7
6.1. Registration of SDP Attributes	7
7. Acknowledgements	7
8. References	7
8.1. Normative References	7
8.2. Informative References	8
Authors' Addresses	9

1. Introduction

Consider that a media sender transmits an original source packet and transmits its duplicate after a certain delay following the original transmission. If a network outage hits the original transmission, the expectation is that the second transmission arrives at the receiver. Alternatively, the second transmission may be hit by an outage and gets dropped, and the original transmission completes successfully. On the receiver side, both transmissions can also arrive and in that case, the receiver (or the node that does the duplicate suppression) needs to identify the duplicate packets and discard them appropriately, producing a duplicate-free stream.

Delayed duplication can be used in a variety of multimedia applications where there is sufficient bandwidth for the duplicated traffic and the application can tolerate the introduced delay. However, it must be used with care since it might easily result in a new series of denial-of-service attacks. Furthermore, delayed duplication must not be used in cases where the primary cause of packet loss is congestion, rather than a network outage due to a temporary link or network element failure. Duplication can make congestion only worse.

One particular use case for delayed duplication is to improve the reliability of real-time video feeds inside a core IP network [IC2011]. Compared to other popular redundancy approaches such as Forward Error Correction (FEC) [RFC6363] and redundant data encoding (e.g., [RFC2198]), delayed duplication is quite easy to implement since it does not require any special type of encoding or decoding.

For duplicate suppression, the receiver has to be able to identify the identical packets. This is straightforward for media packets that carry one or more unique identifiers such as the sequence number field in RTP header [RFC3550]. In non-RTP applications, the receiver can use unique sequence numbers if available or other alternative approaches to compare the incoming packets and discard the duplicate ones.

In this specification, we are not concerned about how the sender should determine the duplication delay. We are not concerned about how the receiver can suppress the duplicate packets and merge the incoming streams to produce a hopefully loss-free and duplication-free output stream (called stream merging), either. These considerations are out of the scope for this specification. Rather, our goal is simply to introduce a new attribute for the Session Description Protocol (SDP) [RFC4566] that indicates that the media stream is to be duplicated and sent two or more times, and also indicates the relative delay for each additional duplication.

In practice, more than two redundant streams are unlikely to be used since the additional delay and increased overhead are not easily justified. However, we define the new attribute in a general way so that it could be used with more than two redundant streams if needed. While the primary focus in this specification is the RTP-based transport, the new attribute is applicable to both RTP and non-RTP streams. Details on duplicating RTP streams are presented in [I-D.begen-avtcore-rtp-duplication].

2. Requirements Notation

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].

3. The 'duplication-delay' Attribute

The following ABNF [RFC5234] syntax formally describes the 'duplication-delay' attribute:

```
delaying-attribute    = "a=duplication-delay:" periods CRLF
periods               = period *( ":" period)
period                = 1*DIGIT ; in milliseconds
```

Figure 1: ABNF syntax for the 'interleaving-period' attribute

The 'duplication-delay' attribute is defined as both a media-level and session-level attribute. It specifies the relative delay for each duplication in milliseconds (ms). If used as a media-level attribute, it MUST be used with the 'ssrc-group' attribute and "DUP" grouping semantics as defined in [I-D.begen-mmusic-redundancy-grouping]. If used as a session-level attribute, it MUST be used with 'group' attribute and "DUP" grouping semantics as defined in [I-D.begen-mmusic-redundancy-grouping].

4. SDP Examples

In the first example below, the multicast stream is duplicated with a duplication delay of 100 ms. The streams have Synchronization Sources (SSRC) of 1000 and 1010, and they are grouped together using the 'ssrc-group' attribute defined in [RFC5576]. The "DUP" grouping semantics are defined in [I-D.begen-mmusic-redundancy-grouping]. The reason for using explicit grouping is that not all the media streams

in the same "m" line are necessarily duplicates of each other.

```
v=0
o=ali 1122334455 1122334466 IN IP4 dup.example.com
s=Delayed Duplication
t=0 0
m=video 30000 RTP/AVP 100
c=IN IP4 233.252.0.1/127
a=source-filter:incl IN IP4 233.252.0.1 198.51.100.1
a=rtpmap:100 MP2T/90000
a=ssrc:1000 cname:chl@example.com
a=ssrc:1010 cname:chl@example.com
a=ssrc-group:DUP 1000 1010
a=duplication-delay:100
a=mid:Group1
```

Note that in actual use, SSRC values, which are random 32-bit numbers, could be much larger than the ones shown in this example.

In the second example below, the multicast stream is duplicated twice. 50 ms after the original transmission, the first duplicate is transmitted and 100 ms after that, the second duplicate is transmitted. In other words, the same packet is transmitted three times over a period of 150 ms.

```
v=0
o=ali 1122334455 1122334466 IN IP4 dup.example.com
s=Delayed Duplication
t=0 0
m=video 30000 RTP/AVP 100
c=IN IP4 233.252.0.1/127
a=source-filter:incl IN IP4 233.252.0.1 198.51.100.1
a=rtpmap:100 MP2T/90000
a=ssrc:1000 cname:chl@example.com
a=ssrc:1010 cname:chl@example.com
a=ssrc:1020 cname:chl@example.com
a=ssrc-group:DUP 1000 1010 1020
a=duplication-delay:50:100
a=mid:Group1
```

In the third example below, the multicast UDP stream is duplicated with a duplication delay of 50 ms. Both streams are sent in the same source-specific multicast (SSM) session but they are sent to different ports. The "DUP" grouping semantics

[I-D.begen-mmusic-redundancy-grouping] are used to describe the redundancy relation.

```
v=0
o=ali 1122334455 1122334466 IN IP4 dup.example.com
s=Delayed Duplication
t=0 0
a=group:DUP S1a S1b
a=duplication-delay:50
m=audio 30000 udp mp4
c=IN IP4 233.252.0.1/127
a=source-filter:incl IN IP4 233.252.0.1 198.51.100.1
a=mid:S1a
m=audio 40000 udp mp4
c=IN IP4 233.252.0.2/127
a=source-filter:incl IN IP4 233.252.0.1 198.51.100.1
a=mid:S1b
```

5. Security Considerations

The 'duplication-delay' attribute is not believed to introduce any significant security risk to multimedia applications. A malevolent third party could use this attribute to misguide the receiver(s) about the duplication delays and/or the number of redundant streams. For example, if the malevolent third party increases the value of the duplication delay, the receiver(s) will unnecessarily incur a longer delay since they will have to wait for the entire period. Or, if the duplication delay is reduced by the malevolent third party, the receiver(s) might not wait long enough for the duplicated transmission and incur unnecessary packet losses. However, these require intercepting and rewriting the packets carrying the SDP description; and if an interceptor can do that, many more attacks are also possible.

In order to avoid attacks of this sort, the SDP description needs to be integrity protected and provided with source authentication. This can, for example, be achieved on an end-to-end basis using S/MIME [RFC5652] [RFC5751] when SDP is used in a signaling packet using MIME types (application/sdp). Alternatively, HTTPS [RFC2818] or the authentication method in the Session Announcement Protocol (SAP) [RFC2974] could be used as well.

Another security risk is due to possible software misconfiguration or a software bug where a large number of duplicates could be unwillingly signaled in the 'duplication-delay' attribute. In applications where this attribute is to be used, it is a good

practice to put a hard limit both on the number of duplicate streams and the total delay introduced due to duplication regardless of what the SDP description specifies.

6. IANA Considerations

The following contact information shall be used for all registrations in this document:

Ali Begen
abegen@cisco.com

Note to the RFC Editor: In the following, replace "XXXX" with the number of this document prior to publication as an RFC.

6.1. Registration of SDP Attributes

This document registers a new attribute name in SDP.

SDP Attribute ("att-field"):	
Attribute name:	duplication-delay
Long form:	Duplication delay for temporally redundant streams
Type of name:	att-field
Type of attribute:	Media or session level
Subject to charset:	No
Purpose:	Specifies the relative duplication delay(s) for redundant stream(s)
Reference:	[RFCXXXX]
Values:	See [RFCXXXX]

7. Acknowledgements

Authors would like to thank Colin Perkins for his suggestions and review.

8. References

8.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.
- [RFC5576] Lennox, J., Ott, J., and T. Schierl, "Source-Specific Media Attributes in the Session Description Protocol (SDP)", RFC 5576, June 2009.
- [I-D.begen-mmusic-redundancy-grouping]
Begen, A., Cai, Y., and H. Ou, "Duplication Grouping Semantics in the Session Description Protocol", draft-begen-mmusic-redundancy-grouping-02 (work in progress), October 2011.

8.2. Informative References

- [RFC6363] Watson, M., Begen, A., and V. Roca, "Forward Error Correction (FEC) Framework", RFC 6363, October 2011.
- [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.
- [I-D.begen-avtcore-rtp-duplication]
Begen, A. and C. Perkins, "Duplicating RTP Streams", draft-begen-avtcore-rtp-duplication-00 (work in progress), October 2011.
- [IC2011] Evans, J., Begen, A., Greengrass, J., and C. Filsfils, "Toward Lossless Video Transport (to appear in IEEE Internet Computing)", November 2011.
- [RFC5652] Housley, R., "Cryptographic Message Syntax (CMS)", STD 70, RFC 5652, September 2009.
- [RFC5751] Ramsdell, B. and S. Turner, "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Message Specification", RFC 5751, January 2010.
- [RFC2818] Rescorla, E., "HTTP Over TLS", RFC 2818, May 2000.

[RFC2974] Handley, M., Perkins, C., and E. Whelan, "Session
Announcement Protocol", RFC 2974, October 2000.

Authors' Addresses

Ali Begen
Cisco
181 Bay Street
Toronto, ON M5J 2T3
Canada

Email: abegen@cisco.com

Yiqun Cai
Cisco
170 W. Tasman Dr.
San Jose, CA 95134
USA

Email: ycai@cisco.com

Heidi Ou
Cisco
170 W. Tasman Dr.
San Jose, CA 95134
USA

Email: hou@cisco.com

mmusic
Internet-Draft
Intended status: Standards Track
Expires: November 1, 2012

P. Capelastegui
Universidad Politecnica de
Madrid
April 30, 2012

3D Video in the Session Description Protocol (SDP)
draft-capelastegui-mmusic-3dv-sdp-00

Abstract

This document defines a mechanism to describe 3D video streams in the Session Description Protocol (SDP). This includes 3D video streams composed of multiple video views, or of a combination of views and depth maps. Several 3D video formats are supported, including simulcast, video-plus-depth, and frame-packing.

A new decoding dependency, "3dd", is defined, describing the association between media stream belonging to a 3D video stream. In addition, a new SDP media-level attribute, "3dvFormat", is defined, describing the format used by media streams within a 3D video stream.

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 November 1, 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
1.1. Requirements Language	3
2. Definitions	3
3. Decoding dependency of 3D video streams	4
4. The "3dvFormat" media attribute	5
4.1. The "depth-map-simulcast" 3D format attribute	5
4.2. The "depth-map-metadata" 3D format attribute	6
4.3. The "stereo-view" 3D format attribute	7
4.4. The "frame-pack" 3D format attribute	8
5. Usage with SDP offer/answer model	9
5.1. Backward compatibility	10
6. Examples	10
6.1. Example session with single 3D video option	10
6.2. Test Scenario: Multiple 3D options	11
7. Formal Grammar	13
7.1. "3dvFormat" media attribute	13
7.2. "depth-map-simulcast" 3D format attribute	13
7.3. "depth-map-metadata" 3D format attribute	13
7.4. "stereo-view" 3D format attribute	13
7.5. "frame-pack" 3D format attribute	13
8. Security Considerations	13
9. IANA Considerations	13
10. Normative References	14
Author's Address	15

1. Introduction

3D video applications convey depth information by showing a different view for each eye of a user. In order to achieve this, 3D video streams need to include additional information compared to conventional 2D video streams, either in the form of extra views, or auxiliary maps such as depth maps, or a combination thereof. These views and maps can be transported in a variety of ways, including, among others: as separate RTP streams (simulcast), frame-packed in a single video stream [HDMIv1.4a], or using the video-plus-depth format [ISO/IEC 23002-3].

The Session Description Protocol (SDP) [RFC4566] lacks the means to describe neither of these transport techniques for 3D video. This document extends SDP to support the description of multimedia sessions using 3D video encapsulated as simulcast streams, using frame-packing techniques, or using the video-plus-depth format.

[RFC5583] defines a mechanism to signal the decoding dependency of media descriptions in SDP. This document extends that mechanism by defining a new SDP decoding dependency type, '3dd', describing the association between media streams belonging to a 3D video stream. In addition, a new SDP media-level attribute, '3dvFormat', is defined to describe the format used by media streams composing a 3D video stream. Several formats for 3D video are described in this specification, including simulcast stereo video, simulcast video and depth map, various frame-packing schemes, and streams using video-plus-depth.

1.1. 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. Definitions

3D video stream: A video stream that conveys depth information by showing different perspectives of a scene to each eye of an observing user. A 3D video stream is typically composed of multiple video streams ('views'), or a combination of video streams and auxiliary data maps such as depth maps

2D video stream: A video stream that lacks 3D depth information.
View: A video stream that represents a specific point of view of the scene in a 3D video stream.
Depth Map: An auxiliary data stream that associates a Z-value to each pixel of a view within a 3D video stream. Depth maps are often encoded as grey scale video streams.
Simulcast: A method for the transmission of 3D video streams that consists on sending additional views and auxiliary data streams as a separate RTP streams.
Video-plus-depth: (also known as MPEG-C Part 3) A method for the transmission of 3D video streams consisting on encapsulating a depth map as metadata within a 2D video stream. This mechanism, standardized in [ISO/IEC 23002-3], is compatible with MPEG-2 and H.264/AVC, and allows for backwards compatibility.
Frame-packing: A method for the transmission of 3D video streams that consists on multiplexing several views and/or auxiliary data within a single video stream, using either spatial multiplexing or time multiplexing. Frame-packing is supported by standards like [HDMIv1.4a] and [ITU-T H.264].

3. Decoding dependency of 3D video streams

The "depend" SDP attribute, defined in [RFC5583] describes the decoding dependency between two or more media descriptions. This specification defines a new dependency type for the "depend" attribute:

- o 3dd: 3D video dependency - indicates that the described media stream belongs to a 3D video stream, and requires other media streams to render the 3D video. When "3dd" is used, all required media streams for the Operation Point MUST be identified by identification-tag and fmt-dependency following the "3dd" string.

Like other dependency types, 3dd is used in combination with the "DDP" grouping semantic, which is defined in [RFC5583], and based on the SDP grouping framework [RFC5888]. Whenever a 3D video stream is composed of multiple media descriptions, these media descriptions MUST be included in the same DDP group.

The media decoding dependency terminology defined in [RFC5583] can be applied to 3D video streams as follows:

- o Media Bitstream: A 3D video stream is considered a Media Bitstream for the purposes of 3dd decoding dependency.
- o Media Partition: Each separate media description composing a 3D video stream is considered a Media Partition. Note that each Media Partition usually contains a single video view or depth map, but can also include multiple of views/maps, e.g. when using frame-packing techniques.
- o Operation Point: A subset of a 3D Media Bitstream that includes all Media Partitions required for reconstruction at a certain point of quality, number of views or depth maps, or other property. Note that a valid Operation Point for a 3D Media Bitstream can be a 2D video lacking any depth information.

4. The "3dvFormat" media attribute

a=3dvFormat:<fmt> <attribute>:<value>

This section defines a new media-level attribute for SDP, "3dvFormat", which can be used to describe the transport format of a media stream in a 3D video stream. This attribute can indicate that a media description corresponds to a specific view within a 3D stream, or to a depth map, or to a combination of views or depth maps encapsulated with frame-packing techniques or with the video-plus-depth mechanism.

A media description can have multiple "3dvFormat" attributes; each attribute is mapped to a media format specified for the media, indicated by <fmt>. Only one "3dvFormat" attribute is allowed per media format.

Each "3dvFormat" attribute indicates a property (known as a "3D format attribute") associated to a media format of its media description. The 3D format attribute consists on an attribute-value pair, with the form "<attribute>:<value>". This specification defines four 3D format attributes: "depth-map-simulcast", "depth-map-metadata", "stereo-view", "and frame-pack".

New 3D format attributes can be defined, but they MUST be registered with IANA, following the registry described in Section 9.

4.1. The "depth-map-simulcast" 3D format attribute

a=3dvFormat:<fmt> depth-map-simulcast:<associated_video>

The 3D format attribute "depth-map-simulcast" indicates that a media stream represents a depth map associated with a view within the same 3D video stream. A depth map described by this attribute is

transmitted as a separate transport stream from its corresponding view.

<associated-video> is the media stream identification (the "a=mid" attribute, as defined in [RFC5888]) of the video stream associated with this depth map.

A media description with the "depth-map-simulcast" 3D format attribute MUST be included in a DDP group. This group MUST include a video stream representing the view associated with the depth map. Finally, the depth map media description MUST include a "depend" attribute with the "3dd" dependency type, indicating dependency to one or more media formats within that video stream.

Example:

```
a=group:DDP 1 2
m=video 1111 RTP/AVP 99
a=rtpmap:99 H264/90000
a=mid:1
m=video 1112 RTP/AVP 99
a=rtpmap:99 H264/90000
a=3dvFormat:99 depth-map-simulcast:1
a=mid:2
a=depend:99 3dd 1:99
```

The example shows two media descriptions forming a 3D video stream, of which the first one (mid:1) represents a video view, and the second one (mid:2) the depth map for that view. The depth map cannot be used without its corresponding view, and this is reflected in the "depend" attribute.

4.2. The "depth-map-metadata" 3D format attribute

```
a=3dvFormat:<fmt> depth-map-metadata:<associated_video>
```

The 3D format attribute "depth-map-metadata" indicates that a media stream represents a depth map associated with a view within the same 3D video stream. A depth map described by this attribute is transmitted as part of the same transport stream as its corresponding view, in the form of metadata. If the view associated with this depth map is a MPEG-2 or H.264/AVC video stream, the depth map follows the format defined in MPEG-C part 3 [ISO/IEC 23002-3].

<associated-video> is the media stream identification (the "a=mid" attribute, as defined in [RFC5888]) of the video stream associated with this depth map.

A media description with the "depth-map-simulcast" 3D format attribute MUST be included in a DDP group. This group MUST include a video stream representing the view associated with the depth map. Finally, the depth map media description MUST include a "depend" attribute with the "3dd" dependency type, indicating dependency to that video stream.

It is important to note that, when a media format with a "depth-map-metadata" is used, the transport information for that media stream such as port, connection address or transport protocol MUST be ignored. In this case, the depth map is transmitted as part of the media stream of its associated view, rather than as a separate stream.

Example:

```
a=group:DDP 1 2
m=video 1111 RTP/AVP 99
a=rtpmap:99 H264/90000
a=mid:1
m=video 1112 RTP/AVP 99 100
a=rtpmap:99 H264/90000
a=3dvFormat:99 depth-map-simulcast:1
a=rtpmap:100 H264/90000
a=3dvFormat:100 depth-map-metadata:1
a=mid:2
a=depend:99 3dd 1:99; 100 3dd 1:99
```

The example shows two media descriptions forming a 3D video stream, of which the first one (mid:1) represents a video view, and the second one (mid:2) the depth map for that view. Two possible configurations for the depth map are offered, one using simulcast (payload type 99), and the other transmitting the depth map as metadata (payload type 100). If the depth map stream is configured as metadata, the port specified in that media description (1112) will be ignored, since the depth map will be transmitted within the video view stream. On the other hand, if the simulcast option is used, the depth map will be transmitted as a separate stream using the specified port and transport, as usual.

4.3. The "stereo-view" 3D format attribute

```
a=3dvFormat:<fmt> stereo-view:<view-type>
```

The 3D format attribute "stereo-view" indicates whether a video stream is associated with the left-eye view or the right-eye view of a stereo 3D video stream.

<view-type> indicates which view is associated with the media stream. It can have the value "left", for the left-eye view, or "right", for the right-eye view.

A media description with the "stereo-view" 3D format attribute MUST be included in a DDP group. This group MUST also include another video stream containing the "stereo-view" 3D format attribute with the other stereo view as value. The media description for either of the two stereo views MUST include a "depend" attribute with the "3dd" dependency type, indicating dependency to the stream corresponding to the other view.

Example:

```
a=group: DDP 1 2
m=video 1111 RTP/AVP 99
a=rtpmap:99 H264/90000
a=3dvFormat:99 stereo-view:left
a=mid:1
m=video 1112 RTP/AVP 99
a=rtpmap:99 H264/90000
a=3dvFormat:99 stereo-view:right
a=mid:2
a=depend:99 3dd 1:99
```

The example shows two media descriptions forming a stereo 3D video stream, of which the first one (mid:1) represents the left view, and the second one (mid:2) the right view. This Media Bitstream can be configured as a 3D video stream composed of two stereo views, or as a 2D video stream including just the left eye view.

4.4. The "frame-pack" 3D format attribute

```
a=3dvFormat:<fmt> frame-pack:<fp-format>
```

The 3d attribute indicates that frame-packing mechanisms are used in a media stream, for the specified media format.

<fp-format> signals which frame-packing mode is applied. It has three possible values: "side-by-side", "top-bottom", and "frame-seq".

Of these frame-pack modes, the first two are based on spatial multiplexing, or dividing each video frame in the stream into two sub-frames, and assigning one view to each sub-frame. In "side-by-side" mode, the left sub-frame corresponds to the left eye view, and the right sub-frame to the right eye view. In "top-bottom" mode, the top sub-frame corresponds to the left eye view, and the lower sub-frame to the right eye view.

On the "frame-seq" (frame sequential) frame-packing mode, time multiplexing is used, so that half the video frames in a stream correspond to the left eye view, and the other half to the right eye view, in alternating order. In order to identify which frame corresponds to each view, additional signalling is required; in H.264/AVC video streams, this is achieved through supplemental enhancement information (SEI) metadata [ITU-T H.264].

5. Usage with SDP offer/answer model

When the extensions defined in this specification are used in the SDP offer/answer model [RFC3264], the following rules apply.

The offerer MAY include more than one "3dvFormat" attribute per media description, and the values of these "3dvFormat" can be different or duplicated. However, each media format MUST NOT have more than one "3dvFormat" attribute.

If the offerer includes a 3D video stream composed of more than one media description, all media descriptions in the stream MUST be included in a DDP group. If the 3D video stream includes streams with 3D format attributes whose description specifies any stream requirements or mandatory dependencies, those requirements or dependencies MUST be respected. Each 3D video stream in the offer SHOULD have at least one Operation Point consisting on a single 2D video stream, as well as any number of Operation Points with 3D video.

An answer MUST NOT include any "3dvFormat" attribute that is not present in the offer.

When a media format in an offered media description has a "3dvFormat" attribute, if the answer contains that media format it MUST also include the "3dvFormat" attribute, with the same parameters as the offer.

To simplify the processing of 3D video configurations, when the answer includes a "3dvFormat" attribute in a media description, the same RTP payload type number used in the offer should also be used in the answer, and the answer MUST NOT include more than one media format for that media description.

If the answerer understands the DDP semantics, it is necessary to take the "depend" attribute into consideration in the Offer/Answer procedure, as indicated in [RFC5583]

5.1. Backward compatibility

Depending on implementation, a node that does not understand DDP grouping or "3d" attributes SHOULD respond to an offer using this grouping or attributes either with a refusal to the request, or with an answer that ignores the grouping or 3D video format attributes.

In case of a refused request, if the offerer has identified that the refusal of the request is caused by the use of 3D video, and it still wishes to initiate a session, it SHOULD generate a new offer without any 3D video streams.

If the request is accepted but the answer is ignoring the grouping attribute, the "depend" attribute, or a "3dvFormat", it should be assumed that the answerer is unable to send or receive 3D video streams. If the offerer still wishes to initiate a session, it SHOULD generate a new offer without any 3D video streams. Alternatively, if the answer does not include more than a single video stream, the offerer MAY initiate the session without generating a new offer, and send and receive that stream as a 2D video stream.

6. Examples

The following examples show SDP Offer/Answer exchanges for sessions with 3D video streams. Only the media descriptions and grouping attributes of the SDP are shown. For each example, two possible answers are considered: one in which the answering device is compatible with this specification, and one with a legacy answering device.

6.1. Example session with single 3D video option

The example shows a session where the 3D video stream is transmitted over a single media stream, so no grouping or decoding dependencies are needed for the SDP. The calling user agent makes a SDP offer with 2 options for configuring the 3D video stream:

- o 2D video stream
- o Single frame-packed video stream, with 2 views multiplexed side-by-side

Offer SDP:

```
m=video 1111 RTP/AVP 99 100
a=rtpmap:99 H264/90000
```

```
a=rtpmap:100 H264/90000
a=3dvFormat:100 frame-pack:side-by-side
```

Answer SDP:

```
m=video 2222 RTP/AVP 100
a=rtpmap:100 H264/90000
a=3dvFormat:100 frame-pack:side-by-side
```

The initial offer includes a media description with two media formats, with one corresponding to a 2D video stream (payload type 99) and the other to a frame-packed 3D video stream (payload type 100). Of these, the answering device chooses the frame-packed media format.

Alternate Answer SDP (legacy device):

```
m=video 2222 RTP/AVP 100
a=rtpmap:100 H264/90000
```

If this SDP offer is received by a legacy device and the session is not rejected, the answer will ignore any 3D video format attributes. In this case, the offerer can initiate the session treating the selected media format as a 2D video stream.

6.2. Test Scenario: Multiple 3D options

The example shows a session where the 3D video stream is transmitted over up to two media streams, and several options for the format of the 3D video stream are offered:

- o 2D video stream
- o Single frame-packed video stream, with 2 views multiplexed side-by-side
- o Single video stream including a depth map as metadata
- o 2 Simulcast streams, with video and depth map
- o 2 Simulcast streams, with 2 stereo views.

Offer SDP:

```
a=group:DDP 1 2
m=video 1111 RTP/AVP 99 100
a=rtpmap:99 H264/90000
a=3dvFormat:99 stereo-view:left
a=rtpmap:100 H264/90000
a=3dvFormat:100 frame-pack:side-by-side
a=mid:1
```

```
m=video 1112 RTP/AVP 99 100 101
a=rtpmap:99 H264/90000
a=3dvFormat:99 depth-map-metadata:1
a=rtpmap:100 H264/90000
a=3dvFormat:100 depth-map-simulcast:1
a=rtpmap:101 H264/90000
a=3dvFormat:101 stereo-view:right
a=mid:2
a=depend:99 3dd 1:99; 100 3dd 1:99; 101 3dd 1:99
```

Answer SDP:

```
a=group:DDP 1 2
m=video 2222 RTP/AVP 99
a=rtpmap:99 H264/90000
a=3d:99 stereo-view:left
a=mid:1
m=video 2223 RTP/AVP 102
a=rtpmap:101 H264/90000
a=3d:101 stereo-view:right
a=mid:2
a=depend:101 3dd 1:99
```

The initial offer includes two media descriptions, the first of which (mid 1) can be transmitted independently, either as a 2D video stream (payload type 99) or as a frame-packed 3D stream (payload type 100). The second media description (mid 2), on the other hand, depends on the first one for all its media formats, and can be configured as a depth map transmitted as metadata (payload type 99), as a simulcast depth map stream (payload type 100), or as a right-eye stereo view (payload-type 101). The answering device chooses the configuration with 2 simulcast stereo views.

Alternate Answer SDP (legacy device)

```
m=video 2222 RTP/AVP 99
a=rtpmap:99 H264/90000
m=video 0 RTP/AVP 102
a=rtpmap:101 H264/90000
```

If this SDP offer is received by a legacy device and the session is not rejected, the answer will ignore any 3D video format attributes, as well as the grouping and dependency attributes. In the example above, the answering device has selected a media format for the first video stream, and disabled the second video stream. In this case, the offerer can initiate the session treating the selected media format as a 2D video stream. If the second video stream had not been disabled, the offerer should send a new offer with a single video

stream.

7. Formal Grammar

The 3d attributes defined in this document use the following Augmented Backus-Naur Form (ABNF) [RFC5234] grammar.

7.1. "3dvFormat media attribute

```
3dvformat-attribute = "a=3dvFormat:" fmt SP 3dvf-type
; fmt is described in [RFC4566]
; fmt is media format (usually RTP payload type)

3dvf-type= depth-scast / depth-meta / st-view / f-pack
```

7.2. "depth-map-simulcast" 3D format attribute

```
depth-scast = "depth-map-simulcast:" identification-tag
; identification-tag is defined in [RFC5888]
```

7.3. "depth-map-metadata" 3D format attribute

```
depth-meta = "depth-map-metadata:" identification-tag
; identification-tag is defined in [RFC5888]
```

7.4. "stereo-view" 3D format attribute

```
st-view = "stereo-view:" view-type
view-type= "left" / "right"
```

7.5. "frame-pack" 3D format attribute

```
f-pack = "frame-pack:" fp-format
fp-format= "side-by-side" / "top-bottom" / "frame-seq"
```

8. Security Considerations

No security issues have been identified for this specification.

9. IANA Considerations

The following contact information shall be used for all registrations included here:

Contact: Pedro Capelastegui
 email: Capelastegui@dit.upm.es
 tel: +34 915 49 57 00 ext. 3024

This document defines the following new semantics for the "depend" SDP attribute. The semantics are registered by IANA under "depend" SDP Attribute Values under "Session Description Protocol (SDP) Parameters":

Token	Semantics	Reference
3dd	3D video dependency	[THIS DOC]

This document defines a new media-level SDP attribute, "3dvFormat". The attribute is registered by IANA under "Session Description Protocol (SDP) Parameters" under "att-field (media level only)".

Attribute name: 3dvFormat
 Long form: 3D video format
 Type of name: att-field
 Type of attribute: media level only
 Subject to charset: no
 Purpose: [THIS DOCUMENT]
 Reference: [THIS DOCUMENT]
 Values: see this document and registrations below

Parameters of the "3dvFormat" SDP attribute MUST be registered under IANA following the "Specification Required" policy [RFC5226]. This document creates a new IANA registry called [REF-1] within the "Session Description Protocol (SDP) Parameters" registry, for that purpose.

The initial entries in the registry are shown below.

Token	Description	Reference
depth-map-simulcast	depth map as separate stream	[THIS DOC]
depth-map-metadata	depth map as metadata	[THIS DOC]
stereo-view	left or right stereo view	[THIS DOC]
frame-pack	frame-packed video stream	[THIS DOC]

10. Normative References

[HDMIv1.4a]

HDMI, "HDMI Specification Version 1.4a", March 2010.

- [ISO/IEC 23002-3]
ISO/IEC JTC1/SC29/WG11, "ISO/IEC FDIS 23002-3
Representation of Auxiliary Video and Supplemental
Information", Doc. N8768, January 2007.
- [ITU-T H.264]
HDMI, "Advanced video coding for generic audiovisual
services", ITU-T Recommendation H.264 and ISO/
IEC 14496-10, 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.
- [RFC5234] Crocker, D. and P. Overell, "Augmented BNF for Syntax
Specifications: ABNF", STD 68, RFC 5234, January 2008.
- [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.

Author's Address

Pedro Capelastegui
Universidad Politecnica de Madrid
ETSI Telecomunicacion
Avenida Complutense, 30
Despacho B.203
Madrid 28040
Spain

Phone: +34 915 49 57 00 ext. 3024
Email: capelastegui@dit.upm.es

mmusic
Internet-Draft
Intended status: Standards Track
Expires: October 11, 2012

B. Greevenbosch
Y. Hui
Huawei
April 9, 2012

Signal 3D format
draft-greevenbosch-mmusic-sdp-3d-format-00

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 October 11, 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	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
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2013

B. Greevenbosch
Y. Hui
Huawei
October 22, 2012

Signal 3D format
draft-greevenbosch-mmusic-sdp-3d-format-01

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 April 25, 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. 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
Internet-Draft
Intended status: Standards Track
Expires: October 11, 2012

B. Greevenbosch
Y. Hui
Huawei Technologies
April 9, 2012

SDP attribute to signal parallax
draft-greevenbosch-mmusic-sdp-parallax-00

Abstract

This document introduces a "ParallaxInfo" attribute in SDP. This attribute can be used in stereoscopic applications, to signal the depth positioning of 2D media data within the 3D space.

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 October 11, 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	4
2. Requirements notation	5
3. Definitions	6
4. The ParallaxInfo attribute	7
5. Example	9
6. Remarks	10
7. Security Considerations	12
8. IANA Considerations	13
9. Normative References	14
Authors' Addresses	15

1. Introduction

To see a 3D scene, the human brain interprets two different views as perceived by the left and right eyes, and fuses these views into a single 3D perception. The depth of the object is perceived by interpreting the horizontal shift of that object between the views. This shift is called "parallax".

In stereoscopic 3D multimedia applications, there are various ways to transmit media streams in 3D. One way is to transmit two different streams, one for the left eye and one for the right eye. These streams are then projected to the relevant eyes using the appropriate technology.

When sending text streams in 3D, the solution mentioned above would imply sending the same text stream twice. Since the two streams would only differ in horizontal positioning, this introduces a lot of unnecessary overhead.

This document specifies a "ParallaxInfo" attribute in SDP [RFC4566], which is used to transfer the parallax information. It eliminates the need to send two different streams separately, as they can be calculated from a single stream and the "ParallaxInfo" attribute value.

The attribute transfers this information as two parameters: one indicating which view (left/right/centre) is carried by the stream, and another to signal the parallax between the objects.

The attribute is not restricted for signalling the parallax of text streams, but it can also be used to place other 2D objects in the 3D space. Examples include a channel logo, an electronic programme guide and on-screen display of an audio volume indicator.

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

L-view

A visual entity that is to be projected to the left eye. In the case of video, the L-view is a video frame designated for the left eye. In the case of text, the L-view is the text positioned for viewing by the left eye.

L-stream

A sequence of L-views, which is streamed to the device.

R-view

A visual entity that is to be projected to the right eye. In the case of video, the R-view is a video frame designated for the right eye. In the case of text, the R-view is the text positioned for viewing by the right eye.

R-stream

A sequence of R-views, which is streamed to the device.

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 sequence of C-views, which is streamed to the device.

stereoscopic device

A device that is able to produce and display different images for the left and right eyes, such that the viewer can experience a 3D view.

2D device

A device that is not able to produce and display different images for the left and right eyes.

2D media stream

A sequence of two dimensional visual entities (such as text or 2D graphics), which is streamed to the device.

4. The ParallaxInfo attribute

The SDP attribute "ParallaxInfo" is used to transmit the depth positioning of 2D media data, such as a text stream, in the 3D space.

The attribute has the following syntax:

```
a=ParallaxInfo:<transmitted position> <parallax>
```

The <transmitted position> indicates whether the L-, C- or R-stream is transmitted, whereas <parallax> indicates the parallax (i.e. shift) between corresponding L- and R-views in pixels, normalised to a screen width of 11520 pixels. To convert the value of <parallax> to match the display video screen width W, it has to be divided by a factor $F=11520/W$.

The <transmitted position> can have the following values:

"L" indicates that the transmitted stream is the L-stream. A stereoscopic device MUST calculate the corresponding R-views by shifting the L-views $\langle\text{parallax}\rangle/F$ pixels towards the right.

"C" indicates that the transmitted stream is the C-stream. A stereoscopic device MUST calculate the corresponding L-views by shifting the C-views $\langle\text{parallax}\rangle/(2*F)$ pixels towards the left, and the R-views by shifting the C-views $\langle\text{parallax}\rangle/(2*F)$ pixels towards the right. <parallax> SHOULD be even. If it is odd, the divided value MUST be rounded off towards zero.

"R" indicates that the transmitted stream is the R-stream. A stereoscopic device MUST calculate the corresponding L-views by shifting the R-views $\langle\text{parallax}\rangle/F$ pixels towards the left.

<parallax> MAY be negative. In this case, the shift direction is reversed.

The "ParallaxInfo" attribute can be a session-level attribute or a media-level attribute. As a session-level attribute, it specifies the default parallax value which can be applied to all the 2D media streams in the session being described. As a media-level attribute, it specifies the parallax value which can be applied to the associated 2D media stream, overriding any session-level parallax value specified.

The stereoscopic device MAY use the session-level attribute value for on-screen display, for example an audio volume indication, channel indication or electronic programme guide.

Notice that a 2D device that does not support the "ParallaxInfo" attribute will ignore it, and therefore display the 2D media data on the position as transmitted.

5. Example

The following is an example of an SDP description of a session with an audio stream, a video stream and a 3GPP timed text stream (see [3gpp-tt]), streamed using RTP as per [RFC4396]. If the display resolution is 1280x720, the parallax is scaled down with a factor $F=11952/1280=9$. The transmitted text stream is the L-stream, and with the example display resolution each R-view is $144/9=16$ pixels on the left of the L-view. This corresponds to the virtual positioning of the text in front of the screen.

```
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=ParallaxInfo:L -180
m=video 49170 RTP/AVP 99
a=rtpmap:99 H264/90000
m=video 52888 RTP/AVP 97
a=rtpmap:97 3gpp-tt/1000
a=ParallaxInfo:L -144
m=audio 52890 RTP/AVP 10
a=rtpmap:10 L16/16000/2
```

Notice that the default value "-180" is overridden by the value "-144" for the text stream. However the "-180" value is still signalled for on-screen display of e.g. volume control and other 2D graphics.

In case each R-view is 24 pixels (216 pixels in the reference resolution) on the right of the associated L-view, i.e. the virtual positioning of the text is behind the screen, then the three lines defined for 3gpp-tt can be replaced as follows:

```
m=video 52888 RTP/AVP 97
a=rtpmap:97 3gpp-tt/1000
a=ParallaxInfo:L 216
```

6. Remarks

A positive parallax value indicates virtual positioning of the 2D media data behind the screen. This is the case when the objects in the L-view are on the left of the same objects in the R-view. Similarly, a negative parallax value indicates that the objects in the R-view are on the left of the same objects in the L-view, and corresponds to virtual positioning in front of the screen.

Since the "ParallaxInfo" attribute indicates a shift of the transmitted stream, it might happen that the L- or R-view trespasses the boundaries of the display. In this case, clipping is necessary, as illustrated by Figure 1.

Similarly, there are areas which are covered by the L-view but not by the R-view and vice versa. These areas need to be filled in a sensible way. Since the "ParallaxInfo" attribute is designed for objects that overlay other video data (e.g. subtitles), it is trivial to fill in uncovered areas by using the underlying video data. However, if there is no underlying video data, other mechanisms to fill in the uncovered areas need to be defined. Definition of these mechanisms are out of scope of this document.

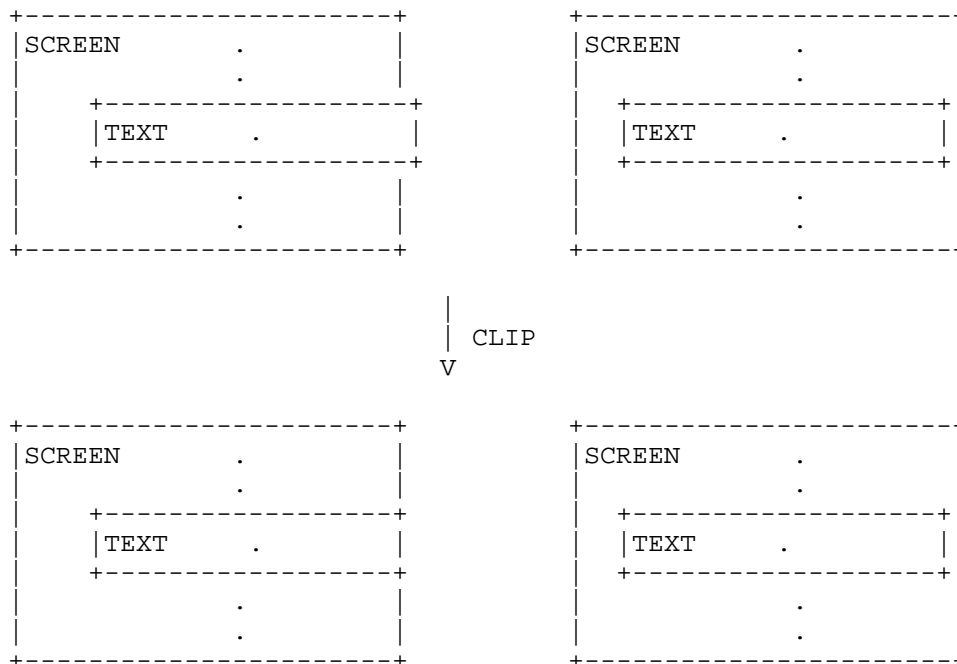


Figure 1

The normalisation of the parallax towards a reference screen width of 11520 pixels was chosen to ensure correct display of the same stream on different video display resolutions. This is especially useful when the underlying video is coded with a scalable video codec, which does not have a fixed video resolution.

7. Security Considerations

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

8. IANA Considerations

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: ParallaxInfo
- o Long-form attribute name: Parallax info for the depth positioning of 2D media data in the 3D space
- o Type of attribute: session level and media level
- o Subject to charset: no

This attribute is used to signal how 2D media data is to be displayed in the 3D space. It indicates the shift of the respective left and right views.

The attribute has the following ABNF (see [RFC4234]) description:

```
ParallaxInfo          = "a=ParallaxInfo:" TransmittedPosition Parallax
TransmittedPosition   = "C"/"L"/"R"
Parallax              = num-val
```

The attribute transfers this information as two parameters:
"TransmittedPosition" indicates which view of the 2D media data (left "L", right "R" or centre "C") is carried by the stream, and
"Parallax" signals the parallax (in pixels) of objects in the 2D media stream.

9. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 4234, October 2005.
- [RFC4396] Rey, J. and Y. Matsui, "RTP Payload Format for 3rd Generation Partnership Project (3GPP) Timed Text", RFC 4396, February 2006.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [3gpp-tt] 3GPP, "Transparent end-to-end packet switched streaming service (PSS); Protocols and codecs (Release 5)", TS 26.234 v5.3.0, December 2002.

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
Internet-Draft
Intended status: Standards Track
Expires: April 25, 2013

B. Greevenbosch
Y. Hui
Huawei Technologies
October 22, 2012

SDP attribute to signal parallax
draft-greevenbosch-mmusic-sdp-parallax-01

Abstract

This document introduces a "ParallaxInfo" attribute in SDP. This attribute can be used in stereoscopic applications, to signal the depth positioning of 2D media data within the 3D space.

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 April 25, 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. Requirements notation	5
3. Definitions	6
4. The ParallaxInfo attribute	7
5. Example	9
6. Remarks	10
7. Security Considerations	12
8. IANA Considerations	13
9. Normative References	14
Authors' Addresses	15

1. Introduction

To see a 3D scene, the human brain interprets two different views as perceived by the left and right eyes, and fuses these views into a single 3D perception. The depth of the object is perceived by interpreting the horizontal shift of that object between the views. This shift is called "parallax".

In stereoscopic 3D multimedia applications, there are various ways to transmit media streams in 3D. One way is to transmit two different streams, one for the left eye and one for the right eye. These streams are then projected to the relevant eyes using the appropriate technology.

When sending text streams in 3D, the solution mentioned above would imply sending the same text stream twice. Since the two streams would only differ in horizontal positioning, this introduces a lot of unnecessary overhead.

This document specifies a "ParallaxInfo" attribute in SDP [RFC4566], which is used to transfer the parallax information. It eliminates the need to send two different streams separately, as they can be calculated from a single stream and the "ParallaxInfo" attribute value.

The attribute transfers this information as two parameters: one indicating which view (left/right/centre) is carried by the stream, and another to signal the parallax between the objects.

The attribute is not restricted for signalling the parallax of text streams, but it can also be used to place other 2D objects in the 3D space. Examples include a channel logo, an electronic programme guide and on-screen display of an audio volume indicator.

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

L-view

A visual entity that is to be projected to the left eye. In the case of video, the L-view is a video frame designated for the left eye. In the case of text, the L-view is the text positioned for viewing by the left eye.

L-stream

A sequence of L-views, which is streamed to the device.

R-view

A visual entity that is to be projected to the right eye. In the case of video, the R-view is a video frame designated for the right eye. In the case of text, the R-view is the text positioned for viewing by the right eye.

R-stream

A sequence of R-views, which is streamed to the device.

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 sequence of C-views, which is streamed to the device.

stereoscopic device

A device that is able to produce and display different images for the left and right eyes, such that the viewer can experience a 3D view.

2D device

A device that is not able to produce and display different images for the left and right eyes.

2D media stream

A sequence of two dimensional visual entities (such as text or 2D graphics), which is streamed to the device.

4. The ParallaxInfo attribute

The SDP attribute "ParallaxInfo" is used to transmit the depth positioning of 2D media data, such as a text stream, in the 3D space.

The attribute has the following syntax:

```
a=ParallaxInfo:<transmitted position> <parallax>
```

The <transmitted position> indicates whether the L-, C- or R-stream is transmitted, whereas <parallax> indicates the parallax (i.e. shift) between corresponding L- and R-views in pixels, normalised to a screen width of 11520 pixels. To convert the value of <parallax> to match the display video screen width W, it has to be divided by a factor $F=11520/W$.

The <transmitted position> can have the following values:

"L" indicates that the transmitted stream is the L-stream. A stereoscopic device MUST calculate the corresponding R-views by shifting the L-views $\langle\text{parallax}\rangle/F$ pixels towards the right.

"C" indicates that the transmitted stream is the C-stream. A stereoscopic device MUST calculate the corresponding L-views by shifting the C-views $\langle\text{parallax}\rangle/(2*F)$ pixels towards the left, and the R-views by shifting the C-views $\langle\text{parallax}\rangle/(2*F)$ pixels towards the right. <parallax> SHOULD be even. If it is odd, the divided value MUST be rounded off towards zero.

"R" indicates that the transmitted stream is the R-stream. A stereoscopic device MUST calculate the corresponding L-views by shifting the R-views $\langle\text{parallax}\rangle/F$ pixels towards the left.

<parallax> MAY be negative. In this case, the shift direction is reversed.

The "ParallaxInfo" attribute can be a session-level attribute or a media-level attribute. As a session-level attribute, it specifies the default parallax value which can be applied to all the 2D media streams in the session being described. As a media-level attribute, it specifies the parallax value which can be applied to the associated 2D media stream, overriding any session-level parallax value specified.

The stereoscopic device MAY use the session-level attribute value for on-screen display, for example an audio volume indication, channel indication or electronic programme guide.

Notice that a 2D device that does not support the "ParallaxInfo" attribute will ignore it, and therefore display the 2D media data on the position as transmitted.

5. Example

The following is an example of an SDP description of a session with an audio stream, a video stream and a 3GPP timed text stream (see [3gpp-tt]), streamed using RTP as per [RFC4396]. If the display resolution is 1280x720, the parallax is scaled down with a factor $F=11952/1280=9$. The transmitted text stream is the L-stream, and with the example display resolution each R-view is $144/9=16$ pixels on the left of the L-view. This corresponds to the virtual positioning of the text in front of the screen.

```
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=ParallaxInfo:L -180
m=video 49170 RTP/AVP 99
a=rtpmap:99 H264/90000
m=video 52888 RTP/AVP 97
a=rtpmap:97 3gpp-tt/1000
a=ParallaxInfo:L -144
m=audio 52890 RTP/AVP 10
a=rtpmap:10 L16/16000/2
```

Notice that the default value "-180" is overridden by the value "-144" for the text stream. However the "-180" value is still signalled for on-screen display of e.g. volume control and other 2D graphics.

In case each R-view is 24 pixels (216 pixels in the reference resolution) on the right of the associated L-view, i.e. the virtual positioning of the text is behind the screen, then the three lines defined for 3gpp-tt can be replaced as follows:

```
m=video 52888 RTP/AVP 97
a=rtpmap:97 3gpp-tt/1000
a=ParallaxInfo:L 216
```

6. Remarks

A positive parallax value indicates virtual positioning of the 2D media data behind the screen. This is the case when the objects in the L-view are on the left of the same objects in the R-view. Similarly, a negative parallax value indicates that the objects in the R-view are on the left of the same objects in the L-view, and corresponds to virtual positioning in front of the screen.

Since the "ParallaxInfo" attribute indicates a shift of the transmitted stream, it might happen that the L- or R-view trespasses the boundaries of the display. In this case, clipping is necessary, as illustrated by Figure 1.

Similarly, there are areas which are covered by the L-view but not by the R-view and vice versa. These areas need to be filled in a sensible way. Since the "ParallaxInfo" attribute is designed for objects that overlay other video data (e.g. subtitles), it is trivial to fill in uncovered areas by using the underlying video data. However, if there is no underlying video data, other mechanisms to fill in the uncovered areas need to be defined. Definition of these mechanisms are out of scope of this document.

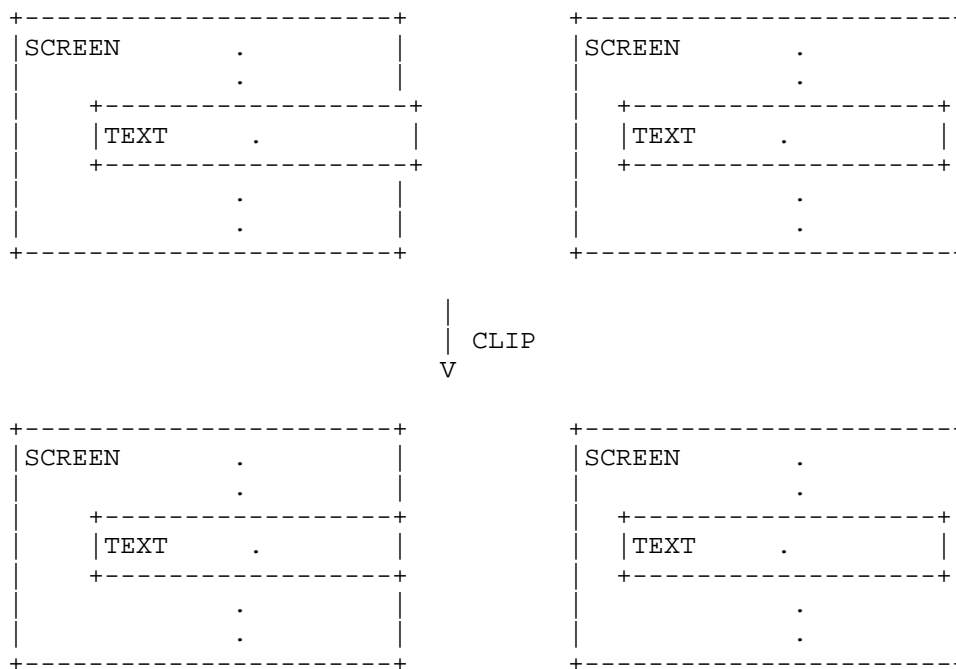


Figure 1

The normalisation of the parallax towards a reference screen width of 11520 pixels was chosen to ensure correct display of the same stream on different video display resolutions. This is especially useful when the underlying video is coded with a scalable video codec, which does not have a fixed video resolution.

7. Security Considerations

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

8. IANA Considerations

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: ParallaxInfo
- o Long-form attribute name: Parallax info for the depth positioning of 2D media data in the 3D space
- o Type of attribute: session level and media level
- o Subject to charset: no

This attribute is used to signal how 2D media data is to be displayed in the 3D space. It indicates the shift of the respective left and right views.

The attribute has the following ABNF (see [RFC4234]) description:

```
ParallaxInfo      = "a=ParallaxInfo:" TransmittedPosition Parallax
TransmittedPosition = "C"/"L"/"R"
Parallax          = num-val
```

The attribute transfers this information as two parameters:
"TransmittedPosition" indicates which view of the 2D media data (left "L", right "R" or centre "C") is carried by the stream, and
"Parallax" signals the parallax (in pixels) of objects in the 2D media stream.

9. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 4234, October 2005.
- [RFC4396] Rey, J. and Y. Matsui, "RTP Payload Format for 3rd Generation Partnership Project (3GPP) Timed Text", RFC 4396, February 2006.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [3gpp-tt] 3GPP, "Transparent end-to-end packet switched streaming service (PSS); Protocols and codecs (Release 5)", TS 26.234 v5.3.0, December 2002.

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: August 27, 2012

C. Holmberg
Ericsson
H. Alvestrand
Google
February 24, 2012

Multiplexing Negotiation Using Session Description Protocol (SDP) Port
Numbers
draft-ietf-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 August 27, 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. 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
Updates: 3264,5888,7941 (if approved)
Intended status: Standards Track
Expires: June 18, 2019

C. Holmberg
Ericsson
H. Alvestrand
Google
C. Jennings
Cisco
December 15, 2018

Negotiating Media Multiplexing Using the Session Description Protocol
(SDP)
draft-ietf-mmusic-sdp-bundle-negotiation-54.txt

Abstract

This specification defines a new Session Description Protocol (SDP) Grouping Framework extension, 'BUNDLE'. The extension can be used with the SDP Offer/Answer mechanism to negotiate the usage of a single transport (5-tuple) for sending and receiving media described by multiple SDP media descriptions ("m=" sections). Such transport is referred to as a BUNDLE transport, and the media is referred to as bundled media. The "m=" sections that use the BUNDLE transport form a BUNDLE group.

This specification updates RFC 3264, to also allow assigning a zero port value to a "m=" section in cases where the media described by the "m=" section is not disabled or rejected.

This specification updates RFC 5888, to also allow an SDP 'group' attribute to contain an identification-tag that identifies a "m=" section with the port set to zero.

This specification defines a new RTP Control Protocol (RTCP) source description (SDS) item and a new RTP header extension that can be used to correlate bundled RTP/RTCP packets with their appropriate "m=" section.

This specification updates RFC 7941, by adding an exception, for the MID RTP header extension, to the requirement regarding protection of an SDS RTP header extension carrying an SDS item for the MID RTP header extension.

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 June 18, 2019.

Copyright Notice

Copyright (c) 2018 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
1.1. Background	4
1.2. BUNDLE Mechanism	4
1.3. Protocol Extensions	5
2. Terminology	6
3. Conventions	8
4. Applicability Statement	8

5.	SDP Grouping Framework BUNDLE Extension	8
6.	SDP 'bundle-only' Attribute	9
7.	SDP Offer/Answer Procedures	10
7.1.	Generic SDP Considerations	10
7.1.1.	Connection Data (c=)	10
7.1.2.	Bandwidth (b=)	11
7.1.3.	Attributes (a=)	11
7.2.	Generating the Initial SDP Offer	12
7.2.1.	Suggesting the Offerer tagged 'm=' section	13
7.2.2.	Example: Initial SDP Offer	13
7.3.	Generating the SDP Answer	14
7.3.1.	Answerer Selection of tagged 'm=' sections	16
7.3.2.	Moving A Media Description Out Of A BUNDLE Group	16
7.3.3.	Rejecting a Media Description in a BUNDLE Group	17
7.3.4.	Example: SDP Answer	18
7.4.	Offerer Processing of the SDP Answer	19
7.5.	Modifying the Session	19
7.5.1.	Adding a Media Description to a BUNDLE group	20
7.5.2.	Moving a Media Description Out of a BUNDLE Group	21
7.5.3.	Disabling a Media Description in a BUNDLE Group	21
8.	Protocol Identification	22
8.1.	STUN, DTLS, SRTP	22
9.	RTP Considerations	23
9.1.	Single RTP Session	23
9.1.1.	Payload Type (PT) Value Reuse	24
9.2.	Associating RTP/RTCP Streams with the Correct SDP Media Description	24
9.3.	RTP/RTCP Multiplexing	30
9.3.1.	SDP Offer/Answer Procedures	30
10.	ICE Considerations	32
11.	DTLS Considerations	33
12.	RTP Header Extensions Consideration	34
13.	Update to RFC 3264	34
13.1.	Original text of section 5.1 (2nd paragraph) of RFC 3264	34
13.2.	New text replacing section 5.1 (2nd paragraph) of RFC 3264	35
13.3.	Original text of section 8.4 (6th paragraph) of RFC 3264	35
13.4.	New text replacing section 8.4 (6th paragraph) of RFC 3264	35
14.	Update to RFC 5888	36
14.1.	Original text of section 9.2 (3rd paragraph) of RFC 5888	36
14.2.	New text replacing section 9.2 (3rd paragraph) of RFC 5888	36
15.	RTP/RTCP extensions for identification-tag transport	36
15.1.	RTCP MID SDES Item	37
15.2.	RTP SDES Header Extension For MID	38
16.	IANA Considerations	38
16.1.	New SDES item	38

16.2.	New RTP SDES Header Extension URI	39
16.3.	New SDP Attribute	39
16.4.	New SDP Group Semantics	40
17.	Security Considerations	40
18.	Examples	41
18.1.	Example: Tagged m= Section Selections	41
18.2.	Example: BUNDLE Group Rejected	43
18.3.	Example: Offerer Adds a Media Description to a BUNDLE Group	45
18.4.	Example: Offerer Moves a Media Description Out of a BUNDLE Group	46
18.5.	Example: Offerer Disables a Media Description Within a BUNDLE Group	48
19.	Acknowledgements	50
20.	Change Log	50
21.	References	61
21.1.	Normative References	61
21.2.	Informative References	64
Appendix A.	Design Considerations	65
A.1.	UA Interoperability	65
A.2.	Usage of Port Number Value Zero	67
A.3.	B2BUA And Proxy Interoperability	67
A.3.1.	Traffic Policing	68
A.3.2.	Bandwidth Allocation	68
A.4.	Candidate Gathering	68
Authors' Addresses	69

1. Introduction

1.1. Background

When the SDP offer/answer mechanism [RFC3264] is used to negotiate the establishment of multimedia communication sessions, if separate transports (5-tuples) are negotiated for each individual media stream, each transport consumes additional resources (especially when Interactive Connectivity Establishment (ICE) [I-D.ietf-ice-rfc5245bis] is used). For this reason, it is attractive to use a single transport for multiple media streams.

1.2. BUNDLE Mechanism

This specification defines a way to use a single transport (BUNDLE transport) for sending and receiving media (bundled media) described by multiple SDP media descriptions ("m=" sections). The address:port combination used by an endpoint for sending and receiving bundled media is referred to as the BUNDLE address:port. The set of SDP attributes that are applied to each "m=" section within a BUNDLE group is referred to as BUNDLE attributes. The same BUNDLE transport

is used for sending and receiving bundled media, which means that the symmetric Real-time Transport Protocol (RTP) mechanism [RFC4961] is always used for RTP-based bundled media.

This specification defines a new SDP Grouping Framework [RFC5888] extension called 'BUNDLE'. The extension can be used with the Session Description Protocol (SDP) Offer/Answer mechanism [RFC3264] to negotiate which "m=" sections will become part of a BUNDLE group. In addition, the offerer and answerer [RFC3264] use the BUNDLE extension to negotiate the BUNDLE addresses:ports (offerer BUNDLE address:port and answerer BUNDLE address:port) and the set of BUNDLE attributes (offerer BUNDLE attributes and answerer BUNDLE attributes) that will be applied to each "m=" section within the BUNDLE group.

The use of a BUNDLE transport allows the usage of a single set of Interactive Connectivity Establishment (ICE) [I-D.ietf-ice-rfc5245bis] candidates for the whole BUNDLE group.

A given BUNDLE address:port MUST only be associated with a single BUNDLE group. If an SDP offer or answer contains multiple BUNDLE groups, the procedures in this specification apply to each group independently. All RTP-based bundled media associated with a given BUNDLE group belong to a single RTP session [RFC3550].

The BUNDLE extension is backward compatible. Endpoints that do not support the extension are expected to generate offers and answers without an SDP 'group:BUNDLE' attribute, and are expected to assign a unique address:port to each "m=" section within an offer and answer, according to the procedures in [RFC4566] and [RFC3264].

1.3. Protocol Extensions

In addition to defining the new SDP Grouping Framework extension, this specification defines the following protocol extensions and RFC updates:

- o The specification defines a new SDP attribute, 'bundle-only', which can be used to request that a specific "m=" section (and the associated media) is used only if kept within a BUNDLE group.
- o The specification updates RFC 3264 [RFC3264], to also allow assigning a zero port value to a "m=" section in cases where the media described by the "m=" section is not disabled or rejected.
- o The specification defines a new RTP Control Protocol (RTCP) [RFC3550] source description (SDS) item, 'MID', and a new RTP SDS header extension that can be used to associate RTP streams with "m=" sections.

- o The specification updates [RFC7941], by adding an exception, for the MID RTP header extension, to the requirement regarding protection of an SDES RTP header extension carrying an SDES item for the MID RTP header extension.

2. Terminology

- o "m=" section: SDP bodies contain one or more media descriptions, referred to as "m=" sections. Each "m=" section is represented by an SDP "m=" line, and zero or more SDP attributes associated with the "m=" line. A local address:port combination is assigned to each "m=" section.
- o 5-tuple: A collection of the following values: source address, source port, destination address, destination port, and transport-layer protocol.
- o Unique address:port: An address:port combination that is assigned to only one "m=" section in an offer or answer.
- o Offerer BUNDLE-tag: The first identification-tag in a given SDP 'group:BUNDLE' attribute identification-tag list in an offer.
- o Answerer BUNDLE-tag: The first identification-tag in a given SDP 'group:BUNDLE' attribute identification-tag list in an answer.
- o Suggested offerer tagged "m=" section: The bundled "m=" section identified by the offerer BUNDLE-tag in an initial BUNDLE offer, before a BUNDLE group has been negotiated.
- o Offerer tagged "m=" section: The bundled "m=" section identified by the offerer BUNDLE-tag in a subsequent offer. The "m=" section contains characteristics (offerer BUNDLE address:port and offerer BUNDLE attributes) applied to each "m=" section within the BUNDLE group.
- o Answerer tagged "m=" section: The bundled "m=" section identified by the answerer BUNDLE-tag in an answer (initial BUNDLE answer or subsequent). The "m=" section contains characteristics (answerer BUNDLE address:port and answerer BUNDLE attributes) applied to each "m=" section within the BUNDLE group.
- o BUNDLE address:port: An address:port combination that an endpoint uses for sending and receiving bundled media.
- o Offerer BUNDLE address:port: the address:port combination used by the offerer for sending and receiving media.

- o Answerer BUNDLE address:port: the address:port combination used by the answerer for sending and receiving media.
- o BUNDLE attributes: IDENTICAL and TRANSPORT multiplexing category SDP attributes. Once a BUNDLE group has been created, the attribute values apply to each bundled "m=" section within the BUNDLE group.
- o Offerer BUNDLE attributes: IDENTICAL and TRANSPORT multiplexing category SDP attributes included in the offerer tagged "m=" section.
- o Answerer BUNDLE attributes: IDENTICAL and TRANSPORT multiplexing category SDP attributes included in the answerer tagged "m=" section.
- o BUNDLE transport: The transport (5-tuple) used by all media described by the "m=" sections within a BUNDLE group.
- o BUNDLE group: A set of bundled "m=" sections, created using an SDP Offer/Answer exchange, which uses a single BUNDLE transport, and a single set of BUNDLE attributes, for sending and receiving all media (bundled media) described by the set of "m=" sections. The same BUNDLE transport is used for sending and receiving bundled media.
- o Bundled "m=" section: An "m=" section, whose identification-tag is placed in an SDP 'group:BUNDLE' attribute identification-tag list in an offer or answer.
- o Bundle-only "m=" section: A bundled "m=" section that contains an SDP 'bundle-only' attribute.
- o Bundled media: All media associated with a given BUNDLE group.
- o Initial BUNDLE offer: The first offer, within an SDP session (e.g. a SIP dialog when the Session Initiation Protocol (SIP) [RFC3261] is used to carry SDP), in which the offerer indicates that it wants to negotiate a given BUNDLE group.
- o Initial BUNDLE answer: The answer to an initial BUNDLE offer in which the offerer indicates that it wants to negotiate a BUNDLE group, and where the answerer accepts the creation of the BUNDLE group. The BUNDLE group is created once the answerer sends the initial BUNDLE answer.
- o Subsequent offer: An offer which contains a BUNDLE group that has been created as part of a previous offer/answer exchange.

- o Subsequent answer: An answer to a subsequent offer.
- o Identification-tag: A unique token value that is used to identify an "m=" section. The SDP 'mid' attribute [RFC5888] in an "m=" section carries the unique identification-tag assigned to that "m=" section. The session-level SDP 'group' attribute [RFC5888] carries a list of identification-tags, identifying the "m=" sections associated with that particular 'group' attribute.

3. Conventions

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 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

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]. Declarative usage of SDP is out of scope of this document, and is thus undefined.

5. SDP Grouping Framework BUNDLE Extension

This section defines a new SDP Grouping Framework [RFC5888] extension, 'BUNDLE'. The BUNDLE extension can be used with the SDP Offer/Answer mechanism to negotiate a set of "m=" sections that will become part of a BUNDLE group. Within a BUNDLE group, each "m=" section uses a BUNDLE transport for sending and receiving bundled media. Each endpoint uses a single address:port combination for sending and receiving the bundled media.

The BUNDLE extension is indicated using an SDP 'group' attribute with a semantics value [RFC5888] of "BUNDLE". An identification-tag is assigned to each bundled "m=" section, and each identification-tag is listed in the SDP 'group:BUNDLE' attribute identification-tag list. Each "m=" section whose identification-tag is listed in the identification-tag list is associated with a given BUNDLE group.

SDP bodies can contain multiple BUNDLE groups. Any given bundled "m=" section MUST NOT be associated with more than one BUNDLE group at any given time.

NOTE: The order of the "m=" sections listed in the SDP 'group:BUNDLE' attribute identification-tag list does not have to be the same as the order in which the "m=" sections occur in the SDP.

The multiplexing category [I-D.ietf-mmusic-sdp-mux-attributes] for the 'group:BUNDLE' attribute is 'NORMAL'.

Section 7 defines the detailed SDP Offer/Answer procedures for the BUNDLE extension.

6. SDP 'bundle-only' Attribute

This section defines a new SDP media-level attribute [RFC4566], 'bundle-only'. 'bundle-only' is a property attribute [RFC4566], and hence has no value.

In order to ensure that an answerer that does not support the BUNDLE extension always rejects a bundled "m=" section in an offer, the offerer can assign a zero port value to the "m=" section. According to [RFC3264] an answerer will reject such an "m=" section. By including an SDP 'bundle-only' attribute in a bundled "m=" section, the offerer can request that the answerer accepts the "m=" section only if the answerer supports the BUNDLE extension, and if the answerer keeps the "m=" section within the associated BUNDLE group.

Name: bundle-only

Value: N/A

Usage Level: media

Charset Dependent: no

Example:

a=bundle-only

Once the offerer tagged "m=" section and the answerer tagged "m=" section have been selected, an offerer and answerer will include an SDP 'bundle-only' attribute in, and assign a zero port value to, every other bundled "m=" section.

The usage of the 'bundle-only' attribute is only defined for a bundled "m=" section with a zero port value. Other usage is unspecified.

Section 7 defines the detailed SDP Offer/Answer procedures for the 'bundle-only' attribute.

7. SDP Offer/Answer Procedures

This section describes the SDP Offer/Answer [RFC3264] procedures for:

- o Negotiating a BUNDLE group; and
- o Suggesting and selecting the tagged "m=" sections (offerer tagged "m=" section and answerer tagged "m=" section); and
- o Adding an "m=" section to a BUNDLE group; and
- o Moving an "m=" section out of a BUNDLE group; and
- o Disabling an "m=" section within a BUNDLE group.

The generic rules and procedures defined in [RFC3264] and [RFC5888] also apply to the BUNDLE extension. For example, if an offer is rejected by the answerer, the previously negotiated addresses:ports, SDP parameters and characteristics (including those associated with a BUNDLE group) apply. Hence, if an offerer generates an offer in order to negotiate a BUNDLE group, and the answerer rejects the offer, the BUNDLE group is not created.

The procedures in this section are independent of the media type or "m=" line proto value assigned to a bundled "m=" section. Section 9 defines additional considerations for RTP based media. Section 6 defines additional considerations for the usage of the SDP 'bundle-only' attribute. Section 10 defines additional considerations for the usage of Interactive Connectivity Establishment (ICE) [I-D.ietf-ice-rfc5245bis] mechanism.

Offers and answers can contain multiple BUNDLE groups. The procedures in this section apply independently to a given BUNDLE group.

7.1. Generic SDP Considerations

This section describes generic restrictions associated with the usage of SDP parameters within a BUNDLE group. It also describes how to calculate a value for the whole BUNDLE group, when parameter and attribute values have been assigned to each bundled "m=" section.

7.1.1. Connection Data (c=)

The "c=" line nettype value [RFC4566] associated with a bundled "m=" section MUST be 'IN'.

The "c=" line addrtype value [RFC4566] associated with a bundled "m=" section MUST be 'IP4' or 'IP6'. The same value MUST be associated with each "m=" section.

NOTE: Extensions to this specification can specify usage of the BUNDLE mechanism for other nettype and addrtype values than the ones listed above.

7.1.2. Bandwidth (b=)

An offerer and answerer MUST use the rules and restrictions defined in [I-D.ietf-mmusic-sdp-mux-attributes] for associating the SDP bandwidth (b=) line with bundled "m=" sections.

7.1.3. Attributes (a=)

An offerer and answerer MUST include SDP attributes in every bundled "m=" section where applicable, following the normal offer/answer procedures for each attribute, with the following exceptions:

- o In the initial BUNDLE offer, the offerer MUST NOT include IDENTICAL and TRANSPORT multiplexing category SDP attributes (BUNDLE attributes) in bundle-only "m=" sections. The offerer MUST include such attributes in all other bundled "m=" sections. In the initial BUNDLE offer each bundled "m=" line can contain a different set of BUNDLE attributes, and attribute values. Once the offerer tagged "m=" section has been selected, the BUNDLE attributes contained in the offerer tagged "m=" section will apply to each bundled "m=" section within the BUNDLE group.
- o In a subsequent offer, or in an answer (initial or subsequent), the offerer and answerer MUST include IDENTICAL and TRANSPORT multiplexing category SDP attributes (BUNDLE attributes) only in the tagged "m=" section (offerer tagged "m=" section or answerer tagged "m=" section). The offerer and answerer MUST NOT include such attributes in any other bundled "m=" section. The BUNDLE attributes contained in the tagged "m=" section will apply to each bundled "m=" section within the BUNDLE group.
- o In an offer (initial BUNDLE offer or subsequent), or in an answer (initial BUNDLE answer or subsequent), the offerer and answerer MUST include SDP attributes of other categories than IDENTICAL and TRANSPORT in each bundled "m=" section that a given attribute applies to. Each bundled "m=" line can contain a different set of such attributes, and attribute values, as such attributes only apply to the given bundled "m=" section in which they are included.

NOTE: A consequence of the rules above is that media-specific IDENTICAL and TRANSPORT multiplexing category SDP attributes which are applicable only to some of the bundled "m=" sections within the BUNDLE group might appear in the tagged "m=" section for which they are not applicable. For instance, the tagged "m=" section might contain an SDP 'rtcp-mux' attribute even if the tagged "m=" section does not describe RTP-based media (but another bundled "m=" section within the BUNDLE group does describe RTP-based media).

7.2. Generating the Initial SDP Offer

The procedures in this section apply to the first offer, within an SDP session (e.g. a SIP dialog when the Session Initiation Protocol (SIP) [RFC3261] is used to carry SDP), in which the offerer indicates that it wants to negotiate a given BUNDLE group. This could occur in the initial offer, or in a subsequent offer, of the SDP session.

When an offerer generates an initial BUNDLE offer, in order to negotiate a BUNDLE group, it MUST:

- o Assign a unique address:port to each bundled "m=" section, following the procedures in [RFC3264], excluding any bundle-only "m=" sections (see below); and
- o Pick a bundled "m=" section as the suggested offerer tagged "m=" section [Section 7.2.1]; and
- o Include SDP attributes in the bundled "m=" sections following the rules in [Section 7.1.3]; and
- o Include an SDP 'group:BUNDLE' attribute in the offer; and
- o Place the identification-tag of each bundled "m=" section in the SDP 'group:BUNDLE' attribute identification-tag list. The offerer BUNDLE-tag indicates the suggested offerer tagged "m=" section.

NOTE: When the offerer assigns unique addresses:ports to multiple bundled "m=" sections, the offerer needs to be prepared to receive bundled media on each unique address:port, until it receives the associated answer and finds out which bundled "m=" section (and associated address:port combination) the answerer has selected as the offerer tagged "m=" section.

If the offerer wants to request that the answerer accepts a given bundled "m=" section only if the answerer keeps the "m=" section within the negotiated BUNDLE group, the offerer MUST:

- o Include an SDP 'bundle-only' attribute [Section 7.2.1] in the "m=" section; and
- o Assign a zero port value to the "m=" section.

NOTE: If the offerer assigns a zero port value to a bundled "m=" section, but does not include an SDP 'bundle-only' attribute in the "m=" section, it is an indication that the offerer wants to disable the "m=" section [Section 7.5.3].

[Section 7.2.2] and [Section 18.1] show an example of an initial BUNDLE offer.

7.2.1. Suggesting the Offerer tagged 'm=' section

In the initial BUNDLE offer, the bundled "m=" section indicated by the offerer BUNDLE-tag is the suggested offerer tagged "m=" section. The address:port combination associated with the "m=" section will be used by the offerer for sending and receiving bundled media if the answerer selects the "m=" section as the offerer tagged "m=" section [Section 7.3.1]. In addition, if the answerer selects the "m=" section as the offerer tagged "m=" section, the BUNDLE attributes included in the "m=" section will be applied to each "m=" section within the negotiated BUNDLE group.

The offerer MUST NOT suggest a bundle-only "m=" section as the offerer tagged "m=" section.

It is RECOMMENDED that the suggested offerer tagged "m=" section is a bundled "m=" section that the offerer believes it is unlikely that the answerer will reject, or move out of the BUNDLE group. How such assumption is made is outside the scope of this document.

7.2.2. Example: Initial SDP Offer

The example shows an initial BUNDLE offer. The offer includes two "m=" sections in the offer, and suggests that both "m=" sections are included in a BUNDLE group. The audio "m=" section is the suggested offerer tagged "m=" section, indicated by placing the identification-tag associated with the "m=" section (offerer BUNDLE-tag) first in the SDP group:BUNDLE attribute identification-id list.

SDP Offer

```
v=0
o=alice 2890844526 2890844526 IN IP6 2001:db8::3
s=
c=IN IP6 2001:db8::3
t=0 0
a=group:BUNDLE foo bar

m=audio 10000 RTP/AVP 0 8 97
b=AS:200
a=mid:foo
a=rtcp-mux
a=rtpmap:0 PCMU/8000
a=rtpmap:8 PCMA/8000
a=rtpmap:97 iLBC/8000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid

m=video 10002 RTP/AVP 31 32
b=AS:1000
a=mid:bar
a=rtcp-mux
a=rtpmap:31 H261/90000
a=rtpmap:32 MPV/90000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid
```

7.3. Generating the SDP Answer

When an answerer generates an answer (initial BUNDLE answer or subsequent) that contains a BUNDLE group the following general SDP grouping framework restrictions, defined in [RFC5888], also apply to the BUNDLE group:

- o The answerer is only allowed to include a BUNDLE group in an initial BUNDLE answer if the offerer requested the BUNDLE group to be created in the corresponding initial BUNDLE offer; and
- o The answerer is only allowed to include a BUNDLE group in a subsequent answer if the corresponding subsequent offer contains a previously negotiated BUNDLE group; and
- o The answerer is only allowed to include a bundled "m=" section in an answer if the "m=" section was indicated as bundled in the corresponding offer; and

- o The answerer is only allowed to include a bundled "m=" section in the same BUNDLE group as the bundled "m=" line in the corresponding offer.

In addition, when an answerer generates an answer (initial BUNDLE answer or subsequent) that contains a BUNDLE group, the answerer MUST:

- o In case of an initial BUNDLE answer, select the offerer tagged "m=" section using the procedures in Section 7.3.1. In case of a subsequent answer, the offerer tagged "m=" section is indicated in the corresponding subsequent offer, and MUST NOT be changed by the answerer; and
- o Select the answerer tagged "m=" section [Section 7.3.1]; and
- o Assign the answerer BUNDLE address:port to the answerer tagged "m=" section; and
- o Include an SDP 'bundle-only' attribute in, and assign a zero port value to, every other bundled "m=" section within the BUNDLE group; and
- o Include SDP attributes in the bundled "m=" sections following the rules in [Section 7.1.3]; and
- o Include an SDP 'group:BUNDLE' attribute in the answer; and
- o Place the identification-tag of each bundled "m=" section in the SDP 'group:BUNDLE' attribute identification-tag list. The answerer BUNDLE-tag indicates the answerer tagged "m=" section [Section 7.3.1].

If the answerer does not want to keep an "m=" section within a BUNDLE group, it MUST:

- o Move the "m=" section out of the BUNDLE group [Section 7.3.2]; or
- o Reject the "m=" section [Section 7.3.3].

The answerer can modify the answerer BUNDLE address:port, add and remove SDP attributes, or modify SDP attribute values, in a subsequent answer. Changes to the answerer BUNDLE address:port and the answerer BUNDLE attributes will be applied to each bundled "m=" section within the BUNDLE group.

NOTE: If a bundled "m=" section in an offer contains a zero port value, but the "m=" section does not contain an SDP 'bundle-only'

attribute, it is an indication that the offerer wants to disable the "m=" section [Section 7.5.3].

7.3.1. Answerer Selection of tagged 'm=' sections

When the answerer selects the offerer tagged "m=" section, it first checks the suggested offerer tagged "m=" section [Section 7.2.1]. The answerer MUST check whether the "m=" section fulfils the following criteria:

- o The answerer will not move the "m=" section out of the BUNDLE group [Section 7.3.2]; and
- o The answerer will not reject the "m=" section [Section 7.3.3]; and
- o The "m=" section does not contain a zero port value.

If all of the criteria above are fulfilled, the answerer MUST select the "m=" section as the offerer tagged "m=" section, and MUST also mark the corresponding "m=" section in the answer as the answerer tagged "m=" section. In the answer the answerer BUNDLE-tag indicates the answerer tagged "m=" section.

If one or more of the criteria are not fulfilled, the answerer MUST pick the next identification-tag in the identification-tag list in the offer, and perform the same criteria check for the "m=" section indicated by that identification-tag. If there are no more identification-tags in the identification-tag list, the answerer MUST NOT create the BUNDLE group. Unless the answerer rejects the whole offer, the answerer MUST apply the answerer procedures for moving an "m=" section out of a BUNDLE group [Section 7.3.2] or rejecting an "m=" section within a BUNDLE group [Section 7.3.3] to every bundled "m=" section in the offer when creating the answer.

[Section 18.1] shows an example of an offerer BUNDLE address:port selection.

[Section 7.3.4] and [Section 18.1] show an example of an answerer tagged "m=" section selection.

7.3.2. Moving A Media Description Out Of A BUNDLE Group

When an answerer generates the answer, if the answerer wants to move a bundled "m=" section out of the negotiated BUNDLE group, the answerer MUST first check the following criteria:

- o In the corresponding offer, the "m=" section is within a previously negotiated BUNDLE group; and

- o In the corresponding offer, the "m=" section contains an SDP 'bundle-only' attribute.

If either criterium above is fulfilled the answerer can not move the "m=" section out of the BUNDLE group in the answer. The answerer can either reject the whole offer, reject each bundled "m=" section within the BUNDLE group [Section 7.3.3], or keep the "m=" section within the BUNDLE group in the answer and later create an offer where the "m=" section is moved out of the BUNDLE group [Section 7.5.2].

NOTE: One consequence of the rules above is that, once a BUNDLE group has been negotiated, a bundled "m=" section can not be moved out of the BUNDLE group in an answer. Instead an offer is needed.

When the answerer generates an answer, in which it moves a bundled "m=" section out of a BUNDLE group, the answerer:

- o MUST assign a unique address:port to the "m=" section; and
- o MUST include any applicable SDP attribute in the "m=" section, using the normal offer/answer procedures for the each Attributes; and
- o MUST NOT place the identification-tag associated with the "m=" section in the SDP 'group:BUNDLE' attribute identification-tag list associated with the BUNDLE group.
- o MUST NOT include an SDP 'bundle-only' attribute to the "m=" section; and

Because an answerer is not allowed to move an "m=" section from one BUNDLE group to another within an answer [Section 7.3], if the answerer wants to move an "m=" section from one BUNDLE group to another it MUST first move the "m=" section out of the current BUNDLE group, and then generate an offer where the "m=" section is added to another BUNDLE group [Section 7.5.1].

7.3.3. Rejecting a Media Description in a BUNDLE Group

When an answerer wants to reject a bundled "m=" section in an answer, it MUST first check the following criterion:

- o In the corresponding offer, the "m=" section is the offerer tagged "m=" section.

If the criterium above is fulfilled the answerer can not reject the "m=" section in the answer. The answerer can either reject the whole offer, reject each bundled "m=" section within the BUNDLE group, or

keep the "m=" section within the BUNDLE group in the answer and later create an offer where the "m=" section is disabled within the BUNDLE group [Section 7.5.3].

When an answerer generates an answer, in which it rejects a bundled "m=" section, the answerer:

- o MUST assign a zero port value to the "m=" section, according to the procedures in [RFC3264]; and
- o MUST NOT place the identification-tag associated with the "m=" section in the SDP 'group:BUNDLE' attribute identification-tag list associated with the BUNDLE group; and
- o MUST NOT include an SDP 'bundle-only' attribute in the "m=" section.

7.3.4. Example: SDP Answer

The example below shows an answer, based on the corresponding offer in [Section 7.2.2]. The answerer accepts both bundled "m=" sections within the created BUNDLE group. The audio "m=" section is the answerer tagged "m=" section, indicated by placing the identification-tag associated with the "m=" section (answerer BUNDLE-tag) first in the SDP group;BUNDLE attribute identification-id list. The answerer includes an SDP 'bundle-only' attribute in, and assigns a zero port value to, the video "m=" section.

SDP Answer

```
v=0
o=bob 2808844564 2808844564 IN IP6 2001:db8::1
s=
c=IN IP6 2001:db8::1
t=0 0
a=group:BUNDLE foo bar

m=audio 20000 RTP/AVP 0
b=AS:200
a=mid:foo
a=rtcp-mux
a=rtpmap:0 PCMU/8000
a=extmap:1 urn:ietf:params:rtp-hdext:sdes:mid

m=video 0 RTP/AVP 32
b=AS:1000
a=mid:bar
a=bundle-only
a=rtpmap:32 MPV/90000
a=extmap:1 urn:ietf:params:rtp-hdext:sdes:mid
```

7.4. Offerer Processing of the SDP Answer

When an offerer receives an answer, if the answer contains a BUNDLE group, the offerer MUST check that any bundled "m=" section in the answer was indicated as bundled in the corresponding offer. If there is no mismatch, the offerer MUST apply the properties (BUNDLE address:port, BUNDLE attributes etc) of the offerer tagged "m=" section (selected by the answerer [Section 7.3.1]) to each bundled "m=" section within the BUNDLE group.

NOTE: As the answerer might reject one or more bundled "m=" sections in an initial BUNDLE offer, or move a bundled "m=" section out of a BUNDLE group, a given bundled "m=" section in the offer might not be indicated as bundled in the corresponding answer.

If the answer does not contain a BUNDLE group, the offerer MUST process the answer as a normal answer.

7.5. Modifying the Session

When a BUNDLE group has previously been negotiated, and an offerer generates a subsequent offer, the offerer MUST:

- o Pick one bundled "m=" section as the offerer tagged "m=" section. The offerer can either pick the "m=" section that was previously selected by the answerer as the offerer tagged "m=" section, or pick another bundled "m=" section within the BUNDLE group; and
- o Assign a BUNDLE address:port (previously negotiated or newly suggest) to the offerer tagged "m=" section; and
- o Include an SDP 'bundle-only' attribute in, and assign a zero port value to, every other bundled "m=" section within the BUNDLE group; and
- o Include SDP attributes in the bundled "m=" sections following the rules in [Section 7.1.3]; and
- o Include an SDP 'group:BUNDLE' attribute in the offer; and
- o Place the identification-tag of each bundled "m=" section in the SDP 'group:BUNDLE' attribute identification-tag list. The offerer BUNDLE-tag indicates the offerer tagged "m=" section.

The offerer MUST NOT pick a given bundled "m=" section as the offerer tagged "m=" section if:

- o The offerer wants to move the "m=" section out of the BUNDLE group [Section 7.5.2]; or
- o The offerer wants to disable the "m=" section [Section 7.5.3].

The offerer can modify the offerer BUNDLE address:port, add and remove SDP attributes, or modify SDP attribute values, in the subsequent offer. Changes to the offerer BUNDLE address:port and the offerer BUNDLE attributes will (if the offer is accepted by the answerer) be applied to each bundled "m=" section within the BUNDLE group.

7.5.1. Adding a Media Description to a BUNDLE group

When an offerer generates a subsequent offer, in which it wants to add a bundled "m=" section to a previously negotiated BUNDLE group, the offerer follows the procedures in Section 7.5. The offerer either picks the added "m=" section, or an "m=" section previously added to the BUNDLE group, as the offerer tagged "m=" section.

NOTE: As described in Section 7.3.2, the answerer can not move the added "m=" section out of the BUNDLE group in its answer. If the answer wants to move the "m=" section out of the BUNDLE group, it will have to first accept it into the BUNDLE group in the answer, and

then send a subsequent offer where the "m=" section is moved out of the BUNDLE group [Section 7.5.2].

7.5.2. Moving a Media Description Out of a BUNDLE Group

When an offerer generates a subsequent offer, in which it want to remove a bundled "m=" section from a BUNDLE group, the offerer:

- o MUST assign a unique address:port to the "m=" section; and
- o MUST include SDP attributes in the "m=" section following the normal offer/answer rules for each attribute; and
- o MUST NOT place the identification-tag associated with the "m=" section in the SDP 'group:BUNDLE' attribute identification-tag list associated with the BUNDLE group; and
- o MUST NOT assign an SDP 'bundle-only' attribute to the "m=" section.

For the other bundled "m=" sections within the BUNDLE group, the offerer follows the procedures in [Section 7.5].

An offerer MUST NOT move an "m=" section from one BUNDLE group to another within a single offer. If the offerer wants to move an "m=" section from one BUNDLE group to another it MUST first move the BUNDLE group out of the current BUNDLE group, and then generate a second offer where the "m=" section is added to another BUNDLE group [Section 7.5.1].

[Section 18.4] shows an example of an offer for moving an "m=" section out of a BUNDLE group.

7.5.3. Disabling a Media Description in a BUNDLE Group

When an offerer generates a subsequent offer, in which it want to disable a bundled "m=" section from a BUNDLE group, the offerer:

- o MUST assign a zero port value to the "m=" section, following the procedures in [RFC4566]; and
- o MUST NOT place the identification-tag associated with the "m=" section in the SDP 'group:BUNDLE' attribute identification-tag list associated with the BUNDLE group; and
- o MUST NOT assign an SDP 'bundle-only' attribute to the "m=" section.

For the other bundled "m=" sections within the BUNDLE group, the offerer follows the procedures in [Section 7.5].

[Section 18.5] shows an example of an offer and answer for disabling an "m=" section within a BUNDLE group.

8. Protocol Identification

Each "m=" section within a BUNDLE group MUST use the same transport-layer protocol. If bundled "m=" sections use different upper-layer protocols on top of the transport-layer protocol, there MUST exist a publicly available specification which describes a mechanism how to associate received data with the correct protocol for this particular protocol combination.

In addition, if received data can be associated with more than one bundled "m=" section, there MUST exist a publicly available specification which describes a mechanism for associating the received data with the correct "m=" section.

This document describes a mechanism to identify the protocol of received data among the STUN, DTLS and SRTP protocols (in any combination), when UDP is used as transport-layer protocol, but it does not describe how to identify different protocols transported on DTLS. While the mechanism is generally applicable to other protocols and transport-layer protocols, any such use requires further specification around how to multiplex multiple protocols on a given transport-layer protocol, and how to associate received data with the correct protocols.

8.1. STUN, DTLS, SRTP

Section 5.1.2 of [RFC5764] describes a mechanism to identify the protocol of a received packet among the STUN, DTLS and SRTP protocols (in any combination). If an offer or answer includes a bundled "m=" section that represents these protocols, the offerer or answerer MUST support the mechanism described in [RFC5764], and no explicit negotiation is required in order to indicate support and usage of the mechanism.

[RFC5764] does not describe how to identify different protocols transported on DTLS, only how to identify the DTLS protocol itself. If multiple protocols are transported on DTLS, there MUST exist a specification describing a mechanism for identifying each individual protocol. In addition, if a received DTLS packet can be associated with more than one "m=" section, there MUST exist a specification which describes a mechanism for associating the received DTLS packets with the correct "m=" section.

[Section 9.2] describes how to associate the packets in a received SRTP stream with the correct "m=" section.

9. RTP Considerations

9.1. Single RTP Session

All RTP-based media within a single BUNDLE group belong to a single RTP session [RFC3550].

Since a single BUNDLE transport is used for sending and receiving bundled media, the symmetric RTP mechanism [RFC4961] MUST be used for RTP-based bundled media.

Since a single RTP session is used for each BUNDLE group, all "m=" sections representing RTP-based media within a BUNDLE group will share a single SSRC numbering space [RFC3550].

The following rules and restrictions apply for a single RTP session:

- o A specific payload type value can be used in multiple bundled "m=" sections only if each codec associated with the payload type number shares an identical codec configuration [Section 9.1.1].
- o The proto value in each bundled RTP-based "m=" section MUST be identical (e.g., RTP/AVPF).
- o The RTP MID header extension MUST be enabled, by including an SDP 'extmap' attribute [RFC8285], with a 'urn:ietf:params:rtp-hdext:sdes:mid' URI value, in each bundled RTP-based "m=" section in every offer and answer.
- o A given SSRC MUST NOT transmit RTP packets using payload types that originate from different bundled "m=" sections.

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 fail to function. Even if done in proper sequence this causes RTP Timestamp rate switching issues [RFC7160]. However, once an SSRC has left the RTP session (by sending an RTCP BYE packet), that SSRC can be reused by another source (possibly associated with a different bundled "m=" section) after a delay of 5 RTCP reporting intervals (the delay is to ensure the SSRC has timed out, in case the RTCP BYE packet was lost [RFC3550]).

[RFC7657] defines Differentiated Services (Diffserv) considerations for RTP-based bundled media sent using a mixture of Diffserv Codepoints.

9.1.1. Payload Type (PT) Value Reuse

Multiple bundled "m=" sections might describe RTP based media. As all RTP based media associated with a BUNDLE group belong to the same RTP session, in order for a given payload type value to be used inside more than one bundled "m=" section, all codecs associated with the payload type number MUST share an identical codec configuration. This means that the codecs MUST share the same media type, encoding name, clock rate and any parameter that can affect the codec configuration and packetization.

[I-D.ietf-mmusic-sdp-mux-attributes] lists SDP attributes, whose attribute values are required to be identical for all codecs that use the same payload type value.

9.2. Associating RTP/RTCP Streams with the Correct SDP Media Description

As described in [RFC3550], RTP packets are associated with RTP streams [RFC7656]. Each RTP stream is identified by an SSRC value, and each RTP packet includes an SSRC field that is used to associate the packet with the correct RTP stream. RTCP packets also use SSRCs to identify which RTP streams the packet relates to. However, a RTCP packet can contain multiple SSRC fields, in the course of providing feedback or reports on different RTP streams, and therefore can be associated with multiple such streams.

In order to be able to process received RTP/RTCP packets correctly, it MUST be possible to associate an RTP stream with the correct "m=" section, as the "m=" section and SDP attributes associated with the "m=" section contains information needed to process the packets.

As all RTP streams associated with a BUNDLE group use the same transport for sending and receiving RTP/RTCP packets, the local address:port combination part of the transport cannot be used to associate an RTP stream with the correct "m=" section. In addition, multiple RTP streams might be associated with the same "m=" section.

An offerer and answerer can inform each other which SSRC values they will use for an RTP stream by using the SDP 'ssrc' attribute [RFC5576]. However, an offerer will not know which SSRC values the answerer will use until the offerer has received the answer providing that information. Due to this, before the offerer has received the answer, the offerer will not be able to associate an RTP stream with the correct "m=" section using the SSRC value associated with the RTP

stream. In addition, the offerer and answerer may start using new SSRC values mid-session, without informing each other using the SDP 'ssrc' attribute.

In order for an offerer and answerer to always be able to associate an RTP stream with the correct "m=" section, the offerer and answerer using the BUNDLE extension MUST support the mechanism defined in Section 15, where the offerer and answerer insert the identification-tag associated with an "m=" section (provided by the remote peer) into RTP and RTCP packets associated with a BUNDLE group.

When using this mechanism, the mapping from an SSRC to an identification-tag is carried in RTP header extensions or RTCP SDES packets, as specified in Section 15. Since a compound RTCP packet can contain multiple RTCP SDES packets, and each RTCP SDES packet can contain multiple chunks, a single RTCP packet can contain several SSRC to identification-tag mappings. The offerer and answerer maintain tables used for routing that are updated each time an RTP/RTCP packet contains new information that affects how packets are to be routed.

However, some legacy implementations may not include this identification-tag in their RTP and RTCP traffic when using the BUNDLE mechanism, and instead use a payload type based mechanism to associate RTP streams with SDP "m=" sections. In this situation, each "m=" section needs to use unique payload type values, in order for the payload type to be a reliable indicator of the relevant "m=" section for the RTP stream. If an implementation fails to ensure unique payload type values it will be impossible to associate the RTP stream using that payload type value to a particular "m=" section. Note that when using the payload type to associate RTP streams with "m=" sections an RTP stream, identified by its SSRC, will be mapped to an "m=" section when the first packet of that RTP stream is received, and the mapping will not be changed even if the payload type used by that RTP stream changes. In other words, the SSRC cannot "move" to a different "m=" section simply by changing the payload type.

Applications can implement RTP stacks in many different ways. The algorithm below details one way that RTP streams can be associated with "m=" sections, but is not meant to be prescriptive about exactly how an RTP stack needs to be implemented. Applications MAY use any algorithm that achieves equivalent results to those described in the algorithm below.

To prepare to associate RTP streams with the correct "m=" section, the following steps MUST be followed for each BUNDLE group:

Construct a table mapping MID to "m=" section for each "m=" section in this BUNDLE group. Note that an "m=" section may only have one MID.

Construct a table mapping SSRCs of incoming RTP streams to "m=" section for each "m=" section in this BUNDLE group and for each SSRC configured for receiving in that "m=" section.

Construct a table mapping the SSRC of each outgoing RTP stream to "m=" section for each "m=" section in this BUNDLE group and for each SSRC configured for sending in that "m=" section.

Construct a table mapping payload type to "m=" section for each "m=" section in the BUNDLE group and for each payload type configured for receiving in that "m=" section. If any payload type is configured for receiving in more than one "m=" section in the BUNDLE group, do not include it in the table, as it cannot be used to uniquely identify an "m=" section.

Note that for each of these tables, there can only be one mapping for any given key (MID, SSRC, or PT). In other words, the tables are not multimaps.

As "m=" sections are added or removed from the BUNDLE groups, or their configurations are changed, the tables above MUST also be updated.

When an RTP packet is received, it MUST be delivered to the RTP stream corresponding to its SSRC. That RTP stream MUST then be associated with the correct "m=" section within a BUNDLE group, for additional processing, according to the following steps:

If the MID associated with the RTP stream is not in the table mapping MID to "m=" section, then the RTP stream is not decoded and the payload data is discarded.

If the packet has a MID, and the packet's extended sequence number is greater than that of the last MID update, as discussed in [RFC7941], Section 4.2.6, update the MID associated with the RTP stream to match the MID carried in the RTP packet, then update the mapping tables to include an entry that maps the SSRC of that RTP stream to the "m=" section for that MID.

If the SSRC of the RTP stream is in the incoming SSRC mapping table, check that the payload type used by the RTP stream matches a payload type included on the matching "m=" section. If so, associate the RTP stream with that "m=" section. Otherwise, the RTP stream is not decoded and the payload data is discarded.

If the payload type used by the RTP stream is in the payload type table, update the incoming SSRC mapping table to include an entry that maps the RTP stream's SSRC to the "m=" section for that payload type. Associate the RTP stream with the corresponding "m=" section.

Otherwise, mark the RTP stream as not for decoding and discard the payload.

If the RTP packet contains one or more contributing source (CSRC) identifiers, then each CSRC is looked up in the incoming SSRC table and a copy of the RTP packet is associated with the corresponding "m=" section for additional processing.

For each RTCP packet received (including each RTCP packet that is part of a compound RTCP packet), the packet is processed as usual by the RTP layer, then associated with the appropriate "m=" sections, and processed for the RTP streams represented by those "m=" sections. This routing is type-dependent, as each kind of RTCP packet has its own mechanism for associating it with the relevant RTP streams.

RTCP packets that cannot be associated with an appropriate "m=" section MUST still be processed as usual by the RTP layer, updating the metadata associated with the corresponding RTP streams. This situation can occur with certain multiparty RTP topologies, or when RTCP packets are sent containing a subset of the SDES information.

Additional rules for processing various types of RTCP packets are explained below.

If the RTCP packet is of type SDES, for each chunk in the packet whose SSRC is found in the incoming SSRC table, deliver a copy of the SDES packet to the "m=" section associated with that SSRC. In addition, for any SDES MID items contained in these chunks, if the MID is found in the table mapping MID to "m=" section, update the incoming SSRC table to include an entry that maps the RTP stream associated with the chunk's SSRC to the "m=" section associated with that MID, unless the packet is older than the packet that most recently updated the mapping for this SSRC, as discussed in [RFC7941], Section 4.2.6.

Note that if an SDES packet is received as part of a compound RTCP packet, the SSRC to "m=" section mapping might not exist until the SDES packet is handled (e.g., in the case where RTCP for a source is received before any RTP packets). Therefore, it can be beneficial for an implementation to delay RTCP packet routing, such that it either prioritizes processing of the SDES item to generate or update the mapping, or buffers the RTCP information

that needs to be routed until the SDES item(s) has been processed. If the implementation is unable to follow this recommendation, the consequence could be that some RTCP information from this particular RTCP compound packet is not provided to higher layers. The impact from this is likely minor, when this information relates to a future incoming RTP stream.

If the RTCP packet is of type BYE, it indicates that the RTP streams referenced in the packet are ending. Therefore, for each SSRC indicated in the packet that is found in the incoming SSRC table, first deliver a copy of the BYE packet to the "m=" section associated with that SSRC, then remove the entry for that SSRC from the incoming SSRC table after an appropriate delay to account for "straggler packets", as specified in [RFC3550], Section 6.2.1.

If the RTCP packet is of type SR or RR, for each report block in the report whose "SSRC of source" is found in the outgoing SSRC table, deliver a copy of the SR or RR packet to the "m=" section associated with that SSRC. In addition, if the packet is of type SR, and the sender SSRC for the packet is found in the incoming SSRC table, deliver a copy of the SR packet to the "m=" section associated with that SSRC.

If the implementation supports RTCP XR and the packet is of type XR, as defined in [RFC3611], for each report block in the report whose "SSRC of source" is found in the outgoing SSRC table, deliver a copy of the XR packet to the "m=" section associated with that SSRC. In addition, if the sender SSRC for the packet is found in the incoming SSRC table, deliver a copy of the XR packet to the "m=" section associated with that SSRC.

If the RTCP packet is a feedback message of type RTPFB or PSFB, as defined in [RFC4585], it will contain a media source SSRC, and this SSRC is used for routing certain subtypes of feedback messages. However, several subtypes of PSFB and RTPFB messages include target SSRC(s) in a section called Feedback Control Information (FCI). For these messages, the target SSRC(s) are used for routing.

If the RTCP packet is a feedback packet that does not include target SSRCs in its FCI section, and the media source SSRC is found in the outgoing SSRC table, deliver the feedback packet to the "m=" section associated with that SSRC. RTPFB and PSFB types that are handled in this way include:

Generic NACK: [RFC4585] (PT=RTPFB, FMT=1).

Picture Loss Indication (PLI): [RFC4585] (PT=PSFB, FMT=1).

Slice Loss Indication (SLI): [RFC4585] (PT=PSFB, FMT=2).

Reference Picture Selection Indication (RPSI): [RFC4585]
(PT=PSFB, FMT=3).

If the RTCP packet is a feedback message that does include target SSRC(s) in its FCI section, it can either be a request or a notification. Requests reference a RTP stream that is being sent by the message recipient, whereas notifications are responses to an earlier request, and therefore reference a RTP stream that is being received by the message recipient.

If the RTCP packet is a feedback request that includes target SSRC(s), for each target SSRC that is found in the outgoing SSRC table, deliver a copy of the RTCP packet to the "m=" section associated with that SSRC. PSFB and RTPFB types that are handled in this way include:

Full Intra Request (FIR): [RFC5104] (PT=PSFB, FMT=4).

Temporal-Spatial Trade-off Request (TSTR): [RFC5104] (PT=PSFB,
FMT=5).

H.271 Video Back Channel Message (VBCM): [RFC5104] (PT=PSFB,
FMT=7).

Temporary Maximum Media Bit Rate Request (TMMBR): [RFC5104]
(PT=RTPFB, FMT=3).

Layer Refresh Request (LRR): [I-D.ietf-avtext-lrr] (PT=PSFB,
FMT=10).

If the RTCP packet is a feedback notification that includes target SSRC(s), for each target SSRC that is found in the incoming SSRC table, deliver a copy of the RTCP packet to the "m=" section associated with the RTP stream with matching SSRC. PSFB and RTPFB types that are handled in this way include:

Temporal-Spatial Trade-off Notification (TSTN): [RFC5104]
(PT=PSFB, FMT=6). This message is a notification in response to a prior TSTR.

Temporary Maximum Media Bit Rate Notification (TMMBN): [RFC5104]
(PT=RTPFB, FMT=4). This message is a notification in response to a prior TMMBR, but can also be sent unsolicited.

If the RTCP packet is of type APP, then it is handled in an application specific manner. If the application does not recognise the APP packet, then it MUST be discarded.

9.3. RTP/RTCP Multiplexing

Within a BUNDLE group, the offerer and answerer MUST enable RTP/RTCP multiplexing [RFC5761] for the RTP-based bundled media (i.e., the same transport will be used for both RTP packets and RTCP packets). In addition, the offerer and answerer MUST support the SDP 'rtcp-mux-only' attribute [I-D.ietf-mmusic-mux-exclusive].

9.3.1. SDP Offer/Answer Procedures

This section describes how an offerer and answerer use the SDP 'rtcp-mux' attribute [RFC5761] and the SDP 'rtcp-mux-only' attribute [I-D.ietf-mmusic-mux-exclusive] to negotiate usage of RTP/RTCP multiplexing for RTP-based bundled media.

RTP/RTCP multiplexing only applies to RTP-based media. However, as described in Section 7.1.3, within an offer or answer the SDP 'rtcp-mux' and SDP 'rtcp-mux-only' attributes might be included in a bundled "m=" section for non-RTP-based media (if such "m=" section is the offerer tagged "m=" section or answerer tagged "m=" section).

9.3.1.1. Generating the Initial SDP BUNDLE Offer

When an offerer generates an initial BUNDLE offer, if the offer contains one or more bundled "m=" sections for RTP-based media (or, if there is a chance that "m=" sections for RTP-based media will later be added to the BUNDLE group), the offerer MUST include an SDP 'rtcp-mux' attribute [RFC5761] in each bundled "m=" section (excluding any bundle-only "m=" sections). In addition, the offerer MAY include an SDP 'rtcp-mux-only' attribute [I-D.ietf-mmusic-mux-exclusive] in one or more bundled "m=" sections for RTP-based media.

NOTE: Whether the offerer includes the SDP 'rtcp-mux-only' attribute depends on whether the offerer supports fallback to usage of a separate port for RTCP in case the answerer moves one or more "m=" sections for RTP-based media out of the BUNDLE group in the answer.

NOTE: If the offerer includes an SDP 'rtcp-mux' attribute in the bundled "m=" sections, but does not include an SDP 'rtcp-mux-only' attribute, the offerer can also include an SDP 'rtcp' attribute [RFC3605] in one or more RTP-based bundled "m=" sections in order to provide a fallback port for RTCP, as described in [RFC5761]. However, the fallback port will only be applied to "m=" sections for

RTP-based media that are moved out of the BUNDLE group by the answerer.

In the initial BUNDLE offer, the address:port combination for RTCP MUST be unique in each bundled "m=" section for RTP-based media (excluding a bundle-only "m=" section), similar to RTP.

9.3.1.2. Generating the SDP Answer

When an answerer generates an answer, if the answerer supports RTP-based media, and if a bundled "m=" section in the corresponding offer contained an SDP 'rtcp-mux' attribute, the answerer MUST enable usage of RTP/RTCP multiplexing, even if there currently are no bundled "m=" sections for RTP-based media within the BUNDLE group. The answerer MUST include an SDP 'rtcp-mux' attribute in the answerer tagged "m=" section, following the procedures for BUNDLE attributes [Section 7.1.3]. In addition, if the "m=" section that is selected as the offerer tagged "m=" section contained an SDP "rtcp-mux-only" attribute, the answerer MUST include an SDP "rtcp-mux-only" attribute in the answerer tagged "m=" section.

In an initial BUNDLE offer, if the suggested offerer tagged "m=" section contained an SDP 'rtcp-mux-only' attribute, the "m=" section was for RTP-based media, and the answerer does not accept the "m=" section in the created BUNDLE group, the answerer MUST either move the "m=" section out of the BUNDLE group [Section 7.3.2], include the attribute in the moved "m=" section and enable RTP/RTCP multiplexing for the media associated with the "m=" section, or reject the "m=" section [Section 7.3.3].

The answerer MUST NOT include an SDP 'rtcp' attribute in any bundled "m=" section in the answer. The answerer will use the port value of the tagged offerer "m=" section sending RTP and RTCP packets associated with RTP-based bundled media towards the offerer.

If the usage of RTP/RTCP multiplexing within a BUNDLE group has been negotiated in a previous offer/answer exchange, the answerer MUST include an SDP 'rtcp-mux' attribute in the answerer tagged "m=" section. It is not possible to disable RTP/RTCP multiplexing within a BUNDLE group.

9.3.1.3. Offerer Processing of the SDP Answer

When an offerer receives an answer, if the answerer has accepted the usage of RTP/RTCP multiplexing [Section 9.3.1.2], the answerer follows the procedures for RTP/RTCP multiplexing defined in [RFC5761]. The offerer will use the port value of the answerer

tagged "m=" section for sending RTP and RTCP packets associated with RTP-based bundled media towards the answerer.

NOTE: It is considered a protocol error if the answerer has not accepted the usage of RTP/RTCP multiplexing for RTP-based "m=" sections that the answerer included in the BUNDLE group.

9.3.1.4. Modifying the Session

When an offerer generates a subsequent offer, the offerer MUST include an SDP 'rtcp-mux' attribute in the offerer tagged "m=" section, following the procedures for IDENTICAL multiplexing category attributes in Section 7.1.3.

10. ICE Considerations

This section describes how to use the BUNDLE grouping extension together with the Interactive Connectivity Establishment (ICE) mechanism [I-D.ietf-ice-rfc5245bis].

The generic procedures for negotiating usage of ICE using SDP, defined in [I-D.ietf-mmusic-ice-sip-sdp], also apply to usage of ICE with BUNDLE, with the following exceptions:

- o When the BUNDLE transport has been established, ICE connectivity checks and keep-alives only need to be performed for the BUNDLE transport, instead of per individual bundled "m=" section within the BUNDLE group.
- o The generic SDP attribute offer/answer considerations [Section 7.1.3] also apply to ICE-related attributes. Therefore, when an offer sends an initial BUNDLE offer (in order to negotiate a BUNDLE group) the offerer include ICE-related media-level attributes in each bundled "m=" section (excluding any bundle-only "m=" section), and each "m=" section MUST contain unique ICE properties. When an answerer generates an answer (initial BUNDLE answer or subsequent) that contains a BUNDLE group, and when an offerer sends a subsequent offer that contains a BUNDLE group, ICE-related media-level attributes are only included in the tagged "m=" section (suggested offerer tagged "m=" section or answerer tagged "m=" section), and the ICE properties are applied to each bundled "m=" section within the BUNDLE group.

NOTE: Most ICE-related media-level SDP attributes belong to the TRANSPORT multiplexing category [I-D.ietf-mmusic-sdp-mux-attributes], and the generic SDP attribute offer/answer considerations for TRANSPORT multiplexing category apply to the attributes. However, in the case of ICE-related attributes, the same considerations also

apply to ICE-related media-level attributes that belong to other multiplexing categories.

NOTE: The following ICE-related media-level SDP attributes are defined in [I-D.ietf-mmusic-ice-sip-sdp]: 'candidate', 'remote-candidates', 'ice-mismatch', 'ice-ufrag', 'ice-pwd', and 'ice-pacing'.

Initially, before ICE has produced selected candidate pairs that will be used for media, there might be multiple transports established (if multiple candidate pairs are tested). Once ICE has selected candidate pairs, they form the BUNDLE transport.

Support and usage of ICE mechanism together with the BUNDLE extension is OPTIONAL, and the procedures in this section only apply when the ICE mechanism is used. Note that applications might mandate usage of the ICE mechanism even if the BUNDLE extension is not used.

NOTE: If the trickle ICE mechanism [I-D.ietf-mmusic-trickle-ice-sip] is used, an offerer and answerer might assign a port value of '9', and an IPv4 address of '0.0.0.0' (or, the IPv6 equivalent ':::') to multiple bundled "m=" sections in the initial BUNDLE offer. The offerer and answerer will follow the normal procedures for generating the offers and answers, including picking a bundled "m=" section as the suggested offerer tagged "m=" section, selecting the tagged "m=" sections etc. The only difference is that media can not be sent until one or more candidates have been provided. Once a BUNDLE group has been negotiated, trickled candidates associated with a bundled "m=" section will be applied to all bundled "m=" sections within the BUNDLE group.

11. DTLS Considerations

One or more media streams within a BUNDLE group might use the Datagram Transport Layer Security (DTLS) protocol [RFC6347] in order to encrypt the data, or to negotiate encryption keys if another encryption mechanism is used to encrypt media.

When DTLS is used within a BUNDLE group, the following rules apply:

- o There can only be one DTLS association [RFC6347] associated with the BUNDLE group; and
- o Each usage of the DTLS association within the BUNDLE group MUST use the same mechanism for determining which endpoints (the offerer or answerer) become DTLS client and DTLS server; and

- o Each usage of the DTLS association within the BUNDLE group MUST use the same mechanism for determining whether an offer or answer will trigger the establishment of a new DTLS association, or whether an existing DTLS association will be used; and
- o If the DTLS client supports DTLS-SRTP [RFC5764] it MUST include the 'use_srtp' extension [RFC5764] in the DTLS ClientHello message [RFC5764]. The client MUST include the extension even if the usage of DTLS-SRTP is not negotiated as part of the multimedia session (e.g., SIP session [RFC3261]).

NOTE: The inclusion of the 'use_srtp' extension during the initial DTLS handshake ensures that a DTLS renegotiation will not be required in order to include the extension, in case DTLS-SRTP encrypted media is added to the BUNDLE group later during the multimedia session.

12. RTP Header Extensions Consideration

When [RFC8285] RTP header extensions are used in the context of this specification, the identifier used for a given extension MUST identify the same extension across all the bundled media descriptions.

13. Update to RFC 3264

This section updates RFC 3264, in order to allow extensions to define the usage of a zero port value in offers and answers for other purposes than removing or disabling media streams. The following sections of RFC 3264 are updated:

- o Section 5.1 (Unicast Streams).
- o Section 8.4 (Putting a Unicast Media Stream on Hold).

13.1. Original text of section 5.1 (2nd paragraph) of RFC 3264

For recvonly and sendrecv streams, the port number and address in the offer indicate where the offerer would like to receive the media stream. For sendonly RTP streams, the address and port number indirectly indicate where the offerer wants to receive RTCP reports. Unless there is an explicit indication otherwise, reports are sent to the port number one higher than the number indicated. The IP address and port present in the offer indicate nothing about the source IP address and source port of RTP and RTCP packets that will be sent by the offerer. A port number of zero in the offer indicates that the stream is offered but MUST NOT be used. This has no useful semantics in an initial offer, but is allowed for reasons of completeness, since the answer can contain a zero port indicating a rejected stream

(Section 6). Furthermore, existing streams can be terminated by setting the port to zero (Section 8). In general, a port number of zero indicates that the media stream is not wanted.

13.2. New text replacing section 5.1 (2nd paragraph) of RFC 3264

For `recvonly` and `sendrecv` streams, the port number and address in the offer indicate where the offerer would like to receive the media stream. For `sendonly` RTP streams, the address and port number indirectly indicate where the offerer wants to receive RTCP reports. Unless there is an explicit indication otherwise, reports are sent to the port number one higher than the number indicated. The IP address and port present in the offer indicate nothing about the source IP address and source port of RTP and RTCP packets that will be sent by the offerer. A port number of zero in the offer by default indicates that the stream is offered but **MUST NOT** be used, but an extension mechanism might specify different semantics for the usage of a zero port value. Furthermore, existing streams can be terminated by setting the port to zero (Section 8). In general, a port number of zero by default indicates that the media stream is not wanted.

13.3. Original text of section 8.4 (6th paragraph) of RFC 3264

RFC 2543 [10] specified that placing a user on hold was accomplished by setting the connection address to 0.0.0.0. Its usage for putting a call on hold is no longer recommended, since it doesn't allow for RTCP to be used with held streams, doesn't work with IPv6, and breaks with connection oriented media. However, it can be useful in an initial offer when the offerer knows it wants to use a particular set of media streams and formats, but doesn't know the addresses and ports at the time of the offer. Of course, when used, the port number **MUST NOT** be zero, which would specify that the stream has been disabled. An agent **MUST** be capable of receiving SDP with a connection address of 0.0.0.0, in which case it means that neither RTP nor RTCP is to be sent to the peer.

13.4. New text replacing section 8.4 (6th paragraph) of RFC 3264

RFC 2543 [10] specified that placing a user on hold was accomplished by setting the connection address to 0.0.0.0. Its usage for putting a call on hold is no longer recommended, since it doesn't allow for RTCP to be used with held streams, doesn't work with IPv6, and breaks with connection oriented media. However, it can be useful in an initial offer when the offerer knows it wants to use a particular set of media streams and formats, but doesn't know the addresses and ports at the time of the offer. Of course, when used, the port number **MUST NOT** be zero, if it would specify that the stream has been disabled. However, an extension mechanism might specify different

semantics of the zero port number usage. An agent MUST be capable of receiving SDP with a connection address of 0.0.0.0, in which case it means that neither RTP nor RTCP is to be sent to the peer.

14. Update to RFC 5888

This section updates RFC 5888 [RFC5888]), in order to allow extensions to allow an SDP 'group' attribute containing an identification-tag that identifies a "m=" section with the port set to zero Section 9.2 (Group Value in Answers) of RFC 5888 is updated.

14.1. Original text of section 9.2 (3rd paragraph) of RFC 5888

SIP entities refuse media streams by setting the port to zero in the corresponding "m" line. "a=group" lines MUST NOT contain identification-tags that correspond to "m" lines with the port set to zero.

14.2. New text replacing section 9.2 (3rd paragraph) of RFC 5888

SIP entities refuse media streams by setting the port to zero in the corresponding "m" line. "a=group" lines MUST NOT contain identification-tags that correspond to "m" lines with the port set to zero, but an extension mechanism might specify different semantics for including identification-tags that correspond to such "m=" lines.

15. RTP/RTCP extensions for identification-tag transport

SDP Offerers and Answerers [RFC3264] can associate identification-tags with "m=" sections within SDP Offers and Answers, using the procedures in [RFC5888]. Each identification-tag uniquely represents an "m=" section.

This section defines a new RTCP SDES item [RFC3550], 'MID', which is used to carry identification-tags within RTCP SDES packets. This section also defines a new RTP SDES header extension [RFC7941], which is used to carry the 'MID' RTCP SDES item in RTP packets.

The SDES item and RTP SDES header extension make it possible for a receiver to associate each RTP stream with a specific "m=" section, with which the receiver has associated an identification-tag, even if those "m=" sections are part of the same RTP session. The RTP SDES header extension also ensures that the media recipient gets the identification-tag upon receipt of the first decodable media and is able to associate the media with the correct application.

A media recipient informs the media sender about the identification-tag associated with an "m=" section through the use of an 'mid'

attribute [RFC5888]. The media sender then inserts the identification-tag in RTCP and RTP packets sent to the media recipient.

NOTE: This text above defines how identification-tags are carried in SDP Offers and Answers. The usage of other signaling protocols for carrying identification-tags is not prevented, but the usage of such protocols is outside the scope of this document.

[RFC3550] defines general procedures regarding the RTCP transmission interval. The RTCP MID SDES item SHOULD be sent in the first few RTCP packets sent after joining the session, and SHOULD be sent regularly thereafter. The exact number of RTCP packets in which this SDES item is sent is intentionally not specified here, as it will depend on the expected packet loss rate, the RTCP reporting interval, and the allowable overhead.

The RTP SDES header extension for carrying the 'MID' RTCP SDES SHOULD be included in some RTP packets at the start of the session and whenever the SSRC changes. It might also be useful to include the header extension in RTP packets that comprise access points in the media (e.g., with video I-frames). The exact number of RTP packets in which this header extension is sent is intentionally not specified here, as it will depend on expected packet loss rate and loss patterns, the overhead the application can tolerate, and the importance of immediate receipt of the identification-tag.

For robustness, endpoints need to be prepared for situations where the reception of the identification-tag is delayed, and SHOULD NOT terminate sessions in such cases, as the identification-tag is likely to arrive soon.

15.1. RTCP MID SDES Item

```

      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
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
|          MID=TBD          |      length      | identification-tag  |...
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+

```

The identification-tag payload is UTF-8 encoded [RFC3629], as in SDP.

The identification-tag is not zero terminated.

[RFC EDITOR NOTE: Please replace TBD with the assigned SDES identifier value.]

15.2. RTP SDES Header Extension For MID

The payload, containing the identification-tag, of the RTP SDES header extension element can be encoded using either the one-byte or two-byte header [RFC7941]. The identification-tag payload is UTF-8 encoded, as in SDP.

The identification-tag is not zero terminated. Note, that the set of header extensions included in the packet needs to be padded to the next 32-bit boundary using zero bytes [RFC8285].

As the identification-tag is included in either an RTCP SDES item or an RTP SDES header extension, or both, there needs to be some consideration about the packet expansion caused by the identification-tag. To avoid Maximum Transmission Unit (MTU) issues for the RTP packets, the header extension's size needs to be taken into account when encoding the media.

It is recommended that the identification-tag is kept short. Due to the properties of the RTP header extension mechanism, when using the one-byte header, a tag that is 1-3 bytes will result in a minimal number of 32-bit words used for the RTP SDES header extension, in case no other header extensions are included at the same time. Note, do take into account that some single characters when UTF-8 encoded will result in multiple octets. The identification-tag MUST NOT contain any user information, and applications SHALL avoid generating the identification-tag using a pattern that enables user- or application identification.

16. IANA Considerations

16.1. New SDES item

[RFC EDITOR NOTE: Please replace RFCXXXX with the RFC number of this document.]

[RFC EDITOR NOTE: Please replace TBD with the assigned SDES identifier value.]

This document adds the MID SDES item to the IANA "RTP SDES item types" registry as follows:

Value:	TBD
Abbrev.:	MID
Name:	Media Identification
Reference:	RFCXXXX

16.2. New RTP SDES Header Extension URI

[RFC EDITOR NOTE: Please replace RFCXXXX with the RFC number of this document.]

This document defines a new extension URI in the RTP SDES Compact Header Extensions sub-registry of the RTP Compact Header Extensions registry sub-registry, according to the following data:

Extension URI: urn:ietf:params:rtp-hdext:sdes:mid
Description: Media identification
Contact: IESG (iesg@ietf.org)
Reference: RFCXXXX

The SDES item does not reveal privacy information about the users. It is simply used to associate RTP-based media with the correct SDP media description ("m=" section) in the SDP used to negotiate the media.

The purpose of the extension is for the offerer to be able to associate received multiplexed RTP-based media before the offerer receives the associated SDP answer.

16.3. New SDP Attribute

[RFC EDITOR NOTE: Please replace RFCXXXX with the RFC number of this document.]

This document defines a new SDP media-level attribute, 'bundle-only', according to the following data:

Attribute name: bundle-only
Type of attribute: media
Subject to charset: No
Purpose: Request a media description to be accepted in the answer only if kept within a BUNDLE group by the answerer.
Appropriate values: N/A
Contact name: IESG
Contact e-mail: iesg@ietf.org
Reference: RFCXXXX
Mux category: NORMAL

16.4. New SDP Group Semantics

[RFC EDITOR NOTE: Please replace RFCXXXX with the RFC number of this document.]

This document registers the following semantics with IANA in the "Semantics for the "group" SDP Attribute" subregistry (under the "Session Description Protocol (SDP) Parameters" registry:

Semantics	Token	Reference
Media bundling	BUNDLE	[RFCXXXX]

Mux category: NORMAL

17. Security Considerations

The security considerations defined in [RFC3264] and [RFC5888] apply to the BUNDLE extension. Bundle does not change which information, e.g., RTP streams, flows over the network, with the exception of the usage of the MID SDES item as discussed below. Primarily it changes which addresses and ports, and thus in which (RTP) sessions the information is flowing. This affects the security contexts being used and can cause previously separated information flows to share the same security context. This has very little impact on the performance of the security mechanism of the RTP sessions. In cases where one would have applied different security policies on the different RTP streams being bundled, or where the parties having access to the security contexts would have differed between the RTP streams, additional analysis of the implications are needed before selecting to apply BUNDLE.

The identification-tag, independent of transport, RTCP SDES packet or RTP header extension, can expose the value to parties beyond the signaling chain. Therefore, the identification-tag values MUST be generated in a fashion that does not leak user information, e.g., randomly or using a per-bundle group counter, and SHOULD be 3 bytes or less, to allow them to efficiently fit into the MID RTP header extension. Note that if implementations use different methods for generating identification-tags this could enable fingerprinting of the implementation making it vulnerable to targeted attacks. The identification-tag is exposed on the RTP stream level when included in the RTP header extensions, however what it reveals of the RTP media stream structure of the endpoint and application was already possible to deduce from the RTP streams without the MID SDES header

extensions. As the identification-tag is also used to route the media stream to the right application functionality it is important that the value received is the one intended by the sender, thus integrity and the authenticity of the source are important to prevent denial of service on the application. Existing SRTP configurations and other security mechanisms protecting the whole RTP/RTCP packets will provide the necessary protection.

When the BUNDLE extension is used, the set of configurations of the security mechanism used in all the bundled media descriptions will need to be compatible so that they can be used simultaneously, at least per direction or endpoint. When using SRTP this will be the case, at least for the IETF defined key-management solutions due to their SDP attributes (a=crypto, a=fingerprint, a=mikey) and their classification in [I-D.ietf-mmusic-sdp-mux-attributes].

The security considerations of "RTP Header Extension for the RTP Control Protocol (RTCP) Source Description Items" [RFC7941] requires that when RTCP is confidentiality protected, then any SDES RTP header extension carrying an SDES item, such as the MID RTP header extension, is also protected using commensurate strength algorithms. However, assuming the above requirements and recommendations are followed, there are no known significant security risks with leaving the MID RTP header extension without confidentiality protection. Therefore, this specification updates RFC 7941 by adding the exception that this requirement MAY be ignored for the MID RTP header extension. Security mechanisms for RTP/RTCP are discussed in Options for Securing RTP Sessions [RFC7201], for example SRTP [RFC3711] can provide the necessary security functions of ensuring the integrity and source authenticity.

18. Examples

18.1. Example: Tagged m= Section Selections

The example below shows:

- o An initial BUNDLE offer, in which the offerer wants to negotiate a BUNDLE group, and indicates the audio m= section as the suggested offerer tagged "m=" section.
- o An initial BUNDLE answer, in which the answerer accepts the creation of the BUNDLE group, selects the audio m= section in the offer as the offerer tagged "m=" section, selects the audio "m=" section in the answer as the answerer tagged "m=" section and assigns the answerer BUNDLE address:port to that "m=" section.

SDP Offer (1)

```
v=0
o=alice 2890844526 2890844526 IN IP6 2001:db8::3
s=
c=IN IP6 2001:db8::3
t=0 0
a=group:BUNDLE foo bar

m=audio 10000 RTP/AVP 0 8 97
b=AS:200
a=mid:foo
a=rtcp-mux
a=rtpmap:0 PCMU/8000
a=rtpmap:8 PCMA/8000
a=rtpmap:97 iLBC/8000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid

m=video 10002 RTP/AVP 31 32
b=AS:1000
a=mid:bar
a=rtcp-mux
a=rtpmap:31 H261/90000
a=rtpmap:32 MPV/90000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid
```

SDP Answer (2)

```
v=0
o=bob 2808844564 2808844564 IN IP6 2001:db8::1
s=
c=IN IP6 2001:db8::1
t=0 0
a=group:BUNDLE foo bar

m=audio 20000 RTP/AVP 0
b=AS:200
a=mid:foo
a=rtcp-mux
a=rtpmap:0 PCMU/8000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid

m=video 0 RTP/AVP 32
b=AS:1000
a=mid:bar
a=bundle-only
a=rtpmap:32 MPV/90000
```

```
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid
```

18.2. Example: BUNDLE Group Rejected

The example below shows:

- o An initial BUNDLE offer, in which the offerer wants to negotiate a BUNDLE group, and indicates the audio m= section as the suggested offerer tagged "m=" section.
- o An initial BUNDLE answer, in which the answerer rejects the creation of the BUNDLE group, generates a normal answer and assigns a unique address:port to each "m=" section in the answer.

SDP Offer (1)

```
v=0
o=alice 2890844526 2890844526 IN IP6 2001:db8::3
s=
c=IN IP6 2001:db8::3
t=0 0
a=group:BUNDLE foo bar

m=audio 10000 RTP/AVP 0 8 97
b=AS:200
a=mid:foo
a=rtcp-mux
a=rtpmap:0 PCMU/8000
a=rtpmap:8 PCMA/8000
a=rtpmap:97 iLBC/8000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid

m=video 10002 RTP/AVP 31 32
b=AS:1000
a=mid:bar
a=rtcp-mux
a=rtpmap:31 H261/90000
a=rtpmap:32 MPV/90000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid
```

SDP Answer (2)

```
v=0
o=bob 2808844564 2808844564 IN IP6 2001:db8::1
s=
c=IN IP6 2001:db8::1
t=0 0

m=audio 20000 RTP/AVP 0
b=AS:200
a=rtcp-mux
a=rtpmap:0 PCMU/8000

m=video 30000 RTP/AVP 32
b=AS:1000
a=rtcp-mux
a=rtpmap:32 MPV/90000
```


18.3. Example: Offerer Adds a Media Description to a BUNDLE Group

The example below shows:

- o A subsequent offer, in which the offerer adds a new bundled "m=" section (video), indicated by the "zen" identification-tag, to a previously negotiated BUNDLE group, indicates the new "m=" section as the offerer tagged "m=" section and assigns the offerer BUNDLE address:port to that "m=" section.
- o A subsequent answer, in which the answerer indicates the new video "m=" section in the answer as the answerer tagged "m=" section and assigns the answerer BUNDLE address:port to that "m=" section.

SDP Offer (1)

```
v=0
o=alice 2890844526 2890844526 IN IP6 2001:db8::3
s=
c=IN IP6 2001:db8::3
t=0 0
a=group:BUNDLE zen foo bar

m=audio 0 RTP/AVP 0 8 97
b=AS:200
a=mid:foo
a=bundle-only
a=rtpmap:0 PCMU/8000
a=rtpmap:8 PCMA/8000
a=rtpmap:97 iLBC/8000
a=extmap:1 urn:ietf:params:rtp-hdext:sdes:mid

m=video 0 RTP/AVP 31 32
b=AS:1000
a=mid:bar
a=bundle-only
a=rtpmap:31 H261/90000
a=rtpmap:32 MPV/90000
a=extmap:1 urn:ietf:params:rtp-hdext:sdes:mid

m=video 10000 RTP/AVP 66
b=AS:1000
a=mid:zen
a=rtcp-mux
a=rtpmap:66 H261/90000
a=extmap:1 urn:ietf:params:rtp-hdext:sdes:mid
```

SDP Answer (2)

```
v=0
o=bob 2808844564 2808844564 IN IP6 2001:db8::1
s=
c=IN IP6 2001:db8::1
t=0 0
a=group:BUNDLE zen foo bar

m=audio 0 RTP/AVP 0
b=AS:200
a=mid:foo
a=bundle-only
a=rtpmap:0 PCMU/8000
a=extmap:1 urn:ietf:params:rtp-hdext:sdes:mid

m=video 0 RTP/AVP 32
b=AS:1000
a=mid:bar
a=bundle-only
a=rtpmap:32 MPV/90000
a=extmap:1 urn:ietf:params:rtp-hdext:sdes:mid

m=video 20000 RTP/AVP 66
b=AS:1000
a=mid:zen
a=rtcp-mux
a=rtpmap:66 H261/90000
a=extmap:1 urn:ietf:params:rtp-hdext:sdes:mid
```

18.4. Example: Offerer Moves a Media Description Out of a BUNDLE Group

The example below shows:

- o A subsequent offer, in which the offerer removes a "m=" section (video), indicated by the "zen" identification-tag, from a previously negotiated BUNDLE group, indicates one of the bundled "m=" sections (audio) remaining in the BUNDLE group as the offerer tagged "m=" section and assigns the offerer BUNDLE address:port to that "m=" section.
- o A subsequent answer, in which the answerer removes the "m=" section from the BUNDLE group, indicates the audio "m=" section in the answer as the answerer tagged "m=" section and assigns the answerer BUNDLE address:port to that "m=" section.

SDP Offer (1)

```
v=0
o=alice 2890844526 2890844526 IN IP6 2001:db8::3
s=
c=IN IP6 2001:db8::3
t=0 0
a=group:BUNDLE foo bar

m=audio 10000 RTP/AVP 0 8 97
b=AS:200
a=mid:foo
a=rtcp-mux
a=rtpmap:0 PCMU/8000
a=rtpmap:8 PCMA/8000
a=rtpmap:97 iLBC/8000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid

m=video 0 RTP/AVP 31 32
b=AS:1000
a=mid:bar
a=bundle-only
a=rtpmap:31 H261/90000
a=rtpmap:32 MPV/90000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid

m=video 50000 RTP/AVP 66
b=AS:1000
a=mid:zen
a=rtcp-mux
a=rtpmap:66 H261/90000
```

SDP Answer (2)

```
v=0
o=bob 2808844564 2808844564 IN IP6 2001:db8::1
s=
c=IN IP6 2001:db8::1
t=0 0
a=group:BUNDLE foo bar

m=audio 20000 RTP/AVP 0
b=AS:200
a=mid:foo
a=rtcp-mux
a=rtpmap:0 PCMU/8000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid
```

```
m=video 0 RTP/AVP 32
b=AS:1000
a=mid:bar
a=bundle-only
a=rtpmap:32 MPV/90000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid
```

```
m=video 60000 RTP/AVP 66
b=AS:1000
a=mid:zen
a=rtcp-mux
a=rtpmap:66 H261/90000
```

18.5. Example: Offerer Disables a Media Description Within a BUNDLE Group

The example below shows:

- o A subsequent offer, in which the offerer disables (by assigning a zero port value) a "m=" section (video), indicated by the "zen" identification-tag, from a previously negotiated BUNDLE group, indicates one of the bundled "m=" sections (audio) remaining active in the BUNDLE group as the offerer tagged "m=" section and assigns the offerer BUNDLE address:port to that "m=" section.
- o A subsequent answer, in which the answerer disables the "m=" section, indicates the audio "m=" section in the answer as the answerer tagged "m=" section and assigns the answerer BUNDLE address:port to that "m=" section.

SDP Offer (1)

```
v=0
o=alice 2890844526 2890844526 IN IP6 2001:db8::3
s=
t=0 0
a=group:BUNDLE foo bar

m=audio 10000 RTP/AVP 0 8 97
c=IN IP6 2001:db8::3
b=AS:200
a=mid:foo
a=rtcp-mux
a=rtpmap:0 PCMU/8000
a=rtpmap:8 PCMA/8000
```

```
a=rtpmap:97 iLBC/8000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid

m=video 0 RTP/AVP 31 32
c=IN IP6 2001:db8::3
b=AS:1000
a=mid:bar
a=bundle-only
a=rtpmap:31 H261/90000
a=rtpmap:32 MPV/90000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid

m=video 0 RTP/AVP 66
a=mid:zen
a=rtpmap:66 H261/90000
```

SDP Answer (2)

```
v=0
o=bob 2808844564 2808844564 IN IP6 2001:db8::1
s=
t=0 0
a=group:BUNDLE foo bar

m=audio 20000 RTP/AVP 0
c=IN IP6 2001:db8::1
b=AS:200
a=mid:foo
a=rtcp-mux
a=rtpmap:0 PCMU/8000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid

m=video 0 RTP/AVP 32
c=IN IP6 2001:db8::1
b=AS:1000
a=mid:bar
a=bundle-only
a=rtpmap:32 MPV/90000
a=extmap:1 urn:ietf:params:rtp-hdrext:sdes:mid

m=video 0 RTP/AVP 66
a=mid:zen
a=rtpmap:66 H261/90000
```

19. Acknowledgements

The usage of the SDP grouping extension for negotiating bundled media is based on similar alternatives proposed by Harald Alvestrand and Cullen Jennings. The BUNDLE extension described in this document is based on the different alternative proposals, and text (e.g., SDP examples) have been borrowed (and, in some cases, modified) from those alternative proposals.

The SDP examples are also modified versions from the ones in the Alvestrand proposal.

Thanks to Paul Kyzivat, Martin Thomson, Flemming Andreassen, Thomas Stach, Ari Keranen, Adam Roach, Christian Groves, Roman Shpount, Suhas Nandakumar, Nils Ohlmeier, Jens Guballa, Raju Makaraju, Justin Uberti, Taylor Brandstetter, Byron Campen and Eric Rescorla for reading the text, and providing useful feedback.

Thanks to Bernard Aboba, Peter Thatcher, Justin Uberti, and Magnus Westerlund for providing the text for the section on RTP/RTCP stream association.

Thanks to Magnus Westerlund, Colin Perkins and Jonathan Lennox for providing help and text on the RTP/RTCP procedures.

Thanks to Charlie Kaufman for performing the Sec-Dir review.

Thanks to Linda Dunbar for performing the Gen-ART review.

Thanks to Spotify for providing music for the countless hours of document editing.

20. Change Log

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

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-51

- o Changes based on IESG reviews.
- o - Clarification of 'initial offer' terminology.
- o - Merging of tagged m- section selection sections.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-50

- o Changes based on IESG reviews.

- o - Adding of tagged m- section concept.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-49

- o Changes based on IESG reviews.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-48

- o Changes based on Sec-Dir review by Charlie Kaufman.

- o - s/unique address/unique address:port

- o Changes based on Gen-ART review by Linda Dunbar.

- o Mux category for group:BUNDLE attribute added.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-47

- o Changes based on AD review by Ben Campbell.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-46

- o Pre-RFC5378 disclaimer removed put back.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-45

- o Mux category for SDP 'group:BUNDLE' attribute added.

- o <https://github.com/cdh4u/draft-sdp-bundle/pull/54>

- o Pre-RFC5378 disclaimer removed.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-44

- o Minor editorial nits based on pull request by Colin P.

- o <https://github.com/cdh4u/draft-sdp-bundle/pull/53>

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-43

- o Changes based on WG chairs review.

- o Text added in order to close GitHub issues by Taylor B.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-42

- o Changes based on final WG review.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-41

- o Update to section 6 o RFC 3264:
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/47>
- o Editorial clarification on BUNDLE address selection:
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/46>

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-40

- o Editorial changes and technical restrictions in order to make the specification more understandable:
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/45>
- o - BUNDLE address is only assigned to m- section indicated by BUNDLE-tag.
- o - bundle-only attribute also used in answers and subsequent offers.
- o - Answerer cannot reject, or remove, the bundled m- section that contains the BUNDLE address.
- o - ICE Offer/Answer sections removed, due to duplicated information.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-39

- o Editorial terminology changes.
- o RFC 5285 reference replaced by reference to RFC 8285.
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/44>
- o - Clarify that an m- section can not be moved between BUNDLE groups without first moving the m- section out of a BUNDLE group.
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/41>
- o - Addition of BUNDLE transport concept.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-38

- o Changes to RTP streaming mapping section based on text from Colin Perkins.

- o The following GitHub pull requests were merged:
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/34>
- o - Proposed updates to RTP processing
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/35>
- o - fixed reference to receiver-id section

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-37

- o The following GitHub pull request was merged:
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/33>

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-36

- o The following GitHub pull requests were merged:
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/32>
- o - extmap handling in BUNDLE.
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/31>
- o - Additional Acknowledgement text added.
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/30>
- o - MID SDES item security procedures updated
- o <https://github.com/cdh4u/draft-sdp-bundle/pull/29>
- o - Appendix B of JSEP moved into BUNDLE.
- o - Associating RTP/RTCP packets with SDP m- lines.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-35

- o Editorial changes on RTP streaming mapping section based on comments from Colin Perkins.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-34

- o RTP streams, instead of RTP packets, are associated with m- lines.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-33

- o Editorial changes based on comments from Eric Rescorla and Cullen Jennings:
- o - Changes regarding usage of RTP/RTCP multiplexing attributes.
- o - Additional text regarding associating RTP/RTCP packets with SDP m- lines.
- o - Reference correction.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-32

- o Editorial changes based on comments from Eric Rescorla and Cullen Jennings:
- o - Justification for mechanism added to Introduction.
- o - Clarify that the order of m- lines in the group:BUNDLE attribute does not have to be the same as the order in which the m- lines are listed in the SDP.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-31

- o Editorial changes based on GitHub Pull requests by Martin Thomson:
- o - <https://github.com/cdh4u/draft-sdp-bundle/pull/2>
- o - <https://github.com/cdh4u/draft-sdp-bundle/pull/1>
- o Editorial change based on comment from Diederick Huijbers (9th July 2016).
- o Changes based on comments from Flemming Andreassen (21st June 2016):
- o - Mux category for SDP bundle-only attribute added.
- o - Mux category considerations editorial clarification.
- o - Editorial changes.
- o RTP SDES extension according to draft-ietf-avtext-sdes-hdr-ext.
- o Note whether Design Considerations appendix is to be kept removed:
- o - Appendix is kept within document.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-30

- o Indicating in the Abstract and Introduction that the document updates RFC 3264.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-29

- o Change based on WGLC comment from Colin Perkins.
- o - Clarify that SSRC can be reused by another source after a delay of 5 RTCP reporting intervals.
- o Change based on WGLC comment from Alissa Cooper.
- o - IANA registry name fix.
- o - Additional IANA registration information added.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-28

- o - Alignment with exclusive mux procedures.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-27

- o - Yet another terminology change.
- o - Mux category considerations added.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-26

- o - ICE considerations modified: ICE-related SDP attributes only added to the bundled m- line representing the selected BUNDLE address.
- o - Reference to draft-ietf-mmusic-ice-sip-sdp added.
- o - Reference to RFC 5245 replaced with reference to draft-ietf-ice-rfc5245bis.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-25

- o - RTP/RTCP mux procedures updated with exclusive RTP/RTCP mux considerations.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-24

- o - Reference and procedures associated with exclusive RTP/RTCP mux added

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-23

- o - RTCP-MUX mandatory for bundled RTP m- lines
- o - Editorial fixes based on comments from Flemming Andreassen

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-22

- o - Correction of Ari's family name
- o - Editorial fixes based on comments from Thomas Stach
- o - RTP/RTCP correction based on comment from Magnus Westerlund
- o -- <http://www.ietf.org/mail-archive/web/mmusic/current/msg14861.html>

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-21

- o - Correct based on comment from Paul Kyzivat
- o -- 'received packets' replaced with 'received data'

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-20

- o - Clarification based on comment from James Guballa
- o - Clarification based on comment from Flemming Andreassen

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-19

- o - DTLS Considerations section added.
- o - BUNDLE semantics added to the IANA Considerations
- o - Changes based on WGLC comments from Adam Roach
- o -- <http://www.ietf.org/mail-archive/web/mmusic/current/msg14673.html>

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-18

- o - Changes based on agreements at IETF#92
- o -- BAS Offer removed, based on agreement at IETF#92.
- o -- Procedures regarding usage of SDP "b=" line is replaced with a reference to to draft-ietf-mmusic-sdp-mux-attributes.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-17

- o - Editorial changes based on comments from Magnus Westerlund.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-16

- o - Modification of RTP/RTCP multiplexing section, based on comments from Magnus Westerlund.
- o - Reference updates.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-15

- o - Editorial fix.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-14

- o - Editorial changes.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-13

- o Changes to allow a newly suggested offerer BUNDLE address to be assigned to each bundled m- line.
- o Changes based on WGLC comments from Paul Kyzivat
- o - Editorial fixes

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-12

- o Usage of SDP 'extmap' attribute added
- o SDP 'bundle-only' attribute scoped with "m=" lines with a zero port value
- o Changes based on WGLC comments from Thomas Stach
- o - ICE candidates not assigned to bundle-only m- lines with a zero port value
- o - Editorial changes
- o Changes based on WGLC comments from Colin Perkins
- o - Editorial changes:
 - o -- "RTP SDES item" -> "RTCP SDES item"
 - o -- "RTP MID SDES item" -> "RTCP MID SDES item"

- o - Changes in section 10.1.1:
- o -- "SHOULD NOT" -> "MUST NOT"
- o -- Additional text added to the Note
- o - Change to section 13.2:
- o -- Clarify that mid value is not zero terminated
- o - Change to section 13.3:
- o -- Clarify that mid value is not zero terminated
- o -- Clarify padding
- o Changes based on WGLC comments from Paul Kyzivat
- o - Editorial changes:
- o Changes based on WGLC comments from Jonathan Lennox
- o - Editorial changes:
- o - Definition of SDP bundle-only attribute aligned with structure in 4566bis draft

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-11

- o Editorial corrections based on comments from Harald Alvestrand.
- o Editorial corrections based on comments from Cullen Jennings.
- o Reference update (RFC 7160).
- o Clarification about RTCP packet sending when RTP/RTCP multiplexing is not used (<http://www.ietf.org/mail-archive/web/mmusic/current/msg13765.html>).
- o Additional text added to the Security Considerations.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-10

- o SDP bundle-only attribute added to IANA Considerations.
- o SDES item and RTP header extension added to Abstract and Introduction.

- o Modification to text updating section 8.2 of RFC 3264.
- o Reference corrections.
- o Editorial corrections.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-09

- o Terminology change: "bundle-only attribute assigned to m= line" to "bundle-only attribute associated with m= line".
- o Editorial corrections.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-08

- o Editorial corrections.
- o - "of"->"if" (8.3.2.5).
- o - "optional"->"OPTIONAL" (9.1).
- o - Syntax/ABNF for 'bundle-only' attribute added.
- o - SDP Offer/Answer sections merged.
- o - 'Request new offerer BUNDLE address' section added

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-07

- o OPEN ISSUE regarding Receiver-ID closed.
- o - RTP MID SDES Item.
- o - RTP MID Header Extension.
- o OPEN ISSUE regarding insertion of SDP 'rtcp' attribute in answers closed.
- o - Indicating that, when rtcp-mux is used, the answerer MUST NOT include an 'rtcp' attribute in the answer, based on the procedures in section 5.1.3 of RFC 5761.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-06

- o Draft title changed.
- o Added "SDP" to section names containing "Offer" or "Answer".

- o Editorial fixes based on comments from Paul Kyzivat (<http://www.ietf.org/mail-archive/web/mmusic/current/msg13314.html>).
- o Editorial fixed based on comments from Colin Perkins (<http://www.ietf.org/mail-archive/web/mmusic/current/msg13318.html>).
- o - Removed text about extending BUNDLE to allow multiple RTP sessions within a BUNDLE group.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-05

- o Major re-structure of SDP Offer/Answer sections, to align with RFC 3264 structure.
- o Additional definitions added.
- o - Shared address.
- o - Bundled "m=" line.
- o - Bundle-only "m=" line.
- o - Offerer suggested BUNDLE mid.
- o - Answerer selected BUNDLE mid.
- o Q6 Closed (IETF#88): An Offerer MUST NOT assign a shared address to multiple "m=" lines until it has received an SDP Answer indicating support of the BUNDLE extension.
- o Q8 Closed (IETF#88): An Offerer can, before it knows whether the Answerer supports the BUNDLE extension, assign a zero port value to a 'bundle-only' "m=" line.
- o SDP 'bundle-only' attribute section added.
- o Connection data nettype/addrtype restrictions added.
- o RFC 3264 update section added.
- o Indicating that a specific payload type value can be used in multiple "m=" lines, if the value represents the same codec configuration in each "m=" line.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-04

- o Updated Offerer procedures (<http://www.ietf.org/mail-archive/web/mmusic/current/msg12293.html>).
- o Updated Answerer procedures (<http://www.ietf.org/mail-archive/web/mmusic/current/msg12333.html>).
- o Usage of SDP 'bundle-only' attribute added.
- o Reference to Trickle ICE document added.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-02

- o Mechanism modified, to be based on usage of SDP Offers with both different and identical port number values, depending on whether it is known if the remote endpoint supports the extension.
- o Cullen Jennings added as co-author.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-01

- o No changes. New version due to expiration.

Changes from draft-ietf-mmusic-sdp-bundle-negotiation-00

- o No changes. New version due to expiration.

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.

21. References

21.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <<https://www.rfc-editor.org/info/rfc2119>>.

- [RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, DOI 10.17487/RFC3264, June 2002, <<https://www.rfc-editor.org/info/rfc3264>>.
- [RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, DOI 10.17487/RFC3550, July 2003, <<https://www.rfc-editor.org/info/rfc3550>>.
- [RFC3605] Huitema, C., "Real Time Control Protocol (RTCP) attribute in Session Description Protocol (SDP)", RFC 3605, DOI 10.17487/RFC3605, October 2003, <<https://www.rfc-editor.org/info/rfc3605>>.
- [RFC3629] Yergeau, F., "UTF-8, a transformation format of ISO 10646", STD 63, RFC 3629, DOI 10.17487/RFC3629, November 2003, <<https://www.rfc-editor.org/info/rfc3629>>.
- [RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, DOI 10.17487/RFC3711, March 2004, <<https://www.rfc-editor.org/info/rfc3711>>.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, DOI 10.17487/RFC4566, July 2006, <<https://www.rfc-editor.org/info/rfc4566>>.
- [RFC4961] Wing, D., "Symmetric RTP / RTP Control Protocol (RTCP)", BCP 131, RFC 4961, DOI 10.17487/RFC4961, July 2007, <<https://www.rfc-editor.org/info/rfc4961>>.
- [RFC5761] Perkins, C. and M. Westerlund, "Multiplexing RTP Data and Control Packets on a Single Port", RFC 5761, DOI 10.17487/RFC5761, April 2010, <<https://www.rfc-editor.org/info/rfc5761>>.
- [RFC5764] McGrew, D. and E. Rescorla, "Datagram Transport Layer Security (DTLS) Extension to Establish Keys for the Secure Real-time Transport Protocol (SRTP)", RFC 5764, DOI 10.17487/RFC5764, May 2010, <<https://www.rfc-editor.org/info/rfc5764>>.
- [RFC5888] Camarillo, G. and H. Schulzrinne, "The Session Description Protocol (SDP) Grouping Framework", RFC 5888, DOI 10.17487/RFC5888, June 2010, <<https://www.rfc-editor.org/info/rfc5888>>.

- [RFC6347] Rescorla, E. and N. Modadugu, "Datagram Transport Layer Security Version 1.2", RFC 6347, DOI 10.17487/RFC6347, January 2012, <<https://www.rfc-editor.org/info/rfc6347>>.
- [RFC7941] Westerlund, M., Burman, B., Even, R., and M. Zanaty, "RTP Header Extension for the RTP Control Protocol (RTCP) Source Description Items", RFC 7941, DOI 10.17487/RFC7941, August 2016, <<https://www.rfc-editor.org/info/rfc7941>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.
- [RFC8285] Singer, D., Desineni, H., and R. Even, Ed., "A General Mechanism for RTP Header Extensions", RFC 8285, DOI 10.17487/RFC8285, October 2017, <<https://www.rfc-editor.org/info/rfc8285>>.
- [I-D.ietf-ice-rfc5245bis]
Keranen, A., Holmberg, C., and J. Rosenberg, "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal", draft-ietf-ice-rfc5245bis-20 (work in progress), March 2018.
- [I-D.ietf-mmusic-sdp-mux-attributes]
Nandakumar, S., "A Framework for SDP Attributes when Multiplexing", draft-ietf-mmusic-sdp-mux-attributes-16 (work in progress), December 2016.
- [I-D.ietf-mmusic-mux-exclusive]
Holmberg, C., "Indicating Exclusive Support of RTP/RTCP Multiplexing using SDP", draft-ietf-mmusic-mux-exclusive-12 (work in progress), May 2017.
- [I-D.ietf-mmusic-ice-sip-sdp]
Petit-Huguenin, M., Nandakumar, S., and A. Keranen, "Session Description Protocol (SDP) Offer/Answer procedures for Interactive Connectivity Establishment (ICE)", draft-ietf-mmusic-ice-sip-sdp-20 (work in progress), April 2018.
- [I-D.ietf-mmusic-trickle-ice-sip]
Ivov, E., Stach, T., Marocco, E., and C. Holmberg, "A Session Initiation Protocol (SIP) Usage for Trickle ICE", draft-ietf-mmusic-trickle-ice-sip-14 (work in progress), February 2018.

21.2. Informative References

- [RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, DOI 10.17487/RFC3261, June 2002, <<https://www.rfc-editor.org/info/rfc3261>>.
- [RFC3611] Friedman, T., Ed., Caceres, R., Ed., and A. Clark, Ed., "RTP Control Protocol Extended Reports (RTCP XR)", RFC 3611, DOI 10.17487/RFC3611, November 2003, <<https://www.rfc-editor.org/info/rfc3611>>.
- [RFC5104] Wenger, S., Chandra, U., Westerlund, M., and B. Burman, "Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)", RFC 5104, DOI 10.17487/RFC5104, February 2008, <<https://www.rfc-editor.org/info/rfc5104>>.
- [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, DOI 10.17487/RFC4585, July 2006, <<https://www.rfc-editor.org/info/rfc4585>>.
- [RFC5576] Lennox, J., Ott, J., and T. Schierl, "Source-Specific Media Attributes in the Session Description Protocol (SDP)", RFC 5576, DOI 10.17487/RFC5576, June 2009, <<https://www.rfc-editor.org/info/rfc5576>>.
- [RFC7160] Petit-Huguenin, M. and G. Zorn, Ed., "Support for Multiple Clock Rates in an RTP Session", RFC 7160, DOI 10.17487/RFC7160, April 2014, <<https://www.rfc-editor.org/info/rfc7160>>.
- [RFC7201] Westerlund, M. and C. Perkins, "Options for Securing RTP Sessions", RFC 7201, DOI 10.17487/RFC7201, April 2014, <<https://www.rfc-editor.org/info/rfc7201>>.
- [RFC7656] Lennox, J., Gross, K., Nandakumar, S., Salgueiro, G., and B. Burman, Ed., "A Taxonomy of Semantics and Mechanisms for Real-Time Transport Protocol (RTP) Sources", RFC 7656, DOI 10.17487/RFC7656, November 2015, <<https://www.rfc-editor.org/info/rfc7656>>.
- [RFC7657] Black, D., Ed. and P. Jones, "Differentiated Services (Diffserv) and Real-Time Communication", RFC 7657, DOI 10.17487/RFC7657, November 2015, <<https://www.rfc-editor.org/info/rfc7657>>.

[I-D.ietf-ice-trickle]

Ivov, E., Rescorla, E., Uberti, J., and P. Saint-Andre,
"Trickle ICE: Incremental Provisioning of Candidates for
the Interactive Connectivity Establishment (ICE)
Protocol", draft-ietf-ice-trickle-21 (work in progress),
April 2018.

[I-D.ietf-avtext-lrr]

Lennox, J., Hong, D., Uberti, J., Holmer, S., and M.
Flodman, "The Layer Refresh Request (LRR) RTCP Feedback
Message", draft-ietf-avtext-lrr-07 (work in progress),
July 2017.

Appendix A. Design Considerations

One of the main issues regarding the BUNDLE grouping extensions has been whether, in SDP Offers and SDP Answers, the same port value can be inserted in "m=" lines associated with a BUNDLE group, as the purpose of the extension is to negotiate the usage of a single transport for media specified by the "m=" sections. Issues with both approaches, discussed in the Appendix have been raised. The outcome was to specify a mechanism which uses SDP Offers with both different and identical port values.

Below are the primary issues that have been considered when defining the "BUNDLE" grouping extension:

- o 1) Interoperability with existing UAs.
- o 2) Interoperability with intermediary Back to Back User Agent (B2BUA) and proxy entities.
- o 3) Time to gather, and the number of, ICE candidates.
- o 4) Different error scenarios, and when they occur.
- o 5) SDP Offer/Answer impacts, including usage of port number value zero.

A.1. UA Interoperability

Consider the following SDP Offer/Answer exchange, where Alice sends an SDP Offer to Bob:

SDP Offer

```
v=0
o=alice 2890844526 2890844526 IN IP4 atlanta.example.com
s=
c=IN IP4 atlanta.example.com
t=0 0

m=audio 10000 RTP/AVP 97
a=rtpmap:97 iLBC/8000
m=video 10002 RTP/AVP 97
a=rtpmap:97 H261/90000
```

SDP Answer

```
v=0
o=bob 2808844564 2808844564 IN IP4 biloxi.example.com
s=
c=IN IP4 biloxi.example.com
t=0 0

m=audio 20000 RTP/AVP 97
a=rtpmap:97 iLBC/8000
m=video 20002 RTP/AVP 97
a=rtpmap:97 H261/90000
```

RFC 4961 specifies a way of doing symmetric RTP but that is a later extension to RTP and Bob can not assume that Alice supports RFC 4961. This means that Alice may be sending RTP from a different port than 10000 or 10002 – some implementations simply send the RTP from an ephemeral port. When Bob's endpoint receives an RTP packet, the only way that Bob knows if the packet is to be passed to the video or audio codec is by looking at the port it was received on. This led some SDP implementations to use the fact that each "m=" section had a different port number to use that port number as an index to find the correct m line in the SDP. As a result, some implementations that do support symmetric RTP and ICE still use an SDP data structure where SDP with "m=" sections with the same port such as:

SDP Offer

```
v=0
o=alice 2890844526 2890844526 IN IP4 atlanta.example.com
s=
c=IN IP4 atlanta.example.com
t=0 0

m=audio 10000 RTP/AVP 97
a=rtpmap:97 iLBC/8000
m=video 10000 RTP/AVP 98
a=rtpmap:98 H261/90000
```

will result in the second "m=" section being considered an SDP error because it has the same port as the first line.

A.2. Usage of Port Number Value Zero

In an SDP Offer or SDP Answer, the media specified by an "m=" section can be disabled/rejected by setting the port number value to zero. This is different from e.g., using the SDP direction attributes, where RTCP traffic will continue even if the SDP "inactive" attribute is indicated for the associated "m=" section.

If each "m=" section associated with a BUNDLE group would contain different port values, and one of those port values would be used for a BUNDLE address:port associated with the BUNDLE group, problems would occur if an endpoint wants to disable/reject the "m=" section associated with that port, by setting the port value to zero. After that, no "m=" section would contain the port value which is used for the BUNDLE address:port. In addition, it is unclear what would happen to the ICE candidates associated with the "m=" section, as they are also used for the BUNDLE address:port.

A.3. B2BUA And Proxy Interoperability

Some back to back user agents may be configured in a mode where if the incoming call leg contains an SDP attribute the B2BUA does not understand, the B2BUA still generates that SDP attribute in the Offer for the outgoing call leg. Consider a B2BUA that did not understand the SDP "rtcp" attribute, defined in RFC 3605, yet acted this way. Further assume that the B2BUA was configured to tear down any call where it did not see any RTCP for 5 minutes. In this case, if the B2BUA received an Offer like:

SDP Offer

```
v=0
o=alice 2890844526 2890844526 IN IP4 atlanta.example.com
s=
c=IN IP4 atlanta.example.com
t=0 0

m=audio 49170 RTP/AVP 0
a=rtcp:53020
```

It would be looking for RTCP on port 49171 but would not see any because the RTCP would be on port 53020 and after five minutes, it would tear down the call. Similarly, a B2BUA that did not understand BUNDLE yet put BUNDLE in its offer may be looking for media on the wrong port and tear down the call. It is worth noting that a B2BUA that generated an Offer with capabilities it does not understand is not compliant with the specifications.

A.3.1. Traffic Policing

Sometimes intermediaries do not act as B2BUAs, in the sense that they don't modify SDP bodies, nor do they terminate SIP dialogs. Still, however, they may use SDP information (e.g., IP address and port) in order to control traffic gating functions, and to set traffic policing rules. There might be rules which will trigger a session to be terminated in case media is not sent or received on the ports retrieved from the SDP. This typically occurs once the session is already established and ongoing.

A.3.2. Bandwidth Allocation

Sometimes intermediaries do not act as B2BUAs, in the sense that they don't modify SDP bodies, nor do they terminate SIP dialogs. Still, however, they may use SDP information (e.g., codecs and media types) in order to control bandwidth allocation functions. The bandwidth allocation is done per "m=" section, which means that it might not be enough if media specified by all "m=" sections try to use that bandwidth. That may either simply lead to bad user experience, or to termination of the call.

A.4. Candidate Gathering

When using ICE, a candidate needs to be gathered for each port. This takes approximately 20 ms extra for each extra "m=" section due to the NAT pacing requirements. All of this gathering can be overlapped

with other things while e.g., a web-page is loading to minimize the impact. If the client only wants to generate TURN or STUN ICE candidates for one of the "m=" lines and then use trickle ICE [I-D.ietf-ice-trickle] to get the non host ICE candidates for the rest of the "m=" sections, it MAY do that and will not need any additional gathering time.

Some people have suggested a TURN extension to get a bunch of TURN allocations at once. This would only provide a single STUN result so in cases where the other end did not support BUNDLE, it may cause more use of the TURN server but would be quick in the cases where both sides supported BUNDLE and would fall back to a successful call in the other cases.

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

Cullen Jennings
Cisco
400 3rd Avenue SW, Suite 350
Calgary, AB T2P 4H2
Canada

Email: fluffy@iii.ca

Network WG
Internet-Draft
Expires: Sept 12, 2012
Intended Status: Standards Track (PS)

James Polk
Subha Dhesikan
Paul Jones
Cisco Systems
March 12, 2012

The Session Description Protocol (SDP) 'trafficclass' Attribute
draft-ietf-mmusic-traffic-class-for-sdp-01

Abstract

This document proposes a new Session Description Protocol (SDP) attribute to identify the traffic class a session is requesting in its offer/answer exchange.

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 12, 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	2
2. Traffic Class Framework and String Definitions	5
3. Traffic Class Attribute Definition	11
4. Offer/Answer Behavior	15
4.1 Offer Behavior	15
4.2 Answer Behavior	15
5. Security considerations	16
6. IANA considerations	17
7. Acknowledgments	19
8. References	19
8.1. Normative References	19
8.2. Informative References	20
Authors' Addresses	20
Appendix	20

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].

1. Introduction

The Session Description Protocol (SDP) [RFC4566] provides a means for an offerer to describe the specifics of a session to an answerer, and for the answerer to respond back with its session specifics to the offerer. These session specifics include offering the codec or codecs to choose from, the specific IP address and port number the offerer wants to receive the RTP stream(s) on/at, the particulars about the codecs the offerer wants considered or mandated, and so on.

There are many facets within SDP to determine the Real-time Transport Protocol (RTP) [RFC3550] details for the session establishment between one or more endpoints, but identifying how the underlying network should process each stream still remains under-specified.

The ability to identify a traffic flow by port number gives an indication to underlying network elements to treat traffic with dissimilar ports in a different way, the same or in groups the same - but different from other ports or groups of ports.

Within the context of realtime communications, the labeling of an RTP session based on media descriptor lines as just a voice and/or video session is insufficient, and provides no guidelines to the underlying network on how to treat the traffic. A more granular labeling helps on several fronts to

- inform application layer elements in the signaling path the intent of this session.
- inform the network on how to treat the traffic if the network is configured to differentiate session treatments based on the type of session the RTP is, including the ability to provide call admission control based on the type of traffic in the network.
- allow network monitoring/management of traffic types realtime and after-the-fact analysis.

Some network operators want the ability to guarantee certain traffic gets a minimum amount of network bandwidth per link or through a series of links that perhaps makes up a network such as a campus or WAN, or a backbone. For example, a call center voice application gets at least 20% of a link as a minimum bandwidth allocation.

Some network operators want the ability to allow certain users or devices access to greater bandwidth during non-busy hours, than during busy hours of the day. For example, all desktop video to operate at 1080p during non-peak times, but curtail a similar session between the same users or devices to 720p or 360p during peak hours. Another example would be to reduce the frames per second (fps) rate, say from 30fps to 15fps. This case is not as clear as accepting or denying similar sessions during different times of the day, but tuning the access to the bandwidth based on the type of session. In other words, tune down the bandwidth for desktop video during peak hours to allow a 3-screen telepresence session that would otherwise look like the same type of traffic (RTP, and more granular, video).

RFC 4594 established a guideline for classifying the various flows in the network and the Differentiated Services Codepoints (DSCP) that apply to many traffic types (table 3 of [RFC4594]), including RTP based voice and video traffic sessions. The RFC also defines the

per hop network behavior that is strongly encouraged for each of these application traffic types based on the traffic characteristics and tolerances to delay, loss and jitter within each traffic class.

Video was broken down into 4 categories in that RFC, and voice into another single category. We do not believe this satisfies the technical and business requirements to accomplish sufficiently unique labeling of RTP traffic.

A question arises about once we properly label the traffic, what does that get us? This is a fair question, but out of scope for this document because that answer lies within other RFCs and IDs in other WGs and/or Areas (specifically the Transport Area). That said, we can discuss some of the ideas here for completeness.

If the application becomes aware of traffic labeling,

- this can be coded into layer 3 mechanisms.
- this can be coded into layer 4 protocols and/or mechanisms.
- this can be coded into a combination of mechanisms and protocols.

The layer 3 mechanism for differentiating traffic is either the port number or the Differentiated Services Codepoint (DSCP) value [RFC2474]. Within the public Internet, if the application is not part of a managed service, the DSCP likely will be best effort (BE). Within the corporate LAN, this is usually completely configurable and a local IT department can take full advantage of this labeling to shape and manage their network as they see fit. Communications between enterprise networks will likely have to take advantage of MPLS.

Within a network core, where only MPLS is used, Diffserv typically does not apply. That said, Diffserv can be used to identify which traffic goes into which MPLS tunnels [RFC4124].

Labeling realtime traffic types using a layer 4 protocol would likely mean RSVP [RFC2205] or NSIS [RFC4080]. RSVP has an Application Identifier (app-ID) defined in [RFC2872] that provides a means for carrying a traffic class label along the media path. An advantage with this mechanism is for the label to inform each domain along the media path what type of traffic this traffic flow is, and allow each domain to adjust the appropriate DSCP (set by each domain for use within that domain). Meaning, if a DSCP is set by an endpoint or a router in the first domain and gets reset by a SP, the far end domain will be able to reset the DSCP to the intended traffic class. There is a proposed extension to RSVP which creates individual profiles for what goes into each app-ID field to describe these traffic classes [ID-RSVP-PROF], which will take advantage of what is described in this document.

There are several proprietary mechanisms to take advantage of this labeling, but none of those will be discussed here.

The idea of traffic - or service - identification is not new; it has been described in [RFC5897]. If that RFC is used as a guideline, identification that leads to stream differentiation can be quite useful. One of the points within RFC 5897 is that users cannot be allowed to assign any identification (fraud is but one reason given). In addition, RFC 5897 recommends that service identification should be done in signaling, rather than guessing or deep packet inspection. The network will have to currently guess or perform deep packet inspection to classify and offer the service as per RFC 4594 since such service identification information is currently not available in SDP and therefore to the network elements. Since SDP understands how each stream is created (i.e., the particulars of the RTP stream), this is the right place to have this service differentiated. Such service differentiation can then be communicated to and leveraged by the network.

[Editor's Note: the words "traffic" and "service" are similar enough that the above paragraph talks about RFC 5897's "service identification", but this document is only wanting to discuss and propose traffic indications in SDP.]

This document proposes a simple attribute line to identify the application a session is requesting in its offer/answer exchange. This document uses previously defined service class strings for consistency between IETF documents.

This document modifies the traffic classes originally created in RFC 4594 in Section 2, incrementing each class with application identifiers and optional adjective strings. Section 3 defines the new SDP attribute "trafficclass". Section 4 discusses the offerer and answerer behavior when generating or receiving this attribute.

2. Traffic Class Framework and String Definitions

The framework of the traffic class attribute will have at least two parts, allowing for several more to be included. The intention is to have a parent class (e.g., Conversational) that merely serves as the anchor point for an application component that when paired together, form the highest level traffic class. An adjective component provides further granularity for the application. There can be more than one adjective within a traffic class label to further refine the uniqueness of a traffic class being described.

The traffic class label will have the following structure,

```
parent.application(.adjective)(.adjective)
```

[Editor's Note: the above is not exactly the ABNF to be used. The order is right. The parent and application MUST appear (each only once) and zero or more adjectives can appear.]

Where

- 1) the 1st component is the human understandable category;
- 2) the 2nd component is the application;
- 3) an optional 3rd component or series of components are adjective(s) used to further refine the application component; and

The construction of the traffic class label for Telepresence video would follow the minimum form of:

Conversational.video.immersive

where there might be one or more adjective after '.immersive'.

There is no traffic class or DSCP value associated with just "Conversational". There is a traffic class associated with "Conversational.video", creating a differentiation between it and a "Conversational.video.immersive" traffic class, which would have DSCP associated with the latter traffic class, depending on local policy. Each parent component is defined below, as are several of application and adjective strings.

[Editor's Note: We're not yet sure how much of what's below will be proposed for IANA registration, but the 5 parent components will be, as well as at least some application components per parent component. Some adjective components will also likely be proposed for IANA registration.

The 5 parent components of the traffic class attribute are as follows:

- o Conversational
- o Multimedia Conferencing
- o Real-Time Interactive
- o Multimedia Streaming
- o Broadcast

The following application components of the traffic class attribute are as follows:

- o Audio
- o Video
- o Text
- o application-sharing
- o Presentation-data

- o Whiteboarding
- o Web (conference) chat/instant messaging
- o Gaming
- o Virtual-desktop (interactive)
- o Remote-desktop
- o Telemetry (e.g., NORAD missile control)
- o Multiplex (i.e., combined streams)
- o Webcast
- o IPTV
- o Live-events (though not the buffered ones)
- o surveillance

The following adjective components of the traffic class attribute are as follows:

- o Immersive
- o avconf
- o Realtime-Text
- o web

Each of the above 3 lists will be defined in the following subsections.

2.1 Conversational Parent Traffic Class

The Conversational traffic class is best suited for applications that require very low delay variation and generally intended to enable real-time, bi-directional person-to-person or multi-directional via an MTP communication, such as the following application components:

- o Audio (voice)**
- o Video**
- o Text (i.e., real-time text required by deaf users)

**The above applications will also be used within Multimedia Streaming and Broadcast

With adjective substrings to the above

Immersive (TP) - An interactive audio-visual communications experience between remote locations, where the users enjoy a strong sense of realism and presence between all participants by optimizing a variety of attributes such as audio and video quality, eye contact, body language, spatial audio, coordinated environments and natural image size.

Desktop-video - An interactive audio-visual communication

experience that is not immersive in nature, though can have a high resolution video component.

Realtime-Text (RTT) - a term for real-time transmission of text in a character-by-character fashion for use in conversational services, often as a text equivalent to voice-based conversational services. Conversational text is defined in the ITU-T Framework for multimedia services, Recommendation F.700 [RFC5194].

Web - for realtime aspects of web conferencing; mutually exclusive of both Immersive and Desktop video experiences

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
Conversational	High priority, typically small packets (large video frames produce large packets), generally sustained high packet rate, low inter-packet transmission interval, usually UDP framed in (S)RTP	Very Low	Very Low	Very Low

Figure 1. Conversational Traffic Characteristics

2.2 Multimedia-Conferencing Parent Traffic Class

Multimedia-Conferencing traffic class is best suited for applications that are generally intended for communication between human users, but are less demanding in terms of delay, packet loss, and jitter than what Conversational applications require. These applications require low to medium delay and may have the ability to change encoding rate (rate adaptive) or transmit data at varying rates, such as the following application component:

- o application-sharing (that webex does or protocols like T.128) - An application that shares the output of one or more running applications or the desktop on a host. This can utilize vector graphics, raster graphics or video.
- o Presentation-data - can be a series of still images or motion video.
- o Whiteboarding - an application enabling the exchange of graphical information including images, pointers and filled and unfilled parametric drawing elements (points, lines, polygons and ellipses).

- o (RTP-based) file transfer
- o Web (conference) chat/instant messaging

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
Multimedia Conferencing	Variable size packets, Variable transmit interval, rate adaptive, reacts to loss, usually TCP-based	Low	Low	Low
		-	-	-
		Medium	Medium	Medium

Figure 2. Multimedia Conferencing Traffic Characteristics

2.3 Realtime-Interactive Parent Traffic Class

Realtime-Interactive traffic class is intended for interactive variable rate inelastic applications that require low jitter and loss and very low delay, such as the following application components:

- o Gaming - interactive player video games with other users on other hosts (e.g., Doom)
- o Virtualized desktop (interactive) - similar to an X-windows station, has no local hard drive, or is operating an application with nlocal storage
- o Remote Desktop - controlling a remote node with local peripherals (i.e., monitor, keyboard and mouse)
- o Telemetry - a communication that allows remote measurement and reporting of information (e.g., post launch missile status or energy monitoring)

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
Realtime Interactive	Inelastic, mostly variable rate, rate increases with user activity	Low	Very Low	Low

Figure 3. Realtime Interactive Traffic Characteristics

2.4 Multimedia-Streaming Parent Traffic Class

Multimedia-Streaming traffic class is best suited for variable rate elastic streaming media applications where a human is waiting for output and where the application has the capability to react to packet loss by reducing its transmission rate, such as the following application components:

- o Audio
- o Video
- o Multiplex (i.e., combined a/v streams)

With adjective substrings to the above (which may or may not get IANA registered)

Webcast

The primary difference from the Multimedia-streaming parent class and the Broadcast parent class is about the length of time for buffering. Buffered streaming audio and/or video which are initiated by SDP, and not HTTP. Buffering here can be from many seconds to hours, and is typically at the destination end (as opposed to Broadcast buffering which is minimal at the destination). The buffering aspect is what differentiates this parent class from the Broadcast class (which has minimal or no buffering).

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
Multimedia Streaming	Variable size packets, elastic with variable rate	Low - Medium	Medium - High	High

Figure 4. Multimedia Streaming Traffic Characteristics

2.5 Broadcast Parent Traffic Class

Broadcast traffic class is best suited for inelastic streaming media Applications, which might have a 'wardrobe malfunction' delay at or near the source but not typically at the destination, that may be of constant or variable rate, requiring low jitter and very low packet loss, such as the following application components:

- o Audio
- o Video
- o Multiplex (i.e., combined a/v streams)

With adjective substrings to the above:

- o IPTV
- o Live events (non-buffered)
- o Video surveillance - one way video from a camera (e.g., observing a parking lot or building exit), typically enabled for long periods of time, usually stored at the destination.

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
Broadcast	Constant and variable rate, inelastic, generally non-bursty flows, generally sustained high packet rate, low inter-packet transmission interval, usually UDP framed in (S)RTP	Very Low	Low - Medium	Low - Medium

Figure 5. Broadcast Traffic Characteristics

3. SDP Attribute Definition

This document proposes the 'trafficclass' session and media-level SDP [RFC4566] attribute. The following is the Augmented Backus-Naur Form (ABNF) [RFC5234] syntax for this attribute, which is based on the SDP [RFC4566] grammar:

```

attribute                =/ traffic-classification

traffic-classification    = "trafficclass" ":" [SP] parent-class
                           "." app-type *( adj-param )

parent-class              = "Broadcast" /
                           "Realtime-Interactive" /
                           "Multimedia-Conferencing" /
                           "Multimedia-Streaming" /
                           "Conversational" /
                           extension-mech

extension-mech            = token

app-type                  = "audio" / "video" / "text" /
                           "application-sharing" /
                           "presentation-data" / "whiteboarding" /
                           "webchat/IM" / "gaming" /
                           "virtual-desktop" / "remote-desktop" /

```

```

        "telemetry" / "multiplex" / "webcast" /
        "IPTV" / "live-events" /
        "surveillance" / extension-mech

adj-param          = "." unqualified-adjective /
                    "." qualified-adjective

unqualified-adjective = "immersive" / "avconf" /
                        "Realtime-Text" / "web" /
                        generic-param ; from RFC3261

qualified-adjective  = qual-category ":" q-adjective

qual-category        = "aq" / extension-mech

q-adjective          = "admitted" / "non-admitted" / "none" /
                        generic-param ; from RFC3261

```

The attribute is named "trafficclass", for traffic classification, identifying which one of the five traffic classes applies to the media stream. There MUST NOT be more than one trafficclass attribute per media line. Confusion would result in where more than one exists per m= line.

The parent classes in this document are an augmented version of the application labels introduced by table 3 of RFC 4595 (which will be rewritten based on the updated labels and treatments expected for each traffic class defined in this document).

Application Labels Defined in RFC 4594	Parent Classes Defined in this document
Broadcast-video	Broadcast
Realtime-Interactive	Realtime-Interactive
Multimedia-Conferencing	Multimedia-Conferencing
Multimedia-Streaming	Multimedia-Streaming
Telephony	Conversational

Figure 6. Label Changes from RFC 4594

As is evident from the changes above, from left to right, two labels are different and each of the meanings are different in this document relative to how RFC 4594 defined them. These differences are articulated in Section 2 of this document.

A parent class is a human understandable categorization, and MUST NOT be the only part of the traffic class label present in the attribute. The parent class string MUST always be paired with an application type, with a "." as the component separator.

The application types (app-type) define the application of a particular traffic flow. The application types are listed both in the ABNF and defined in Section 2 of this document. Not every combination parent class is paired with application types, at least as defined in this document. Section 2.1 through 2.5 list many of the expected combinations.

For additional application type granularity, adjective components can be added (also listed in Section 2). One or more adjectives can be within the same traffic class attribute. It is also permitted to include one or more non-IANA registered adjective component, but these MUST be prefaced by the additional delimiter "_", creating a possibility such as

```
parent-class.application-type.adjective._non-standard-adjective
                                   ^^^^
                                   See the underscore
```

For example, this is valid:

```
m=video 50000 RTP/AVP 112
a=trafficclass Conversational.video.immersive._foo._bar
```

where both "foo" and "bar" are not IANA registered adjectives, but "immersive" is IANA registered. However, including non-registered adjectives without the "_" delimiter are not valid, such as the following:

```
m=video 50000 RTP/AVP 112
a=trafficclass Conversational.video.immersive.foo.bar
```

There is no limit to the number of adjectives allowed, without regard for whether they are registered or not. These non-registered adjectives can be vendor generated, or merely considered to be proprietary in nature.

It is important to note that the order of component types matter, but not the order of the adjective components. There might be local significance to the ordering though. In other words, the parent class component MUST be before the application component, which MUST be before the adjective component.

Some algorithm such as alphabetizing the list and matching the understood strings SHOULD be used.

Adjectives can be either unqualified or qualified. Qualified

adjectives have a designation it is qualified and a ":" separating the string component into two parts. We define this qualifying designation to have the form of a two or three letter qualifier, in which the last letter is always "q" (i.e., for "qualified").

We are proposing in this document to have a single qualified adjective indicating whether this trafficclass has had or will have capacity-admission applied to it. Here we define the admission qualifier ("aq") with three possible values for this adjective: admitted, non-admitted and none, that will have the form

aq:admitted|non-admitted|none

Like all adjectives, it is OPTIONAL to include this adjective in any trafficclass attribute, and has the following meanings:

- admitted - capacity admission mechanisms or protocols are to be or were used for the full amount of bandwidth in relation to this m= line.
- non-admitted - capacity admission mechanisms or protocols were attempted but failed in relation to this m= line. This does not mean the flow described by this m= line failed. It just failed to attain the capacity admission mechanism or protocol necessary for a predictable quality of service, and is likely to continue with only a class of service marking or best effort.
- none - no capacity admission mechanisms or protocols are or were attempted in relation to this m= line.

The default for any flow generated from an m= line not having a trafficclass adjective of 'aq:admitted' or 'aq:non-admitted' MUST be the equivalent of 'aq:none', whether or not it is present.

Any parent class, application, or adjective string component within this attribute that is not understood MUST be ignored, leaving all that is understood to be processed. Ignored string components SHOULD NOT be deleted, as a downstream entity could understand the component(s) and use it/them.

Not understanding the parent class string SHOULD mean that this attribute is ignored.

The following is an example of media level description with a 'trafficclass' attribute:

```
m=video 50000 RTP/AVP 112
a=trafficclass conversational.video.immersive.aq:admitted
```

The above indicates a telepresence session that has had capacity admission process applied to its media flow.

4. Offer/Answer Behavior

Through the inclusion of the 'trafficclass' attribute, an offer/answer exchange identifies the application type for use by endpoints within a session. Policy elements can use this attribute to determine the acceptability and/or treatment of that session through lower layers. One specific use-case is for setting of the DSCP specific for that application type (say a Broadcast instead of a Conversational video), decided on a per domain basis - instead of exclusively by the offering domain.

4.1 Offer Behavior

Offerers include the 'trafficclass' attribute with a single string comprised of two or more components (from the list in Section 2) to obtain configurable and predictable classification between the answerer and the offerer. The offerer can also include a private set of components, or a combination of IANA registered and private components within a single domain (e.g., enterprise networks).

Offerers of this 'trafficclass' attribute MUST NOT change the label in transit (e.g., wrt to B2BUAs). Session Border Controllers (SBC) at domain boundaries can change this attribute through local policy.

Offers containing a 'trafficclass' label not understood are ignored by default (i.e., as if there was no 'trafficclass' attribute in the offer).

4.2 Answer Behavior

Upon receiving an offer containing a 'trafficclass' attribute, if the offer is accepted, the answerer will use this attribute to classify the session or media (level) traffic accordingly towards the offerer. This answer does not need to match the traffic class in the offer, though this will likely be the case most of the time.

In order to understand the traffic class attribute, the answerer MUST check several components within the attribute, such as

1 - does the answerer understand the parent component?

If not, the attribute SHOULD be ignored.

If yes, it checks the application component.

2 - does the answerer understand the application component?

If not, the answerer needs to check if it has a local policy to proceed without an application component. The default for this

situation is as if the parent component was not understood, the attribute SHOULD be ignored.

If yes, it checks to see if there are any other components present in this attribute to start its classification.

- 3 - does the answerer understand the adjective component or components if any are present?

If not present, process and match the trafficclass label value as is.

If yes, determine if there is more than one. Search for each that is understood. Any adjectives not understood are to be ignored, as if they are not present.

The answerer will answer the offer with its own 'trafficclass' attribute, which will likely be the same value, although this is not mandatory (at this time). The Offerer will process the received answer just as the answerer processed the offer. In other words, the processing steps and rules are identical for each end.

The answerer should expect to receive RTP packets marked as indicated by its 'trafficclass' attribute in the answer itself.

An Answer MAY have a 'trafficclass' attribute when one was not in the offer. This will at least aid the local domain, and perhaps each domain the session transits, to categorize the application type of this RTP session.

Answerers that are middleboxes can use the 'trafficclass' attribute to classify the RTP traffic within this session however local policy determines. In other words, this attribute can help in deciding which DSCP an RTP stream is assigned within a domain, if the answerer were an inbound SBC to a domain.

5. Security considerations

RFC 5897 [RFC5897] discusses many of the pitfalls of service classification, which is similar enough to this idea of traffic classification to apply here as well. That document highly recommends the user not being able to set any classification. Barring a hack within an endpoint (i.e., to intentionally misclassifying (i.e., lying) about which classification an RTP stream is), this document's solution makes the classification part of the signaling between endpoints, which is recommended by RFC 5897.

6. IANA considerations

6.1 Registration of the SDP 'trafficclass' Attribute

This document requests IANA to register the following SDP att-field under the Session Description Protocol (SDP) Parameters registry:

Contact name: jmpolk@cisco.com

Attribute name: trafficclass

Long-form attribute name: Traffic Classification

Type of attribute: Session and Media levels

Subject to charset: No

Purpose of attribute: To indicate the Traffic Classification application for this session

Allowed attribute values: IANA Registered Tokens

Registration Procedures: Specification Required

Type	SDP Name	Reference
----	-----	-----
att-field (both session and media level)		
	trafficclass	[this document]

6.2 The Traffic Classification Application Type Registration

This document requests IANA to create a new registry for the traffic application classes similar to the following table within the Session Description Protocol (SDP) Parameters registry:

Registry Name: "trafficclass" SDP Application Type Attribute Values

Reference: [this document]

Registration Procedures: Specification Required

Parent Values	Reference
-----	-----
Broadcast	[this document]
Realtime-Interactive	[this document]
Multimedia-Conferencing	[this document]
Multimedia-Streaming	[this document]
Conversational	[this document]

6.3 The Traffic Classification Application Type Registration

This document requests IANA to create a new registry for the traffic application classes similar to the following table within the Session Description Protocol (SDP) Parameters registry:

Registry Name: "trafficclass" Attribute Application Type Values
Reference: [this document]
Registration Procedures: Specification Required

Application Values	Reference
-----	-----
Audio	[this document]
Video	[this document]
Text	[this document]
Application-sharing	[this document]
Presentation-data	[this document]
Whiteboarding	[this document]
Webchat/instant messaging	[this document]
Gaming	[this document]
Virtualized-desktop	[this document]
Remote-desktop	[this document]
Telemetry	[this document]
Multiplex	[this document]
Webcast	[this document]
IPTV	[this document]
Live-event	[this document]
surveillance	[this document]

6.4 The Traffic Classification Unqualified Adjective Registration

This document requests IANA to create a new registry for the traffic unqualified adjective values similar to the following table within the Session Description Protocol (SDP) Parameters registry:

Registry Name: "trafficclass" Attribute Unqualified Adjective Values
Reference: [this document]
Registration Procedures: Specification Required

Application Values	Reference
-----	-----
Immersive	[this document]
Desktop-video	[this document]
Realtime-Text	[this document]
web	[this document]

6.5 The Traffic Classification Attribute Qualified Adjective Values Registration

This document requests IANA to create a new registry Qualified Adjective Values similar to the following table within the Session

Description Protocol (SDP) Parameters registry:

Registry Name: "trafficclass" Attribute Qualified Adjective Values

Reference: [this document]

Registration Procedures: Specification Required

Qualification Category	Attribute Values	Reference
-----	-----	-----
AQ	Admitted	[this document]
AQ	Non-admitted	[this document]
AQ	None	[this document]

7. Acknowledgments

To Dave Oran, Toerless Eckert, Henry Chen, David Benham, David Benham, Mo Zanty, Michael Ramalho, Glen Lavers, Charles Ganzhorn, and Greg Edwards for their comments and suggestions.

8. References

8.1. Normative References

- [RFC2119] S. Bradner, "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997
- [RFC2205] R. Braden, Ed., L. Zhang, S. Berson, S. Herzog, S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997
- [RFC2474] K. Nichols, S. Blake, F. Baker, D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers ", RFC 2474, December 1998
- [RFC2872] Y. Bernet, R. Pabbati, "Application and Sub Application Identity Policy Element for Use with RSVP", RFC 2872, June 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.
- [RFC4080] R. Hancock, G. Karagiannis, J. Loughney, S. Van den Bosch, "Next Steps in Signaling (NSIS): Framework", RFC 4080, June 2005
- [RFC4124] F. Le Faucheur, Ed., " Protocol Extensions for Support of Diffserv-aware MPLS Traffic Engineering ", RFC 4124, June 2005
- [RFC4566] M. Handley, V. Jacobson, C. Perkins, "SDP: Session

Description Protocol", RFC 4566, July 2006

[RFC5234] Crocker, D., Ed., and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.

[RFC5865] F. Baker, J. Polk, M. Dolly, "A Differentiated Services Code Point (DSCP) for Capacity-Admitted Traffic", RFC 5865, May 2010

[RFC5897] J. Rosenberg, "Identification of Communications Services in the Session Initiation Protocol (SIP)", RFC 5897, June 2010

8.2. Informative References

[RFC4594] J. Babiarz, K. Chan, F Baker, "Configuration Guidelines for Diffserv Service Classes", RFC 4594, August 2006

[ID-RSVP-PROF] J. Polk, S. Dhesikan, "Resource Reservation Protocol (RSVP) Application-ID Profiles for Voice and Video Streams", work in progress, Mar 2011

Author's Addresses

James Polk
3913 Treemont Circle
Colleyville, Texas, USA
+1.818.271.3552

mailto: jmpolk@cisco.com

Subha Dhesikan
170 W Tasman St
San Jose, CA, USA
+1.408-902-3351

mailto: sdhesika@cisco.com

Paul E. Jones

mailto: paulej@packetizer.com

Appendix - Changes from Previous Versions

A.1 From -00 to -01

These are the following changes made between the WG -00 version and the -01 version:

- removed the non-SDP applications Netflix and VOD
- switched the adjective 'desktop' to 'avconf'
- Labeled each of the figures.
- clarified the differences between Multimedia-Streaming and Broadcast parent categories.
- defined Video surveillance
- added the concept of a 'qualified' adjective, and modified the ABNF.
- deleted the idea of a 'cac-class' as a separate component, and made the equivalent a qualified adjective.
- modified the answerer behavior because of the removal of the 'cac-class' component.
- created an IANA registry for qualified adjectives
- general clean-up of the doc.

Did **not** do the following in this version:

- add the ability to have more than one trafficclass attribute based on the codec chosen, as feedback indicated this was a bad idea.
- no swap of the Multimedia-Conferencing parent category with the offered Collaboration parent category, as doing this did not solve any perceived problems.
- add more to the 'how does this get processed' portion of Section 3. That will come in the next revision.

Network WG
Internet-Draft
Expires: January 3, 2015
Intended Status: Standards Track (PS)

James Polk
Subha Dhesikan
Paul Jones
Cisco Systems
July 3, 2014

The Session Description Protocol (SDP) 'trafficclass' Attribute
draft-ietf-mmusic-traffic-class-for-sdp-05

Abstract

This document proposes a new Session Description Protocol (SDP) attribute to identify the traffic class a session is requesting in its offer/answer exchange.

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 4, 2015.

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	2
2.	Traffic Class Framework and Component Definitions	5
3.	Traffic Class Attribute Definition	6
3.1	Categories within the SDP Traffic Class Label	8
3.2	Applications within the SDP Traffic Class Label	9
3.3	Adjectives within the SDP Traffic Class Label	9
3.3.1	Qualified Adjectives	9
4.	Matching Categories with Applications and Adjectives	11
4.1	Conversational Category Traffic Class	11
4.2	Multimedia-Conferencing Category Traffic Class	12
4.3	Realtime-Interactive Category Traffic Class	14
4.4	Multimedia-Streaming Category Traffic Class	15
4.5	Broadcast Category Traffic Class	17
4.6	Intermittent Category Traffic Class	18
5.	Offer/Answer Behavior	19
5.1	Offer Behavior	20
5.2	Answer Behavior	20
6.	Security considerations	21
7.	IANA considerations	21
8.	Acknowledgments	25
9.	References	25
9.1.	Normative References	25
9.2.	Informative References	26
	Authors' Addresses	26
	Appendix	27

1. Introduction

The Session Description Protocol (SDP) [RFC4566] provides a means for an offerer to describe the specifics of a session to an answerer, and for the answerer to respond back with its session specifics to the offerer. These session specifics include offering the codec or codecs to choose from, the specific IP address and port number the offerer wants to receive the RTP stream(s) on/at, the particulars about the codecs the offerer wants considered or mandated, and so on.

There are many facets within SDP to determine the Real-time Transport Protocol (RTP) [RFC3550] details for the session establishment between one or more endpoints, but identifying how the underlying network should process each stream still remains under-specified.

The ability to identify a traffic flow by port number gives an

indication to underlying network elements to treat traffic with dissimilar ports in a different way, the same or in groups the same - but different from other ports or groups of ports.

Within the context of realtime communications, the labeling of an RTP session based on media descriptor lines as just a voice and/or video session is insufficient, and provides no guidelines to the underlying network on how to treat the traffic. A more granular labeling helps on several fronts to

- inform application layer elements in the signaling path the intent of this session.
- inform the network on how to treat the traffic if the network is configured to differentiate session treatments based on the type of session the RTP is, including the ability to provide call admission control based on the type of traffic in the network.
- allow network monitoring/management of traffic types realtime and after-the-fact analysis.

Some network operators want the ability to guarantee certain traffic gets a minimum amount of network bandwidth per link or through a series of links that make up a network such as a campus or WAN, or a backbone. For example, a call center voice application might get at least 20% of the available link bandwidth.

Some network operators want the ability to allow certain users or devices access to greater bandwidth during non-busy hours than during busy hours of the day. For example, all desktop video might operate at 1080p during non-peak times, but a similar session might be curtailed between the same users or devices to 720p or 360p during peak hours. Another example would be to reduce the frames per second (fps) rate, say from 30fps to 15fps. This case is not as clear as accepting or denying similar sessions during different times of the day, but tuning the access to the bandwidth based on the type of session. In other words, tune down the bandwidth for desktop video during peak hours to allow a 3-screen Telepresence session that would otherwise look like the same type of traffic (RTP, and more granular, video).

RFC 4594 established a guideline for classifying the various flows in the network and the Differentiated Services Codepoint (DSCP) values that apply to many traffic types (table 3 of [RFC4594]), including RTP based voice and video traffic sessions. The RFC also defined the per hop network behavior that is strongly encouraged for each of these application traffic types based on the traffic characteristics and tolerances to delay, loss and jitter within each traffic class.

Video was broken down into four categories in that RFC, and voice in another single category. We do not believe this satisfies the

technical and business requirements to accomplish sufficiently unique labeling of RTP traffic.

If the application becomes aware of traffic labeling,

- this can be coded into layer 3 mechanisms.
- this can be coded into layer 4 protocols and/or mechanisms.
- this can be coded into a combination of mechanisms and protocols.

A lower layer mechanism for differentiating traffic is either the port number or the Differentiated Services Codepoint (DSCP) value [RFC2474]. Within the public Internet, if the application is not part of a managed service, the DSCP value likely will be best effort (BE), or reset to BE, at ingress to a provider's network. Within the corporate LAN, this is usually completely configurable and a local IT department can take full advantage of this labeling to shape and manage their network as they see fit.

Within a network core, DiffServ typically does not apply. That said, DiffServ can be used to identify which traffic goes into which MPLS tunnel [RFC4124].

Labeling realtime traffic types using a layer 4 protocol would likely involve RSVP [RFC2205] or NSIS [RFC4080]. RSVP has an Application Identifier (app-ID) defined in [RFC2872] that provides a means for carrying a traffic class label along the media path. An advantage of this mechanism is that the label can inform each domain along the media path what type of traffic this traffic flow is, and allow each domain to adjust the appropriate DSCP value (set by each domain for use within that domain). Meaning, if a DSCP value is set by an endpoint or a router in the first domain and gets reset by a service provider, the far-end domain will be able to reset the DSCP value appropriate for the intended traffic class. There is a proposed extension to RSVP which creates individual profiles for what goes into each app-ID field to describe these traffic classes [ID-RSVP-PROF], which will take advantage of what is described in this document.

There are several proprietary mechanisms that can take advantage of this labeling, but none of those will be discussed here.

The idea of traffic - or service - identification is not new; it has been described in [RFC5897]. If that RFC is used as a guideline, identification that leads to stream differentiation can be quite useful. One of the points within RFC 5897 is that users cannot be allowed to assign any identification (fraud is one reason given). In addition, RFC 5897 recommends that service identification should be done in signaling, rather than guessing or deep packet inspection. Currently, any network would have to guess or perform deep packet inspection to classify traffic and offer the service as per

RFC 4594 as such service identification information is currently not available in SDP and therefore to the network elements. Since SDP understands how each stream is created (i.e., the particulars of the RTP stream), this is the right place to have this service differentiated. Such service differentiation can then be communicated to and leveraged by the network.

[Editor's Note: the words "traffic" and "service" are similar enough that the above paragraph talks about RFC 5897's "service identification", but this document only discusses and proposes traffic indications in SDP.]

This document proposes a simple attribute line to identify the application a session is requesting in its offer/answer exchange. This document uses previously defined service class strings for consistency between IETF documents.

This document utilizes the traffic classes originally created in RFC 4594 in Section 2, enhancing each class with application identifiers and optional adjective strings. Section 3 defines the new SDP attribute "trafficclass". Section 4 discusses the offerer and answerer behavior when generating or receiving this attribute.

1.1 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].

2. Traffic Class Framework and Component Definitions

The framework of the traffic class attribute will have at least two parts, called components, allowing for several more to be included further distinguishing a particular session's traffic classification from another session's traffic classification. The amount of indicated differentiation between sessions is not a goal, and should only have additional components for differentiation if there is a need to uniquely identify traffic in different sessions.

The intention is to have a category component (e.g., conversational) that identifies the traffic pattern for a session. Is the traffic within a session one-way or two-way? Can the traffic be buffered before reaching the destination or not? What is this session's tolerance to packet loss and can there be retransmissions?

The application component (e.g., video) identifies the basic type of traffic within a category. Is it media or data packets? If media, which type of media? If data packets, which application of data packets are in this session?

The optional adjective component(s) (e.g., immersive) help to

further refine the traffic within a session by providing more description. For instance, if a session is two-way voice, what additional information can be given about this particular session to refine its description? Is it part of a conference or telepresence session? Is it just standalone voice call? Has a capacity admission protocol or mechanism been applied to this session?

The 'traffic class label' (TCL) will have the following structure,

```
category.application(.adjective)(.adjective)...
```

[Editor's Note: the above is not the exact ABNF to be used.
The order is right. The category and application
MUST appear first (each only once) and zero or more
adjectives can appear following the application
component.]

Where

- 1) the 1st component is the category, and is mandatory;
- 2) the 2nd component is the application, and is mandatory;
- 3) an optional 3rd component or series of components are adjective(s) used to further refine the application component;

The construction of the traffic class label for Telepresence video would follow the minimum form of:

```
conversational.video.immersive
```

where there might be one or more adjective after '.immersive'.

There is no traffic class or DSCP value associated with just "conversational". There is a traffic class associated with "conversational.video", creating a differentiation between it and a "conversational.video.immersive" traffic class, which would have DSCP associated with the latter traffic class, depending on local policy. Each category component is defined below, as are several of application and adjective strings. This is shown in [ID-RSVP-PROF] for the RSVP mapping of distinguishable traffic types.

Mapping a specific Traffic Class Label to a DSCP value might be accomplished in any of the following ways:

- o statically within the offerer and/or answerer; or
- o taken from a local mapping table/file, which might be downloaded once, periodically or as changes in the network are observed; or
- o from feedback from the network.

3. Traffic Class Attribute Definition

This document defines the 'trafficclass' media-level SDP attribute. The following is the Augmented Backus-Naur Form (ABNF) [RFC5234] syntax for this attribute, which is based on the SDP [RFC4566] grammar:

```

attribute                =/ traffic-class-label

traffic-class-label      = "trafficclass" ":" [SP] category
                           "." application *( "." adjective )

category                 = "broadcast" /
                           "realtime-interactive" /
                           "multimedia-conferencing" /
                           "multimedia-streaming" /
                           "conversational" /
                           "intermittent" / tcl-token

application              = tcl-token

adjective                = classified-adjective /
                           unclassified-adjective

classified-adjective     = tcl-token ":" tcl-token

unclassified-adjective   = tcl-token

tcl-token                = ALPHA *( [ "-" ] ALPHA / DIGIT )

```

A TCL "component" is any of the following:

- category,
- application, or
- adjective (which is the only optional component, and can have zero or more of these type of components)

The attribute is named "trafficclass", for traffic classification, identifying which one of the six categories applies to the media stream associated with this m-line. There MUST NOT be more than one category component per SDP media line.

The categories in this document are an augmented version of the application labels introduced by table 3 of RFC 4594 (which will be rewritten based on the updated labels and treatments expected for each traffic class defined in this document).

Application Labels Defined in RFC 4594	Category Classes Defined in this document
broadcast-video	broadcast

realtime-interactive	realtime-interactive	
multimedia-conferencing	multimedia-conferencing	
multimedia-streaming	multimedia-streaming	
telephony	conversational	

Figure 1. Label Differences from RFC 4594

As is evident from the changes above, from left to right, two labels are different and each of the meanings are different in this document relative to how RFC 4594 defined them. These differences are articulated in Section 4 of this document.

Applications and adjectives are defined using the syntax of "tcl-token" defined above.

RFC 4566 defined SDP as case sensitive. Everything is here as well.

An algorithm such as alphabetizing the list of components and matching the understood strings SHOULD be used for determining the traffic within a session. Strings not understood by an entity MUST be ignored during processing, but MUST NOT be deleted.

Any category, application, or adjective string component within this attribute that is not understood MUST be ignored, leaving all that is understood to be processed. Ignored components SHOULD NOT be deleted, as a downstream entity could understand the component(s) and use it/them during processing.

The following is an example of media level description with a 'trafficclass' attribute:

```
m=video 50000 RTP/AVP 112
a=trafficclass:conversational.video.immersive.aq:admitted
```

The above indicates the video part of a Telepresence session that has had capacity admission process applied to its media flow.

3.1 Categories within the SDP Traffic Class Label

The category component within the traffic class attribute describes the type of communication that will occur within that session. It answers these questions, is the traffic

- one-way or two-or-more-way interactive?
- buffered or (virtually) non-buffered?

- media or non-media (data)?

The six category components of the traffic class attribute defined within this specification are as follows:

- o conversational
- o multimedia-conferencing
- o realtime-interactive
- o multimedia-streaming
- o broadcast
- o intermittent

Sections 4.1 through 4.6 define each of the above.

The category component **MUST NOT** be the only component present in a traffic class attribute. The category component **MUST BE** paired with an 'application' component to give enough meaning to the traffic class labeling goal.

Not understanding the category component **SHOULD** mean that this attribute is ignored, because of the information about the expected behavior of this communication flow is identified by or within that component.

3.2 Applications within the SDP Traffic Class Label

The application component identifies the application of a particular traffic flow, for example, audio or video. The application types are listed and defined in Section 4 of this document. Not every category is paired with every application listed, at least as defined in this document. One or more applications are inappropriate in one or more categories.

Section 4.1 through 4.6 list many of the expected combinations.

3.3 Adjectives within the SDP Traffic Class Label

For additional application type granularity, adjective components can be added. One or more adjectives can be within the same traffic class attribute to provide more differentiation.

It is important to note that while the order of component types matter, the order of the adjective components do not. In other words, the category class component **MUST** be before the application component, which **MUST** be before any and all adjective component(s).

There is no limit to the number of adjectives allowed.

Adjective components come in two versions, unqualified and

qualified. One has a prefix (qualified), the other (unqualified) does not. A defined qualified adjective MUST NOT appear without its qualifier name, even in future extensions to this specification. Some implementations will likely perform a search within this attribute for the presence of qualifiers, which might be as simple as searching for the ":" COLON character. Implementations will be confused with inconsistent coding, therefore strict adherence is necessary.

3.3.1 Qualified Adjectives

Adjectives can be either unqualified or qualified. Qualified adjectives have a delimiter ":" character between the "qualifier name" and the "qualifier value". As one example, we introduce in this specification the "admission qualifier" and it has a qualifier name of "aq". We also define several possible qualifier values for the admission qualifier, namely "admitted", "non-admitted", "partial", and "none". When present in a TCL component, the qualified adjectives look like these admission qualifier adjectives:

```
aq:admitted
aq:non-admitted
aq:partial
aq:none
```

Defining some adjectives as qualified adjectives allows entities processing the traffic class label to potentially recognize a particular qualifier name and act on it, even if it does not understand the qualifier value. In the future, a new admission qualifier value might be defined, e.g. "foo", and entities could at least recognize the admission qualifier adjective, even if it did not understand the qualifier value "foo".

Like all adjectives, it is OPTIONAL to include the admission qualifier adjective in any trafficclass attribute.

The admission qualifier and its qualifier values are defined as:

- aq - 'admission qualifier' - this is the qualifier name for the admission qualifier adjectives, wherein the following qualifier values indicate the admission status for the traffic flow described by this m-line.
- admitted - capacity admission mechanisms or protocols are to be or were used for the full amount of bandwidth in relation to this m= line.
- non-admitted - capacity admission mechanisms or protocols were attempted but failed in relation to this m= line. This does not mean the flow described by this m= line failed. It just failed to attain the capacity admission

mechanism or protocol necessary for a predictable quality of service, and is likely to continue with only a class of service marking or best effort.

- partial - capacity admission mechanisms or protocols are to be or were used for the part of the amount of bandwidth in relation to this m= line. All traffic above a certain amount will have no capacity admission mechanisms applied. In other words, there is more traffic sent than was agreed to. The burden is on the sender and receiver to deal with any sent and lost information.
- none - no capacity admission mechanisms or protocols are or were attempted in relation to this m= line.

The default for any flow generated from an m-line not having a trafficclass adjective of 'aq:admitted' or 'aq:non-admitted' MUST be the equivalent of 'aq:none', whether or not it is present.

4. Matching Categories with Applications and Adjectives

This section describes each component within this document, as well as provides the combinations of categories and applications and adjectives. Given that not every combination makes sense, we express the limits here - which will be IANA registered. The majority of these TCLs in this document are found in [ID-RSVP-PROF], where RSVP is appropriate. Look at that other document for example usage of a specified TCL here.

4.1 Conversational Category Traffic Class

The "conversational" traffic class is best suited for applications that require very low delay variation and generally intended to enable realtime, bi-directional person-to-person or multi-directional via an MCU communication. Conversational flows are inelastic, and with few exceptions, use a UDP transport.

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
conversational	High priority, typically consistent sized packets (small audio samples produce small packets and large video samples produce large packets), generally sustained at a high packet rate, low inter-packet transmission interval	Very Low	Very Low	Very Low

Figure 2. Conversational Traffic Characteristics

The following application components are appropriate for use with the Conversational category:

- o audio (voice)
- o video
- o multiplex (i.e., combined a/v streams) an application wherein media of different forms (e.g., audio and video) is multiplexed within the same media flow.

With adjective substrings to the above

immersive (TP) - An interactive audio-visual communications experience between remote locations, where the users enjoy a strong sense of realism and presence between all participants by optimizing a variety of attributes such as audio and video quality, eye contact, body language, spatial audio, coordinated environments and natural image size.

avconf - An interactive audio-visual communication experience that is not immersive in nature, though can have a high resolution video component.

Category	Application	Adjective
conversational	audio	immersive
		avconf
		aq:admitted
		aq:non-admitted
		aq:partial
		aq:none
	video	immersive
		avconf
		aq:admitted
		aq:non-admitted
		aq:partial
		aq:none
	multiplex	immersive
		avconf
		aq:admitted
		aq:non-admitted
		aq:partial
		aq:none

+-----+-----+-----+-----+

Figure 3. Conversational Applications and Adjective Combinations

4.2 Multimedia-Conferencing Category Traffic Class

The "multimedia-conferencing" traffic class is best suited for applications that are generally intended for communication between human users, but are less demanding in terms of delay, packet loss, and jitter than what conversational applications require. These applications require low to medium delay and may have the ability to change encoding rate (rate adaptive) or transmit data at varying rates.

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
multimedia- conferencing	Variable size packets, Variable transmit interval, rate adaptive, reacts to loss, often one-way or unidirectional	Low	Low	Low
		- Medium	- Medium	- Medium

Figure 4. Multimedia Conferencing Traffic Characteristics

Multimedia-conferencing flows are not media flows which are conversational in nature. Multimedia-conferencing flows are those data flows that are typically transmitted in parallel to currently active conversational media flows. For example, a two-way conference session in which the users share a presentation. The presentation part of that conference call uses the Multimedia-conferencing category, whereas the audio and any video uses the conversational category indication.

The following application components are appropriate for use with the Multimedia-Conferencing category:

- o application-sharing (that webex does or protocols like T.128) - An application that shares the output of one or more running applications or the desktop on a host. This can utilize vector graphics, raster graphics or video.
- o presentation-data - can be a series of still images; could be at a rapid or busty rate, just not a continuous 24 fps or greater.
- o presentation-video - motion video that is transmitted and rendered as part of a presentation.
- o presentation-audio - the audio that is transmitted and rendered as

part of a presentation.

- o whiteboarding - an application enabling the exchange of graphical information including images, pointers and filled and unfilled parametric drawing elements (points, lines, polygons and ellipses).
- o (RTP-based) file-transfer as defined in RFC 5547
- o instant messaging

Category	Application	Adjective
multimedia-conferencing	application-sharing	aq:admitted aq:non-admitted aq:partial aq:none
	whiteboarding	aq:admitted aq:non-admitted aq:partial aq:none
	presentation-data	aq:admitted aq:non-admitted aq:partial aq:none
	presentation-video	aq:admitted aq:non-admitted aq:partial aq:none
	presentation-audio	aq:admitted aq:non-admitted aq:partial aq:none
	instant-messaging	aq:admitted aq:non-admitted aq:partial aq:none
	file-transfer	aq:admitted aq:non-admitted aq:partial aq:none

Figure 5. Multimedia Conferencing Applications and Adjective Combinations

4.3 Realtime-Interactive Category Traffic Class

The "Realtime-Interactive" traffic class is intended for interactive variable rate inelastic applications that require low jitter and loss and very low delay. Many of the applications that use the Realtime-Interactive category use TCP or SCTP, even gaming, because lost packets is information that is still required - therefore it is retransmitted.

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
realtime- interactive	Inelastic, mostly variable rate, rate increases with user activity	Low	Very Low	Low

Figure 6. Realtime Interactive Traffic Characteristics

The following application components are appropriate for use with the Realtime-Interactive category:

- o gaming - interactive player video games with other users on other hosts (e.g., Doom)
- o remote-desktop - controlling a remote node with local peripherals (i.e., monitor, keyboard and mouse)
- o telemetry - a communication that allows remote measurement and reporting of information (e.g., post launch missile status or energy monitoring)

With adjective substrings to the above

- o virtual - To be used with the remote-desktop application component specifically when the traffic is a virtual desktop similar to an X-windows station, has no local hard drive, or is operating a computer application with no local storage.

Category	Application	Adjective
realtime-interactive	gaming	aq:admitted aq:non-admitted aq:partial aq:none
	remote-desktop	virtual aq:admitted aq:non-admitted

		aq:partial
		aq:none
	telemetry	aq:admitted
		aq:non-admitted
		aq:partial
		aq:none

Figure 7. Realtime-Interactive Applications and Adjective Combinations

4.4 Multimedia-Streaming Category Traffic Class

The "multimedia-streaming" traffic class is best suited for variable rate elastic streaming media applications where a human is waiting for output and where the application has the capability to react to packet loss by reducing its transmission rate.

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
multimedia-streaming	Variable size packets, elastic with variable rate	Low - Medium	Medium - High	High

Figure 8. Multimedia Streaming Traffic Characteristics

The following application components are appropriate for use with the Multimedia-Streaming category:

- o audio (see Section 4.1)
- o video (see Section 4.1)
- o webcast - is a media file distributed over the Internet or enterprise network using streaming media technology.
- o multiplex (see Section 4.1)

The primary difference between the multimedia-streaming category and the broadcast category is the length of time for buffering. Buffered streaming of audio and/or video which is often initiated by HTTP, and not SDP. Buffering here can be from many seconds to hours, and is typically at the destination end (as opposed to Broadcast buffering which is minimal at the destination). The buffering aspect is what differentiates this category class from the broadcast category (which has minimal or no buffering).

Category	Application	Adjective
multimedia-streaming	audio	aq:admitted aq:non-admitted aq:partial aq:none
	video	aq:admitted aq:non-admitted aq:partial aq:none
	webcast	aq:admitted aq:non-admitted aq:partial aq:none
	multiplex	aq:admitted aq:non-admitted aq:partial aq:none

Figure 9. Multimedia Streaming Applications and Adjective Combinations

4.5 Broadcast Category Traffic Class

The "broadcast" traffic class is best suited for inelastic streaming media Applications, which might have a 'wardrobe malfunction' delay at or near the source but not typically at the destination, that may be of constant or variable rate, requiring low jitter and very low packet loss.

See Section 4.4 for the difference between Multimedia-Streaming and Broadcast; it all has to do with buffering.

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
broadcast	Constant and variable rate, inelastic, generally non-bursty flows, generally sustained high packet rate, low inter-packet transmission interval	Very Low	Low - Medium	Low - Medium

Figure 10. Broadcast Traffic Characteristics

The following application components are appropriate for use with the Broadcast category:

- o audio (see Section 4.1)
- o video (see Section 4.1)
- o multiplex (see Section 4.1)

With adjective substrings to the above:

- o live (non-buffered) - refers to various types of media broadcast without a significant delay, typically measured in milliseconds to a few seconds only.
- o surveillance - one way audio from a microphone or video from a camera (e.g., observing a parking lot or building exit), typically enabled for long periods of time, usually stored at the destination.

Category	Application	Adjective
broadcast	audio	surveillance
		live
		aq:admitted
		aq:non-admitted
		aq:partial
		aq:none
	video	surveillance
		live
		aq:admitted
		aq:non-admitted
		aq:partial
		aq:none
	multiplex	surveillance
		live
		aq:admitted
		aq:non-admitted
		aq:partial
		aq:none

Figure 11. Broadcast Applications and Adjective Combinations

4.6 Intermittent Category Traffic Class

The "intermittent" traffic class is best suited for inconstant rate applications such as those from a sensor device, where tolerance to loss, delay and jitter is often medium to high in nature. This category is not to be used for bulk file transfers, rather it can be sometimes bursty for brief periods of time, but then not produce traffic for short or long (i.e., hours or days) durations. Nor is this category to be used for any kind of regular paced rate of transmission, no matter how long the interval.

Traffic Class Name	Traffic Characteristics	Tolerance to		
		Loss	Delay	Jitter
intermittent	Inconstant rate, infrequent but sometimes bursty flows, generally non-regular, variable inter-packet transmission interval	Medium - High	Medium - High	High

Figure 12. Intermittent Traffic Characteristics

The following application components are appropriate for use with the Broadcast category:

- o text (i.e., text required by deaf users) a term for seemingly real-time transmission of text in a character-by-character fashion, often as a text equivalent to voice-based conversational services, without the timing constraints of conversational text is defined in the ITU-T Framework for multimedia services, Recommendation F.700 [RFC5194].
- o sensor - a flow containing information obtained from a sensor, such as a temperature or motion sensor.

With adjective substrings to the above:

- o there are no defined adjectives for the 'sensor' application at this time. There are many examples one could think would be viable adjectives, such as light, motion, temperature, magnetic fields, gravity, humidity, moisture, vibration, pressure, electrical fields, and other physical aspects of the external environment measured by the sensor.

Category	Application	Adjective
intermittent	sensor	(undefined at this time)

	text	aq:admitted aq:non-admitted aq:partial aq:none
--	------	---

Figure 13. Intermittent Applications and Adjective Combinations

5. Offer/Answer Behavior

Through the inclusion of the 'trafficclass' attribute, an offer/answer exchange identifies the application type(s) for use by the endpoints within the media streams of a session. Signaling elements can use this attribute to determine the acceptability and/or treatment of that session through lower layers, communicating a desired treatment for a particular flow to endpoints using SDP, interacting with network elements using some unspecified mechanism, or having endpoints communicate with network elements using some unspecified mechanism.

In order to understand the traffic class attribute, the SDP entity MUST check several components within the Traffic Class Label. By understand, we mean that the value of each component of the TCL is recognized, i.e., both the category and application components MUST be a recognized set for a TCL to be understood. Adjectives that are not recognized are simply ignored and MAY be discarded, however many there are. Adjectives which are not understood SHOULD NOT be discarded, as each/any adjective might be understood by some or all other downstream nodes in the signaling path.

The following pertains to both the receiver of an offer and receiver of an answer when either or both contain a Traffic Class Label attribute.

- 1 - can the receiver of the SDP containing a trafficclass attribute successfully process the category component?

If not, the attribute SHOULD be ignored.

If yes, it checks the application component.

- 2 - can the receiver of the SDP containing a trafficclass attribute successfully process the application component?

If not, the answerer needs to check if it has a local policy to proceed without an application component. The default for this situation is as if the category component was not understood, meaning the attribute SHOULD be ignored.

If yes, it checks to see if there are any adjective components

present in this attribute to start its classification.

- 3 - can the receiver of the SDP containing a trafficclass attribute successfully process the adjective component or components if any are present?

If not present, process and match the trafficclass label value as is.

If yes, determine if there is more than one. Search for each that is understood. Any adjectives not understood are to be ignored, as if they are not present. Match all remaining understood components according to local policy and process attribute.

5.1 Offer Behavior

Offerers include the 'trafficclass' attribute within a single string per m= line comprised of at least a category and application component (see Section 4) to establish the non-generic classification of the media stream between the answerer and the offerer. The offerer can also include one or more adjective components, which might be a combination of registered and private adjectives to further refine the identification of what this particular media stream is.

Session Border Controllers (SBC) at domain boundaries can change this attribute through local policy.

5.2 Answer Behavior

Upon receiving an offer containing a 'trafficclass' attribute, if the offer is accepted - including the ability to process the 3 bulleted rules in Section 5.0, the answerer will use this attribute to classify the media level traffic accordingly towards the offerer.

The answerer will answer the offer with its own 'trafficclass' attribute, which will likely be the same value, although this is not mandatory (at this time). The Offerer will process the received answer just as the answerer processed the offer. In other words, the processing steps and rules are identical for each end (see Section 5).

An Answer MAY have a 'trafficclass' attribute when one was not in the offer. This will at least aid the local domain, and perhaps each domain the session transits, to categorize and in some cases group the media-types of this session.

6. Security considerations

The security considerations from RFC 4566 are also applicable, particularly since intermediary devices might be able to look at an m= line and determine, not only is it audio, but that it is presentation-audio (i.e., 'multimedia-conferencing.presentation-audio') versus conversational audio.

RFC 5897 [RFC5897] discusses many of the pitfalls of service classification, which is similar enough to this idea of traffic classification to apply here as well. That document highly recommends the user not being able to set any classification. Barring a hack within an endpoint (i.e., to intentionally misclassifying (i.e., lying) about which classification an RTP stream is), this document's solution makes the classification part of the signaling between endpoints, which is recommended by RFC 5897.

7. IANA considerations

7.1 Registration of the SDP 'trafficclass' Attribute

This document requests IANA to register the following SDP att-field under the Session Description Protocol (SDP) Parameters registry:

Contact name: jmpolk@cisco.com

Attribute name: trafficclass

Long-form attribute name: Traffic Classification

Type of attribute: Media levels

Subject to charset: No

Purpose of attribute: To indicate the Traffic Classification application for this session

Allowed attribute values: IANA Registered Tokens

Registration Procedures: (there are multiple RFC5226 registration procedures; see below within each sub-section)

Designated Experts: James Polk (jmpolk@cisco.com)
Paul Jones (paulej@packetizer.com)

Type	SDP Name	Reference
----	-----	-----
att-field	(media level)	

trafficclass

[this document]

7.2 The Traffic Classification Category Registration

This document requests IANA to create a new registry for the traffic category classes similar to the following table within the Session Description Protocol (SDP) Parameters registry:

Registry Name: "trafficclass" SDP Category Attribute Values
Reference: [this document]
Registration Procedures: Standards Action Required [RFC5226]

Category Values	Reference
-----	-----
broadcast	[this document]
realtime-interactive	[this document]
multimedia-conferencing	[this document]
multimedia-streaming	[this document]
conversational	[this document]
intermittent	[this document]

7.3 The Traffic Classification Application Type Registration

This document requests IANA to create a new registry for the traffic application classes similar to the following table within the Session Description Protocol (SDP) Parameters registry:

Registry Name: "trafficclass" SDP Application Attribute Values
Reference: [this document]
Registration Procedures: Specification Required [RFC5226]

Application Values	Reference
-----	-----
audio	[this document]
video	[this document]
text	[this document]
application-sharing	[this document]
presentation-data	[this document]
presentation-video	[this document]
presentation-audio	[this document]
whiteboarding	[this document]
instant-messaging	[this document]
gaming	[this document]
remote-desktop	[this document]
telemetry	[this document]
multiplex	[this document]
webcast	[this document]
sensor	[this document]

7.4 The Traffic Classification Adjective Registration

This document requests IANA to create a new registry for the traffic adjective values similar to the following table within the Session Description Protocol (SDP) Parameters registry:

Registry Name: "trafficclass" SDP Adjective Attribute Values
Reference: [this document]
Registration Procedures: Specification Required [RFC5226]

Adjective Values	Reference
-----	-----
immersive	[this document]
avconf	[this document]
realtime	[this document]
web	[this document]
virtual	[this document]
live	[this document]
surveillance	[this document]
aq:admitted	[this document]
aq:non-admitted	[this document]
aq:partial	[this document]
aq:none	[this document]

7.5 The Traffic Classification Component Mapping

7.5.1 Broadcast Applications and Adjective Combinations

This document requests IANA to create a new registry for the Broadcast category mapping similar to Figure 11 in Section 4.5 of this document within the Session Description Protocol (SDP) Parameters registry:

Registry Name: Broadcast Applications and Adjective Combinations
Table
Reference: [this document]
Registration Procedures: Specification Required [RFC5226]

7.5.2 Realtime Interactive Applications and Adjective Combinations

This document requests IANA to create a new registry for the Realtime Interactive category mapping similar to Figure 7 in Section 4.3 of this document within the Session Description Protocol (SDP) Parameters registry:

Registry Name: Realtime Interactive Applications and Adjective
Combinations Table
Reference: [this document]
Registration Procedures: Specification Required [RFC5226]

7.5.3 Multimedia Conferencing Applications and Adjective Combinations

This document requests IANA to create a new registry for the Multimedia Conferencing category mapping similar to Figure 5 in Section 4.2 of this document within the Session Description Protocol (SDP) Parameters registry:

Registry Name: Multimedia Conferencing Applications and Adjective Combinations Table

Reference: [this document]

Registration Procedures: Specification Required [RFC5226]

7.5.4 Multimedia-Streaming Applications and Adjective Combinations

This document requests IANA to create a new registry for the Multimedia-Streaming category mapping similar to Figure 9 in Section 4.4 of this document within the Session Description Protocol (SDP) Parameters registry:

Registry Name: Multimedia-Streaming Applications and Adjective Combinations Table

Reference: [this document]

Registration Procedures: Specification Required [RFC5226]

7.5.5 Conversational Applications and Adjective Combinations

This document requests IANA to create a new registry for the conversational category mapping similar to Figure 3 in Section 4.1 of this document within the Session Description Protocol (SDP) Parameters registry:

Registry Name: Conversational Applications and Adjective Combinations Table

Reference: [this document]

Registration Procedures: Specification Required [RFC5226]

7.5.6 Intermittent Applications and Adjective Combinations

This document requests IANA to create a new registry for the intermittent category mapping similar to Table 13 in Section 4.6 of this document within the Session Description Protocol (SDP) Parameters registry:

Registry Name: Intermittent Applications and Adjective Combinations Table

Reference: [this document]

Registration Procedures: Specification Required [RFC5226]

7.6 Designated Expert Reviewers

The following will be the designated expert reviewers of new 'trafficclass' registry requests:

- James Polk <jmpolk@cisco.com>
- Paul E. Jones <paulej@packetizer.com>

There SHALL remain two designated Expert reviewers at all times. The MMUSIC WG chairs should be consulted for replacements, if one or both are needed.

8. Acknowledgments

To Dave Oran, Toerless Eckert, Henry Chen, David Benham, David Benham, Mo Zanty, Michael Ramalho, Glen Lavers, Charles Ganzhorn, Paul Kyzivat, Greg Edwards, Charles Eckel, Dan Wing, Cullen Jennings and Peter Saint-Andre for their comments and suggestions.

9. References

9.1. Normative References

- [RFC2119] S. Bradner, "Key words for use in RFCs to Indicate Requirement Levels", RFC 2119, March 1997
- [RFC2205] R. Braden, Ed., L. Zhang, S. Berson, S. Herzog, S. Jamin, "Resource ReSerVation Protocol (RSVP) -- Version 1 Functional Specification", RFC 2205, September 1997
- [RFC2474] K. Nichols, S. Blake, F. Baker, D. Black, "Definition of the Differentiated Services Field (DS Field) in the IPv4 and IPv6 Headers ", RFC 2474, December 1998
- [RFC2872] Y. Bernet, R. Pabbati, "Application and Sub Application Identity Policy Element for Use with RSVP", RFC 2872, June 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.
- [RFC4080] R. Hancock, G. Karagiannis, J. Loughney, S. Van den Bosch, "Next Steps in Signaling (NSIS): Framework", RFC 4080, June 2005
- [RFC4124] F. Le Faucheur, Ed., " Protocol Extensions for Support of Diffserv-aware MPLS Traffic Engineering ", RFC 4124,

June 2005

- [RFC4566] M. Handley, V. Jacobson, C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006
- [RFC5226] T. Narten, H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", RFC 5226, May 2008
- [RFC5234] Crocker, D., Ed., and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", STD 68, RFC 5234, January 2008.
- [RFC5547] M. Garcia-Martin, M. Isomaki, G. Camarillo, S. Loreto, P. , Kyzivat, "A Session Description Protocol (SDP) Offer/Answer Mechanism to Enable File Transfer ", RFC 5547, May 2009
- [RFC5865] F. Baker, J. Polk, M. Dolly, "A Differentiated Services Code Point (DSCP) for Capacity-Admitted Traffic", RFC 5865, May 2010
- [RFC5897] J. Rosenberg, "Identification of Communications Services in the Session Initiation Protocol (SIP)", RFC 5897, June 2010

9.2. Informative References

- [RFC4594] J. Babiarez, K. Chan, F Baker, "Configuration Guidelines for Diffserv Service Classes", RFC 4594, August 2006
- [ID-RSVP-PROF] J. Polk, S. Dhesikan, "Resource Reservation Protocol (RSVP) Application-ID Profiles for Voice and Video Streams", work in progress, Feb 2013

Author's Addresses

James Polk
3913 Treemont Circle
Colleyville, Texas, USA
+1.818.271.3552

mailto: jmpolk@cisco.com

Subha Dhesikan
170 W Tasman St
San Jose, CA, USA
+1.408-902-3351

mailto: sdhesika@cisco.com

Paul E. Jones
7025 Kit Creek Rd.
Research Triangle Park, NC, USA
+1 919 476 2048

mailto: paulej@packetizer.com

Appendix - Changes from Previous Versions

A.1 From -04 to -05

These are the following changes made between the WG -03 version and the -04 version:

- general clean-up of text.
- added presentation-video and presentation-audio to the multimedia-conferencing section.
- brought forward the text describing how a SDP entity handles receiving a session description containing the trafficclass attribute to Section 5, from 5.2.
- added RFC 5547 as a normative reference.
- expended the security considerations section.

A.2 From -03 to -04

These are the following changes made between the WG -03 version and the -04 version:

- minimal text changes.
- introduced the "intermittent" category based on IETF86 feedback in the WG.

A.3 From -02 to -03

These are the following changes made between the WG -02 version and the -03 version:

- Rearranged a fair amount of text
- Separated and defined the components into separate subsections.
- built 5 different tables, one per category, that lists within each category - what applications are appropriate as well as what adjectives are appropriate for each application within that

category.

- added the 'partial' admission qualifier for those flows that have only part of their respective flow admitted (i.e., CAC'd).

A.4 From -01 to -02

These are the following changes made between the WG -01 version and the -02 version:

- converged the use of terms 'parent' and 'category' to just 'category' for consistency.
- changed ABNF to reflect extensibility by not having applications and adjectives named in the ABNF, rather have them merely IANA registered.
- merged the qualified and unqualified adjective sections into a single section on adjectives, but allowing some to have a preceding qualifier.
- text clean-up

A.5 From -00 to -01

These are the following changes made between the WG -00 version and the -01 version:

- removed the non-SDP applications Netflix and VOD
- switched the adjective 'desktop' to 'avconf'
- Labeled each of the figures.
- clarified the differences between Multimedia-Streaming and Broadcast category categories.
- defined Video surveillance
- added the concept of a 'qualified' adjective, and modified the ABNF.
- deleted the idea of a 'cac-class' as a separate component, and made the equivalent a qualified adjective.
- modified the answerer behavior because of the removal of the 'cac-class' component.
- created an IANA registry for qualified adjectives
- general clean-up of the doc.

Network Working Group
Internet-Draft
Intended status: Informational
Expires: December 7, 2012

E. Iovov
Jitsi
H. Kaplan
Acme Packet
D. Wing
Cisco
June 5, 2012

Latching: Hosted NAT Traversal (HNT) for Media in Real-Time
Communication
draft-iovov-mmusic-latching-00

Abstract

This document describes behavior of signalling intermediaries in RTC deployments, sometimes referred to as Session Border Controllers (SBCs), when performing Hosted NAT Traversal (HNT). HNT is a set of mechanisms, such as media relaying and latching, that such intermediaries use to enable other RTC devices behind NATs to communicate with each other. This document is non-normative, and is only written to explain HNT in order to provide a reference to the IETF community, as well as an informative description to manufacturers, and users.

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 December 7, 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	4
3. Background	4
4. Impact on Signaling	5
5. Media Behavior, Latching	6
6. Security Considerations	10
7. References	12
7.1. Normative References	12
7.2. Informative References	12
Authors' Addresses	13

1. Introduction

Network Address Translators (NATs) are widely used in the Internet by consumers and organizations. Although specific NAT behaviors vary, this document uses the term "NAT" for devices that map any IPv4 or IPv6 address and transport port number to another IPv4 or IPv6 address and transport port number. This includes consumer NATs, Firewall-NATs, IPv4-IPv6 NATs, Carrier-Grade NATs, etc.

Protocols like SIP [RFC3261], and others that try to use a more direct path for media than with signalling, are difficult to use across NATs. They use IP addresses and transport port numbers encoded in bodies such as SDP [RFC4566] as well as, in the case of SIP, various header fields. Such addresses and ports are unusable unless all peers in a session are located behind the same NAT.

Mechanisms such as STUN [RFC5389], TURN [RFC5766], and ICE [RFC5245], did not exist when protocols like SIP began being deployed. Session Border Controllers (SBCs) that were already being used by SIP domains for other SIP and media-related purposes began to use proprietary mechanisms to enable SIP devices behind NATs to communicate across the NATs.

The term often used for this behavior is Hosted NAT Traversal (HNT), although some manufacturers sometimes use other names such as "Far-end NAT Traversal" or "NAT assist" instead. The systems which perform HNT are frequently SBCs as described in [RFC5853], although other systems such as media gateways and "media proxies" sometimes perform the same role. For the purposes of this document, all such systems are referred to as SBCs, and the NAT traversal behavior is called HNT.

As of this document's creation time, a vast majority of SIP domains use HNT to enable SIP devices to communicate across NATs, despite the publication of ICE. There are many reasons for this, but those reasons are not relevant to this document's purpose and will not be discussed. It is however worth pointing out that the current deployment levels of HNT and NATs themselves make an exclusive adoption of ICE highly unlikely in the foreseeable future.

The purpose of this document is to describe the mechanisms often used for HNT at the SDP and media layer, in order to aid understanding the implications and limitations imposed by it. Although the mechanisms used in HNT are not novel to experts, publication in an IETF document is useful as a means of providing common terminology and a reference for related documents.

In no way does this document try to make a case for HNT or present it

as a solution that is somehow superior to alternatives such as ICE.

It is also worth mentioning that there are purely signaling-layer components of HNT as well. One such component is briefly described for SIP in [RFC5853], but that is not the focus of this document. The SIP signaling-layer component of HNT is typically more implementation-specific and deployment-specific than the SDP and media components. For the purposes of this document it is hence assumed that signaling intermediaries handle traffic in way that allows protocols such as SIP to function correctly across the NATs.

The rest of this document is going to focus primarily on use of HNT for SIP. However, the mechanisms described here are relatively generic and are often used with other protocols, such as XMPP [RFC6120], MGCP, H.248/MEGACO, and H.323.

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. Background

The general problems with NAT traversal for protocols such as SIP are:

1. The addresses and port numbers encoded in SDP bodies (or their equivalents) by NATed User Agents (UAs) are not usable across the Internet, because they represent the private addressing information of the UA rather than the addresses/ports that will be mapped to/from by the NAT.
2. The policies inherent in NATs, and explicit in Firewalls, are such that packets from outside the NAT cannot reach the UA until the UA sends packet out first.
3. Some NATs apply endpoint dependent filtering on incoming packets, as described in [RFC4787] and thus a UA may only be able to receive packets from the same remote peer IP:port as it sends packets out to.

In order to overcome these issues, signaling intermediaries such as SIP SBCs on the public side of the NATs perform HNT for both signaling and media. An example deployment model of HNT and SBCs is shown in Figure 1.

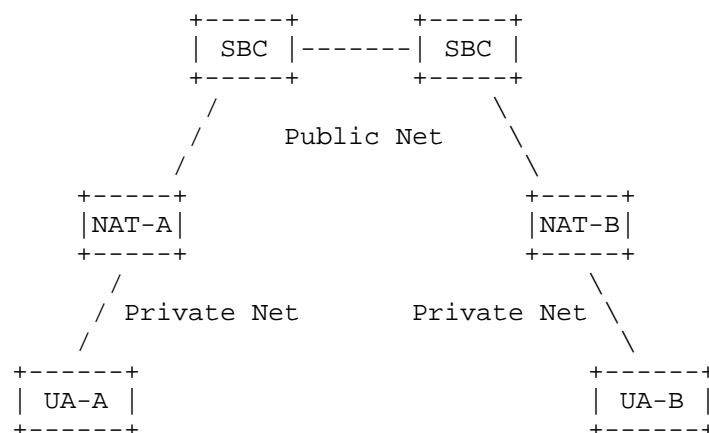


Figure 1: Logical Deployment Paths

4. Impact on Signaling

Along with codec and other media-layer information, session establishment signaling also conveys, potentially private and non-globally routable addressing information. Signaling intermediaries would hence modify such information so that peer UAs are given the (public) addressing information of a media relay controlled by the intermediary.

Quite often, the IP address of the newly introduced media relay may be the same as that of the signaling intermediary (e.g. the SIP SBC) or it may be a completely different one. In almost all cases however, the new address would belong to the same IP address family as the one used for signaling, since it is known to work for that UA.

The port numbers introduced in the signaling by the intermediary are typically allocated dynamically. Allocation strategies are entirely implementation dependent and they often vary from one product to the next.

The offer/answer media negotiation model [RFC3264] is such that once an offer is sent, the generator of the offer needs to be prepared to receive media on the advertised address/ports. In practice such media may or may not be received, depending on the implementations participating in a given session, local policies, and call scenario. For example if a SIP SDP Offer originally came from a UA behind a NAT, the SIP SBC cannot send media to it until an SDP Answer is given to the UA and latching (Section 5) occurs. Another example is when a

SIP SBC sends an SDP Offer in a SIP INVITE to a residential customer's UA and receives back SDP in a 18x response, the SBC may decide not to send media to that customer UA until a SIP 200 response for policy reasons, to prevent toll-fraud.

5. Media Behavior, Latching

An UA behind a NAT streams media from a private address:port set that once packets cross the NAT, will be mapped to a public set. The UA however is not typically aware of the public mapping and would often advertise in the private address:port couple in signaling. This way, when the signalling intermediary performing HNT receives the private addressing information from the UA it will not know what address/ports to send media to. Therefore media relays used in HNT would often use a mechanism called "latching".

Historically, "latching" only referred to the process by which SBCs "latch" onto UDP packets from a given UA for security purposes, and "symmetric-latching" is when the latched address:ports are used to send media back to the UA. Today most people talk about them both as "latching", and thus this document does as well.

The latching mechanism works as follows:

1. After receiving an offer from a NATed UA, a signaling intermediary located on the public Internet would allocate a set of IP address:ports on a media relay. The set would then be advertised to the remote party so that it would use it for all media it wished to send toward the UA.
2. Next, after receiving an answer to its offer, the signaling server would allocate a second address:port set on the media relay. It would advertise this second set to the UA and use it for all media traffic to and from the UA.
3. The media relay receives the media packets on the allocated ports, and uses their source address and port as a destination for all media bound in the opposite direction. In other words, it "latches" or locks on these source address:port set.
4. This way all media streamed by the UA would be received on the second address:port set. The source addresses and ports of the traffic would belong to the public interface of the NAT in front of the UA and anything that the relay sends there would find its way to it.
5. Similarly the source of the stream originating at the remote party would be latched upon and used for media going in that direction.
6. Latching is usually done only once per peer and not allowed to change or cause a re-latching until a new offer and answer get exchanged.

Figure 2 describes how latching occurs for SIP where HNT is provided by an SBC connected to two networks: 38.2.2/24 facing towards the UAC network and 198.51.100/24 facing towards the UAS network.

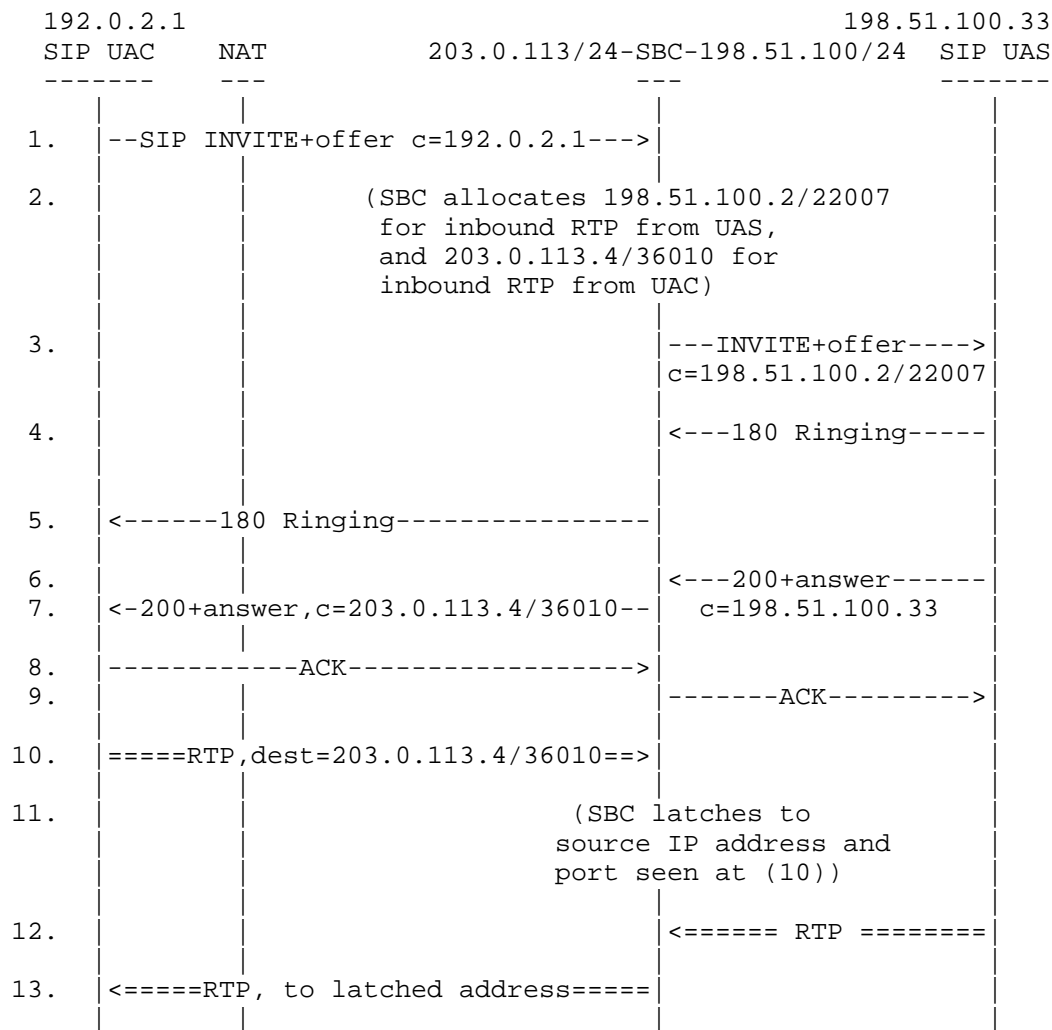


Figure 2: Latching by a SIP SBC across two interfaces

While XMPP implementations often rely on ICE to handle NAT traversal, there are some that also support a non-ICE transport called Raw UDP [XEP-0177]. Figure 3 describes how latching occurs for one such XMPP implementation where HNT is provided by an XMPP server on the public

internet.

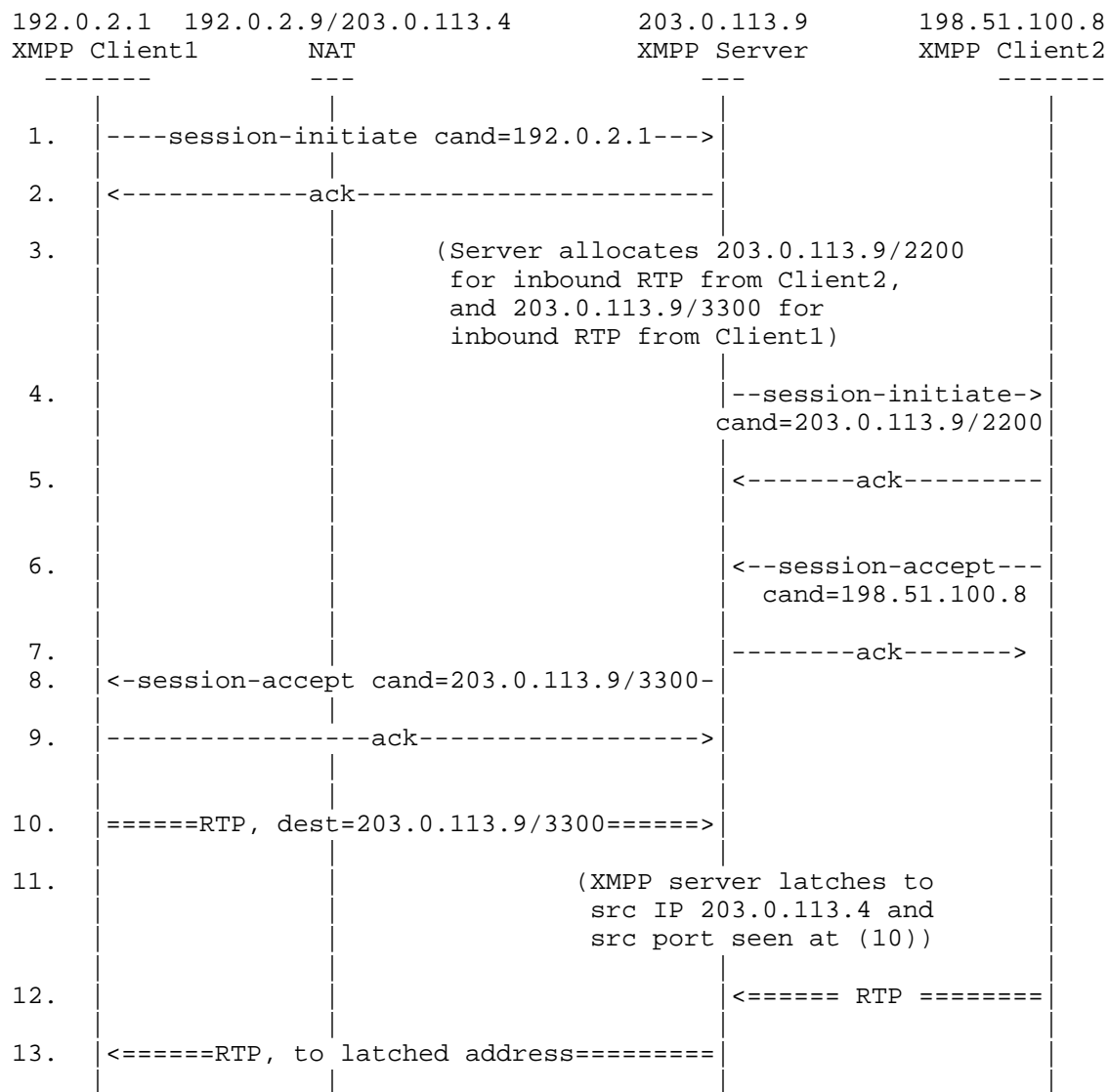


Figure 3: Latching by a SIP SBC across two interfaces

The above is a general description, and some details vary between implementations or configuration settings. For example, some intermediaries perform additional logic before latching on received

packet source information to prevent malicious attacks or latching erroneously to previous media senders - often called "rogue-rtp" in the industry.

It is worth pointing out that latching is not an exclusively "server affair" and some clients may also use it in cases where they are configured with a public IP address and they are contacted by a NATed client with no other NAT traversal means.

In order for latching to function correctly, the UA behind the NAT needs to support symmetric RTP. That is, it needs to use the same ports for sending data as the ones it listens on for inbound packets. Today this is the case for with, for example, almost all SIP and XMPP clients. Also UAs need to make sure they can begin sending media packets independently and without waiting for packets to arrive first. In theory, it is possible that some UAs would not send packets out first; for example if a SIP session begins in 'inactive' or 'recvonly' SDP mode from the UA behind the NAT. In practice, however, SIP sessions from regular UAs (the kind that one could find behind a NAT) virtually never begin in an inactive or recvonly mode, for obvious reasons. The media direction would also be problematic if the SBC side indicated 'inactive' or 'sendonly' modes when it sent SDP to the UA. However SBCs providing HNT would always be configured to avoid this.

Given that, in order for latching to work properly, media relays need to begin receiving media before they start sending, it is possible for deadlocks to occur. This can happen when the UAC and the UAS in a session are connected to different signalling intermediaries that both provide HNT. In this case the media relays controlled by the signalling servers could end up each waiting upon the other to initiate the streaming. To prevent this relays would often attempt to start streaming toward the address:port sets provided in the offer/answer even before receiving any inbound traffic. If the entity they are streaming to is another HNT performing server it would have provided its relay's public address and ports and the early stream would find its target.

Although many SBCs only support UDP-based media latching, and in particular RTP/RTCP, many SBCs support TCP-based media latching as well. TCP-based latching is more complicated, and involves forcing the UA behind the NAT to be the TCP client and sending the initial SYN-flagged TCP packet to the SBC (i.e., be the 'active' mode side of a TCP-based media session). If both UAs of a TCP-based media session are behind NATs, then SBCs typically force both UAs to be the TCP clients, and the SBC splices the TCP connections together. TCP splicing is a well-known technique, and described in [tcp-splicing].

HNT and latching in particular are generally found to be working reliably but they do have obvious caveats. The first one usually raised by IETF members is that UAs are not aware of it occurring. This makes it impossible for the mechanism to be used with protocols such as ICE that try various traversal techniques in an effort to choose the one the best suits a particular situation. Overwriting address information in in offers and answers may actually completely prevent UAs from using ICE because of the ice-mismatch rules described in [RFC5245]

The second issue raised by IETF members is that it causes media to go through a relay instead of directly over the IP-routed path between the two participating UAs. While this adds obvious drawbacks such as reduced scalability and often increased latency, it is also considered a benefit by SBC administrators: if a customer pays for "phone" service, for example, the media is what is truly being paid for, and the administrators usually like to be able to detect that media is flowing correctly, evaluate its quality, know if and why it failed, etc. Also in some cases routing media through operator controlled relays may route media over paths explicitly optimized for media and hence offer better performance than regular Internet routing.

6. Security Considerations

The security implications for HNT are complicated. The mechanism itself needs to be concerned with latching to incorrect and possibly malicious sources. A malicious source could, for example, attempt a resource exhaustion attack by flooding all possible media-latching UDP ports on the SBC in order to prevent calls from succeeding. SBCs have various mechanisms to prevent this from happening., or alert an administrator, but a sufficiently sophisticated attacker may be able to bypass them for some time. The most common example is typically referred to as "restricted-latching", whereby the SBC will not latch to any packets from a source public IP other than the one the SIP UA uses for SIP signaling. In some cases the limitation may be loosened to allow media from a range of IPs belonging to the same network. This way the SBC simply ignores and does not latch onto packets coming from the attacker. If the attacker knows the public source IP of the real SIP UA making a call, then they could still flood all of the SBC's public IPs and ports with packets spoofing that real SIP UA's public source IP. However, this would only disturb media that IP (or range of IPs) rather than all calls that the SBC is servicing.

A malicious source could send media packets to an SBC media-latching UDP port in the hopes of being latched-to for the purpose of receiving media for a given SIP session. SBCs have various

mechanisms to prevent this as well. Restricted latching for example would also help in this case since the attacker can't make the SBC send media packets back to themselves since the SBC will not latch onto the attackers packets. There could still be an issue if the attacker happens to be either (1) in the IP routing path and thus can spoof the same IP as the real UA and get the media coming back, in which case the attacker hardly needs to attack at all to begin with, or (2) the attacker is behind the same NAT as the real SIP UA, in which case the attacker's packets will be latched-to by the SBC and the SBC will send media back to the attacker. In this latter case, which is a corner-case to begin with, in practice the real SIP UA will end the call anyway, because the human won't hear anything and will hang up. EXCEPT, if it's not a human but rather an answering machine, it may not hang up (though most answering machines do hang up when they don't get media). The attacker could also redirect all media to the real SIP UA after receiving it, in which case the attack would likely remain undetected and succeed. Naturally, SRTP [RFC3711] would prevent such an attack from being useful, and should be used independently of HNT.

For SIP clients, HNT is usually transparent in the sense that the SIP UA does not know it occurs. In certain cases it may be detectable, such as when ICE is supported by the SIP UA and the SBC modifies the default connection address and media port numbers in SDP, thereby disabling ICE due to the mismatch condition. Even in that case, however, the SIP UA only knows a middlebox is relaying media, but not necessarily that it is performing latching/HNT. [TODO: need to explain further]

In order to perform HNT, the SBC has to modify SDP to and from the SIP UA behind a NAT, and thus the SIP UA cannot use S/MIME [RFC5751], and it cannot sign a sending request or verify a received request using [RFC4474] unless the SBC re-signs the request. However it is sometimes argued that, neither S/MIME nor [RFC4474] are widely deployed and that this may not be a real concern.

From a privacy perspective, media relaying is sometimes seen as a way of protecting one's IP address and not revealing it to the remote party. That kind of IP address masking is often perceived as important. However, this is no longer an exclusive advantage of HNT since it can also be accomplished by client-controlled relaying mechanisms such as TURN [RFC5766], if the client explicitly wishes to do so.

7. References

7.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

7.2. Informative References

- [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.
- [RFC3489] Rosenberg, J., Weinberger, J., Huitema, C., and R. Mahy, "STUN - Simple Traversal of User Datagram Protocol (UDP) Through Network Address Translators (NATs)", RFC 3489, March 2003.
- [RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.
- [RFC4474] Peterson, J. and C. Jennings, "Enhancements for Authenticated Identity Management in the Session Initiation Protocol (SIP)", RFC 4474, August 2006.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC4787] Audet, F. and C. Jennings, "Network Address Translation (NAT) Behavioral Requirements for Unicast UDP", BCP 127, RFC 4787, January 2007.
- [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.
- [RFC5751] Ramsdell, B. and S. Turner, "Secure/Multipurpose Internet Mail Extensions (S/MIME) Version 3.2 Message Specification", RFC 5751, January 2010.

- [RFC5766] Mahy, R., Matthews, P., and J. Rosenberg, "Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN)", RFC 5766, April 2010.
- [RFC5853] Hautakorpi, J., Camarillo, G., Penfield, R., Hawrylyshen, A., and M. Bhatia, "Requirements from Session Initiation Protocol (SIP) Session Border Control (SBC) Deployments", RFC 5853, April 2010.
- [RFC6120] Saint-Andre, P., "Extensible Messaging and Presence Protocol (XMPP): Core", RFC 6120, March 2011.
- [RFC6189] Zimmermann, P., Johnston, A., and J. Callas, "ZRTP: Media Path Key Agreement for Unicast Secure RTP", RFC 6189, April 2011.
- [XEP-0177] Beda, J., Saint-Andre, P., Hildebrand, J., and S. Egan, "XEP-0177: Jingle Raw UDP Transport Method", XEP XEP-0177, December 2009.

Authors' Addresses

Emil Ivov
Jitsi
Strasbourg 67000
France

Email: emcho@jitsi.org

Hadriel Kaplan
Acme Packet
100 Crosby Drive
Bedford, MA 01730
USA

Email: hkaplan@acmepacket.com

Dan Wing
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134
USA

Email: dwing@cisco.com

Network Working Group
Internet-Draft
Updates: 5245, 6544 (if approved)
Intended status: Standards Track
Expires: January 17, 2013

A. Keranen
J. Arkko
Ericsson
July 16, 2012

Update on Candidate Address Selection for
Interactive Connectivity Establishment (ICE)
draft-keranen-mmusic-ice-address-selection-01

Abstract

This document revisits the rules on how candidate addresses are selected and combined when the Interactive Connectivity Establishment (ICE) NAT traversal method is used. This document updates RFCs 5245 and 6544.

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	3
2. Terminology	3
3. Changes to Candidate Address Selection	4
4. Negotiating Address Selection Scheme	4
5. Security Considerations	5
6. IANA Considerations	6
7. References	6
7.1. Normative References	6
7.2. Informative References	7
Appendix A. Acknowledgments	7
Authors' Addresses	7

1. Introduction

When Interactive Connectivity Establishment (ICE) [RFC5245] [RFC6544] is used for NAT traversal, both endpoints gather multiple IP addresses and ports, also called candidate addresses, and test for connectivity between them. One of the principles of ICE is to gather all possible candidate addresses and pair them with all the addresses of the peer in order to test all combinations and get high probability for successful NAT traversal.

A prioritization formula is used by ICE so that most preferred address pairs are tested first, and if a sufficiently good pair is discovered, the tests can be stopped. Addresses obtained from local network interfaces, called host candidates, are recommended as high-priority ones to be tested first since if they work, they provide usually the best path between the two hosts. With IPv4 this approach works well since interfaces usually have just a single unicast IP address. However, with IPv6 addressing architecture [RFC4291] interfaces commonly have multiple addresses: global, link-local, Unique Local (ULA) [RFC4193], etc.

The ICE specification recommends to use the rules defined in [RFC3484] as part of the prioritization formula for IPv6 candidates, but does not give much further advice on how to handle different kind of IPv6 addresses. However, if all different kind of IPv6 addresses are paired with each other, some of the combinations will never work and may unnecessarily delay the completion of the ICE process.

This document updates the ICE rules defined in [RFC5245] and [RFC6544] on how candidate addresses are selected and how they should be combined with each other in order to maintain high performance for the ICE NAT traversal process.

2. Terminology

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 RFC 2119 [RFC2119].

This document uses the same terminology as ICE (see Section 3 of [RFC5245]) and the following:

Local relayed candidate: a relayed candidate (obtained, e.g., from a TURN server) and included in an ICE offer or answer the agent has or will send.

3. Changes to Candidate Address Selection

This document proposes the following updates to the rules for selecting and combining IPv6 candidate addresses:

- o Instead of RFC 3484 rules, the rules defined in [I-D.ietf-6man-rfc3484bis] MUST be used for determining the candidate priorities. If operating system address preferences are available (e.g., via appropriate API extension), those SHOULD be used instead of default preferences.
- o Deprecated IPv4-compatible IPv6 addresses [RFC4291] and IPv6 site-local unicast addresses [RFC3879] MUST NOT be included in the address candidates.
- o Candidate addresses from link-local addresses MUST NOT be combined with any other candidates except other link-local candidates.
- o Candidate addresses from Unique Local Addresses (ULAs) MUST NOT be combined with any other candidates except other ULA candidates.
- o IPv4-mapped IPv6 addresses MUST NOT be included in the offered candidates unless the application using ICE does not support IPv4 (i.e., is an IPv6-only application [RFC4038]).

The following updates pertain to both IPv4 and IPv6 addresses:

- o Addresses from a loopback interface MUST NOT be included in the candidate addresses.
- o Local relayed candidates MUST NOT be combined with remote host candidates from IPv4 private address space [RFC1918] or IPv6 link-local addresses or ULAs.

4. Negotiating Address Selection Scheme

The prioritization method for the candidate address pairs used by ICE results in matching checklists for both endpoints and hence both endpoints start the checks for the same candidate pair roughly at the same time. This is important since in many scenarios a connectivity check initiated by both endpoints for the same pair is needed before a check for the pair succeeds. Also, some NAT devices have very short timeouts for their address translation bindings and a binding created by a connectivity check from one endpoint may expire before the corresponding connectivity check from the other endpoint is sent if there is a long delay between the two checks.

Depending on how different candidates are paired and whether RFC 3484 or the revised version of it [I-D.ietf-6man-rfc3484bis] is used, the endpoints may end up with different priorities and checklists. Therefore, the endpoints need to agree on how the address selection and pairing is done.

To indicate that the address selection and pairing rules defined in this document are used, the ICE offerer **MUST** include ice-options attribute with "bis-candidates" option identifier in the Session Description Protocol (SDP) [RFC4566] ICE offer. If the ICE offer does not include this option tag, the answerer **SHOULD NOT** utilize the updated rules defined in this document. If the offer included the option tag and the answerer supports this specification, the answerer **SHOULD** add the same option tag to the response and use the updated rules.

If the ICE answer does not contain the option tag, the offerer **SHOULD NOT** use the updated rules. However, even if the other endpoint does not indicate support for the updated rules, loopback addresses or the deprecated IPv6 addresses **SHOULD NOT** be included in the candidates.

5. Security Considerations

The general security considerations for ICE have been documented in Section 18 of [RFC5245] and Section 12 of [RFC6544]. The general security considerations for IPv6 address selection rules have been documented in [I-D.ietf-6man-rfc3484bis]. The vulnerabilities in ICE and RFC3484bis relate to attempts to hijack sessions opened through ICE, denial-of-service attacks, and accidental disclosure of private information. Mechanisms described in [RFC5245] and [RFC6544] - such as validated TCP connections - are designed to protect against connection hijacking.

Denial-of-service attacks can not be completely eliminated, but the amplification capabilities of ICE are limited through a maximum value of concurrently probed connections.

Any address probing mechanism opens up the possibility of outsiders learning the correlation between different IP addresses. For instance, the existence of a privacy address [RFC4941] in the candidate set along with other, more stable addresses will tell at least the peer and maybe eavesdroppers that the addresses are related.

This specification introduces no specific new security concerns beyond these, as it only attempts to unify the algorithms associated with candidate address pair selection. However, where address

selection rules in a node are configured through an external mechanism, as suggested in [I-D.ietf-6man-rfc3484bis], this opens up another avenue for introducing incorrect addresses into the probing mechanism. The resulting system is only as secure as its weakest component. For instance, even if sufficient security mechanisms are in place in ICE, vulnerabilities in the configuration mechanisms for the 3484bis priority tables may introduce weaknesses in the ability of ICE to select the right addresses.

6. IANA Considerations

IANA is requested to register "bis-candidates" option identifier under the "ICE Options" [RFC6336] registry. The required registration information is as follows:

Option identifier: bis-candidates

Contact: Ari Keranen, ari.keranen@ericsson.com

Change control: IETF

Description: Existence of this option identifier indicates that the revised rules (defined in RFCXXXX) are used for candidate address selection.

Reference: RFCXXXX

[RFC editor: replace XXXX with the RFC number of this document]

7. References

7.1. Normative References

- [RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [RFC4193] Hinden, R. and B. Haberman, "Unique Local IPv6 Unicast Addresses", RFC 4193, October 2005.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", RFC 4291, February 2006.

- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC5245] Rosenberg, J., "Interactive Connectivity Establishment (ICE): A Protocol for Network Address Translator (NAT) Traversal for Offer/Answer Protocols", RFC 5245, April 2010.
- [RFC6336] Westerlund, M. and C. Perkins, "IANA Registry for Interactive Connectivity Establishment (ICE) Options", RFC 6336, July 2011.
- [RFC6544] Rosenberg, J., Keranen, A., Lowekamp, B., and A. Roach, "TCP Candidates with Interactive Connectivity Establishment (ICE)", RFC 6544, March 2012.
- [I-D.ietf-6man-rfc3484bis] Thaler, D., Draves, R., Matsumoto, A., and T. Chown, "Default Address Selection for Internet Protocol version 6 (IPv6)", draft-ietf-6man-rfc3484bis-06 (work in progress), June 2012.

7.2. Informative References

- [RFC3484] Draves, R., "Default Address Selection for Internet Protocol version 6 (IPv6)", RFC 3484, February 2003.
- [RFC3879] Huitema, C. and B. Carpenter, "Deprecating Site Local Addresses", RFC 3879, September 2004.
- [RFC4038] Shin, M-K., Hong, Y-G., Hagino, J., Savola, P., and E. Castro, "Application Aspects of IPv6 Transition", RFC 4038, March 2005.
- [RFC4941] Narten, T., Draves, R., and S. Krishnan, "Privacy Extensions for Stateless Address Autoconfiguration in IPv6", RFC 4941, September 2007.

Appendix A. Acknowledgments

The authors would like to thank Jan Melen, Dan Wing, and Jonathan Lennox for comments, reviews and valuable input to the document.

Authors' Addresses

Ari Keranen
Ericsson
Jorvas 02420
Finland

Email: ari.keranen@ericsson.com

Jari Arkko
Ericsson
Jorvas 02420
Finland

Email: jari.arkko@piuha.net

MMUSIC
Internet-Draft
Intended status: BCP
Expires: October 26, 2012

Muthu A M. Perumal
Cisco Systems
Parthasarathi. Ravindran
Sonus Networks
April 24, 2012

Offer/Answer Considerations for G.723 Annex A and G.729 Annex B
draft-muthu-mmusic-offer-answer-g723-g729-00

Abstract

[RFC4856] describes the annexa parameter for G723 and the annexb parameter for G729, G729D and G729E. However, the specification does not describe the offerer and answerer behavior when the value of the annexa or annexb parameter does not match in the SDP offer and answer. This document provides the offer/answer considerations for these parameters.

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 October 26, 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. Offer/Answer Considerations	4
3.1. Offer/Answer Considerations for G723 Annex A	4
3.2. Offer/Answer Considerations for G.729 Annex B, G.729D Annex B and G.729E Annex B	4
4. Examples	5
4.1. Offer with G279 annexb=yes and answer with G279 annexb=no	5
4.2. Offer with G279 annexb=yes and answer with G729 and no annexb parameter	6
4.3. Offer with G279 and no annexb parameter and answer with G729 annexb=no	6
5. Security Considerations	7
6. IANA Considerations	7
7. Acknowledgement	7
8. Normative References	7
Authors' Addresses	8

1. Introduction

[RFC4856] describes the annexa parameter for G723 as follows:

annexa: indicates that Annex A, voice activity detection, is used or preferred. Permissible values are "yes" and "no" (without the quotes); "yes" is implied if this parameter is omitted.

Also, [RFC4856] describes the annexb parameter for G729, G729D and G729E as follows:

annexb: indicates that Annex B, voice activity detection, is used or preferred. Permissible values are "yes" and "no" (without the quotes); "yes" is implied if this parameter is omitted.

However, it does not have any normative statement for the case where the value of this parameter does not match in the SDP offer and answer. For example, if the offer has G729 with annexb=yes and the answer has G729 with annexb=no, it can be interpreted in two different ways:

- o The offerer and answerer proceed as if G729 is negotiated with annexb=yes.
- o The offerer and answerer proceed as if G729 is negotiated with annexb=no.

Since [RFC4856] does not state it clearly, various implementations have interpreted the offer/answer in their own ways, resulting in a different codec being chosen to call failure, when the parameter value does not match in the offer and answer.

[RFC3264] requires SDP extensions that define new fmtp parameters to specify their proper interpretation in offer/answer. But, [RFC4856] does not specify it for the Annex A flavor of G.723 and the Annex B flavors of G.729, G729D and G729E.

This document describes the offer/answer considerations for these parameters and provides the necessary clarifications.

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. Offer/Answer Considerations

[RFC3551] states that

Receivers MUST accept comfort noise frames if restriction of their use has not been signaled. The MIME registration for G729 in RFC 3555 specifies a parameter that MAY be used with MIME or SDP to restrict the use of comfort noise frames.

Based on the above it is best to not use comfort noise frames if the SDP offer or answer indicates no support for it.

3.1. Offer/Answer Considerations for G723 Annex A

When the offer or answer has G723 and the annexa parameter is absent, it MUST be considered as if the offer or answer has G723 with annexa=yes.

When the offer has G723 with annexa=yes and the answer has G723 with annexa=no, the offerer and answerer MUST proceed as if G723 is negotiated with annexa=no.

When the offer has G723 with annexa=no then the answer MUST NOT have annexa=yes for G723. Thus the annexa parameter can be turned off by the answerer, but cannot be turned on.

Open item: Should the above be restated as follows?

When the offer has G723 with annexa=no then the answer MUST have annexa=no for G723.

This is technically correct, but are there implementations that omit the annexa parameter in answer and expect the least common denominator to be used?

When the offer has G723 with no annexa parameter and the answer has G723 with annexa=yes, the offerer and answerer MUST proceed as if G723 is negotiated with annexa=yes.

3.2. Offer/Answer Considerations for G.729 Annex B, G.729D Annex B and G.729E Annex B

In this section G729 represents any of G729 or G729D or G729E.

When the offer or answer has G729 and the annexb parameter is absent, it MUST be considered as if the offer or answer has G729 with annexb=yes.

When the offer has G729 with annexb=yes and the answer has G729 with annexb=no, the offerer and answerer MUST proceed as if G729 is

negotiated with annexb=no.

When the offer has G729 with annexb=no then the answer MUST NOT have annexb=yes for G729. Thus the annexb parameter can be turned off by the answerer, but cannot be turned on.

Open item: Should the above be restated as follows?

When the offer has G729 with annexb=no then the answer MUST have annexb=no for G729.

This is technically correct, but are there implementations that omit the annexb parameter in answer and expect the least common denominator to be used?

When the offer has G.729 with no annexb parameter and the answer has G.729 with annexb=yes, the offerer and answerer MUST proceed as if G.729 is negotiated with annexb=yes.

4. Examples

4.1. Offer with G279 annexb=yes and answer with G279 annexb=no

[Offer with G279 annexb=yes]

```
v=0
o=alice 2890844526 2890844526 IN IP4 host.atlanta.example.com
s=
c=IN IP4 host.atlanta.example.com
t=0 0
m=audio 49170 RTP/AVP 18
a=rtpmap:18 G729/8000
a=fmtp:18 annexb=yes
```

[Answer with G729 annexb=no]

```
v=0
o=bob 1890844326 1890844326 IN IP4 host.bangalore.example.com
s=
c=IN IP4 host.bangalore.example.com
t=0 0
m=audio 19140 RTP/AVP 18
a=rtpmap:18 G729/8000
a=fmtp:18 annexb=no
```

In the above example the offerer and answerer proceed as if G729 is negotiated with annexb=no.

4.2. Offer with G279 annexb=yes and answer with G729 and no annexb parameter

[Offer with G279 annexb=yes]

```
v=0
o=alice 2890844526 2890844526 IN IP4 host.atlanta.example.com
s=
c=IN IP4 host.atlanta.example.com
t=0 0
m=audio 49170 RTP/AVP 18
a=rtpmap:18 G729/8000
a=fmtp:18 annexb=yes
```

[Answer with G729 and no annexb parameter]

```
v=0
o=bob 1890844326 1890844326 IN IP4 host.bangalore.example.com
s=
c=IN IP4 host.bangalore.example.com
t=0 0
m=audio 19140 RTP/AVP 18
a=rtpmap:18 G729/8000
```

In the above example the offerer and answerer proceed as if G729 is negotiated with annexb=yes.

4.3. Offer with G279 and no annexb parameter and answer with G729 annexb=no

[Offer with G279 and no annexb parameter]

```
v=0
o=alice 2890844526 2890844526 IN IP4 host.atlanta.example.com
s=
c=IN IP4 host.atlanta.example.com
t=0 0
m=audio 49170 RTP/AVP 18
a=rtpmap:18 G729/8000
```

[Answer with G729 annexb=no]

```
v=0
o=bob 1890844326 1890844326 IN IP4 host.bangalore.example.com
s=
c=IN IP4 host.bangalore.example.com
t=0 0
m=audio 19140 RTP/AVP 18
a=rtpmap:18 G729/8000
a=fmtp:18 annexb=no
```

In the above example the offerer and answerer proceed as if G729 is negotiated with annexb=no.

5. Security Considerations

There is no extra security consideration apart from what is described in [RFC4856].

6. IANA Considerations

There is no IANA consideration for this draft.

7. Acknowledgement

Thanks to Flemming Andreassen (Cisco), Ali C. Begen (Cisco), Paul Kyzivat, Roni Even (Huawei), Kevin Riley (Sonus), Ashish Sharma (Sonus), Kevin P. Fleming (Digium) and Harprit S. Chhatwal (InnoMedia) for their valuable inputs and comments. Martin Dolly (ATT) and Hadriel Kaplan (Acme Packet) also provided useful suggestions in the MIC at IETF-83.

8. 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.
- [RFC4856] Casner, S., "Media Type Registration of Payload Formats in the RTP Profile for Audio and Video Conferences", RFC 4856, February 2007.

[RFC3551] Schulzrinne, H. and S. Casner, "RTP Profile for Audio and Video Conferences with Minimal Control", STD 65, RFC 3551, July 2003.

Authors' Addresses

Muthu Arul Mozhi Perumal
Cisco Systems
Cessna Business Park
Sarjapur-Marathahalli Outer Ring Road
Bangalore, Karnataka 560103
India

Email: mperumal@cisco.com

Parthasarathi Ravindran
Sonus Networks
Prestige Shantiniketan - Business Precinct
Whitefield Road
Bangalore, Karnataka 560066
India

Email: pravindran@sonusnet.com

MMUSIC
Internet-Draft
Intended status: Standards Track
Expires: January 17, 2013

D. Wing
P. Patil
T. Reddy
Cisco
July 16, 2012

Mobility with ICE (MICE)
draft-wing-mmusic-ice-mobility-01

Abstract

This specification describes how endpoint mobility can be achieved using ICE. Two mechanisms are shown, one where both endpoints support ICE and another where only one endpoint supports ICE.

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	3
2. Notational Conventions	3
3. Mobility using ICE	4
3.1. Gaining an Interface	4
3.2. Receiving ICE Mobility event	5
3.3. Losing an Interface	6
3.4. New STUN Attributes MOBILITY-EVENT and MOBILITY-SUPPORT	6
4. Mobility using TURN	6
4.1. Creating an Allocation	7
4.1.1. Sending an Allocate Request	7
4.1.2. Receiving an Allocate Request	8
4.1.3. Receiving an Allocate Success Response	8
4.1.4. Receiving an Allocate Error Response	8
4.2. Refreshing an Allocation	9
4.2.1. Sending a Refresh Request	9
4.2.2. Receiving a Refresh Request	9
4.2.3. Receiving a Refresh Response	9
4.3. New STUN Attribute MOBILITY-TICKET	10
4.4. New STUN Error Response Code	10
5. IANA Considerations	10
6. Security Considerations	10
6.1. Considerations for ICE mechanism	10
6.2. Considerations for TURN mechanism	11
7. Acknowledgements	11
8. References	11
8.1. Normative References	11
8.2. Informative References	11
Authors' Addresses	11

1. Introduction

When moving between networks, an endpoint has to change its IP address. This change breaks upper layer protocols such as TCP and RTP. Various techniques exist to prevent this breakage, all tied to making the endpoint's IP address static (e.g., Mobile IP, Proxy Mobile IP, LISP). Other techniques exist, which make the upper layer protocol ambivalent to IP address changes (e.g., SCTP). The mechanisms described in this document are in that last category.

ICE [RFC5245] ensures two endpoints have a working media path between them, and is typically used by Internet-connected interactive media systems (e.g., SIP endpoints). ICE does not expect either the local host or the remote host to change their IP addresses. Although ICE does allow an "ICE restart", this is done by sending a re-INVITE which goes over the SIP signaling path. The SIP signaling path is often slower than the media path (which needs to be recovered as quickly as possible), consumes an extra half round trip, and incurs an additional delay if the mobility event forces the endpoint to re-connect with its SIP proxy. Thus, this document attempts to perform mobility entirely on the media path.

A TURN [RFC5766] server relays media packets and is used for a variety of purposes, including overcoming NAT and firewall traversal issues and IP address privacy. The existing TURN specification does not allow the client address to change, especially if multiple clients share the same TURN username (e.g., the same credentials are used on multiple devices).

This document proposes two mechanisms to achieve RTP mobility: a mechanism where both endpoints support ICE, and a mechanism where only one endpoint supports ICE. When both endpoints support ICE, ICE itself can be used to provide mobility. When only one endpoint supports ICE, a TURN server provides mobility. Both mobility techniques work across and between network types (e.g., between 3G and wired Internet access), so long as the client can still access the remote ICE peer or TURN server.

Readers are assumed to be familiar with ICE [RFC5245].

2. Notational 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 [RFC2119].

This note uses terminology defined in [RFC5245].

3. Mobility using ICE

When both endpoints support ICE, ICE itself can provide mobility functions. One of the primary aspects of ICE is its address gathering, wherein ICE has each endpoint determine all of the IP addresses and ports that might be usable for that endpoint and communicate that list of addresses and ports to its peer, usually over SDP. That enables the next primary aspect of ICE, which is its connectivity checks: each ICE endpoint sends a connectivity check to that list of addresses and ports. A connectivity check may unknowingly traverse a NAT, which means the ICE endpoint receiving the connectivity check cannot validate the source IP address or port of the connectivity against the list of IP addresses and ports provided by the ICE peer. In fact, if the source IP address and port is not known to the ICE endpoint, it is added to the list of candidates (Section 7.2.1.3 of [RFC5245]).

ICE Mobility takes advantage of that existent ICE functionality. Media can be switched to the new interface before or after the previous interface is lost.

When an interface is lost, media traffic might or might not be utilizing that interface. If media traffic is currently traversing the interface, this is considered a "break before make", because the host has not already moved its media traffic to a different interface.

Endpoints that support ICE Mobility perform ICE normally, and MUST also include the MOBILITY-SUPPORT attribute in all of their STUN requests and their STUN responses. The inclusion of this attribute allows the ICE peer to determine if it can achieve mobility using ICE or needs to use TURN (or needs to use some other mechanism, such as Mobile IP). To force the use of TURN to achieve ICE mobility, the ICE endpoint SHOULD NOT respond to ICE connectivity checks that have an IP address and port different from the TURN server, unless those connectivity checks contain the MOBILITY-SUPPORT attribute. In this way, the remote peer will think those other candidates are invalid (because its connectivity checks did not succeed).

After concluding ICE and moving to the ICE completed state (see Section 8 of [RFC5245] either endpoint or both endpoints can initiate ICE Mobility, no matter if it was the Controlling Agent or the Controlled Agent during normal ICE processing.

3.1. Gaining an Interface

When gaining an interface which is suitable to send media by the host's policy (if any), the ICE endpoint performs ICE Mobility. ICE

Mobility is performed by:

1. The ICE endpoint clears its ICE check list.
2. The ICE endpoint initiating an ICE connectivity check on the new interface, with the MOBILITY-EVENT attribute.
3. If this interface is the only suitable interface for media (that is, other suitable interfaces have been lost), then the connectivity check from the previous step SHOULD also include the USE-CANDIDATE attribute to signals an aggressive nomination (see Section 2.6 of [RFC5245]), and media MAY immediately begin flowing over that interface.
4. The ICE endpoint performs Sections 7.2.1.3, 7.2.1.4, and 7.2.1.5 of [RFC5245].
5. If the ICE connectivity check succeeds then ICE agents creates a new pair, adds the pair to the valid list and marks it as selected. The ICE agent can now send media using the newly selected candidate pair, even if it is running in Regular Nomination mode.
6. Once ICE connectivity checks for all of the media streams are completed, the controlling ICE endpoint follows the procedures in Section 11.1 of [RFC5245], specifically to send updated offer if the candidates in the m and c lines for the media stream (called the DEFAULT CANDIDATES) do not match ICE's SELECTED CANDIDATES (also see Appendix B.9 of [RFC5245]).

3.2. Receiving ICE Mobility event

A STUN Binding Request containing the MOBILITY-EVENT attribute MAY be received by an ICE endpoint. If this is received before the endpoint is in the ICE Concluded state, it should be silently discarded.

The agent remembers the highest-priority nominated pairs in the Valid list for each component of the media stream, called the previous selected pairs. It continues sending media to that address until it finishes with the steps described below. Because those packets might not be received due to the mobility event, it MAY cache a copy of those packets.

The ICE endpoint clears its ICE check list.

The ICE endpoint performs Sections 7.2.1.3, 7.2.1.4, and 7.2.1.5 of [RFC5245].

3.3. Losing an Interface

When an interface is lost, the SDP MAY be updated, so that the remote ICE host does not waste its efforts with connectivity checks to that address, as those checks will fail. Because it can be argued that this is merely an optimization, and that the interface loss might be temporary (and soon regained), and that ICE has reasonable accommodation for candidates where connectivity checks timeout, this specification does not strongly encourage updating the SDP to remove a lost interface. Likewise, this specification recommends that ICE candidate addresses be maintained actively, subject to the host's policy. For example, battery operated hosts have a strong incentive to not maintain mappings to TURN servers, as that maintenance requires periodic keepalive messages. As another example, a host that is receiving media over IPv6 may not want to persist with keeping a NATted IPv4 mapping alive (because that consumes a NAT mapping that could be more useful to a host actively utilizing the mapping for real traffic).

Note: this differs from Section 8.3 of [RFC5245], which encourages abandoning un-used candidates.

Note: A future version of this document will have more normative language in this section.

3.4. New STUN Attributes MOBILITY-EVENT and MOBILITY-SUPPORT

Two new attributes are defined by this section: MOBILITY-EVENT and MOBILITY-SUPPORT.

The MOBILITY-EVENT attribute indicate the sender experienced a mobility event. This attribute has no value, thus the attribute length field MUST always be 0. Rules for sending and interpretation of receiving are described above.

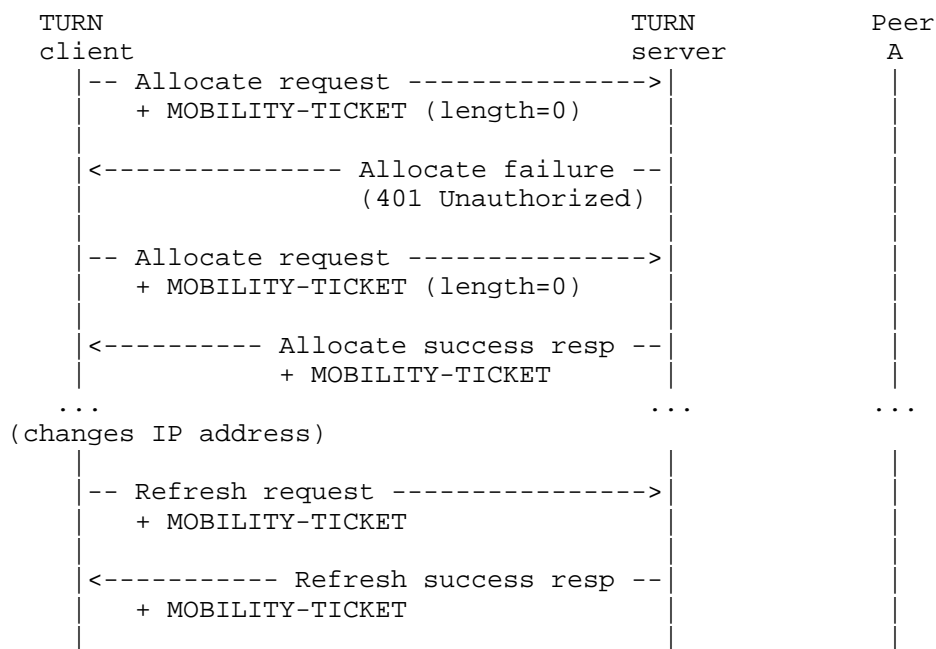
The MOBILITY-SUPPORT attribute indicates the sender supports ICE Mobility, as defined in this document. This attribute has no value, thus the attribute length field MUST always be 0. Rules for sending and interpretation of receiving are described above.

4. Mobility using TURN

To achieve mobility, a TURN client should be able to retain an allocation on the TURN server across changes in the client IP address as a consequence of movement to other networks.

When the client sends the initial Allocate request to the TURN

server, it will also include the new STUN attribute MOBILITY-TICKET (with zero length value), which indicates that the client is capable of mobility and desires a ticket. The TURN server provisions a ticket that is sent inside the new STUN attribute MOBILITY-TICKET in the Allocate Success response to the client. The ticket will be used by the client when it wants to refresh the allocation but with a new client IP address and port. It also ensures that the allocation can only be refreshed this way by the same client. When a client's IP address changes due to mobility, it presents the previously obtained ticket in a Refresh Request to the TURN server. If the ticket is found to be valid, the TURN server will retain the same relayed address/port for the new IP address/port allowing the client to continue using previous channel bindings -- thus, the TURN client does not need to obtain new channel bindings. Any data from external peer will be delivered by the TURN server to this new IP address/port of the client. The TURN client will continue to send application data to its peers using the previously allocated channelBind Requests.



4.1. Creating an Allocation

4.1.1. Sending an Allocate Request

In addition to the process described in Section 6.1 of [RFC5766], the client includes the MOBILITY-TICKET attribute with length 0. This

indicates the client is a mobile node and wants a ticket.

4.1.2. Receiving an Allocate Request

In addition to the process described in Section 6.2 of [RFC5766], the server does the following:

If the MOBILITY-TICKET attribute is included, and has length zero, and the TURN session mobility is forbidden by local policy, the server MUST reject the request with the new Mobility Forbidden error code. Following the rules specified in [RFC5389], if the server does not understand the MOBILITY-TICKET attribute, it ignores the attribute.

If the server can successfully process the request create an allocation, the server replies with a success response that includes a STUN MOBILITY-TICKET attribute. TURN server stores it's session state, such as 5-tuple and NONCE, into a ticket that is encrypted by a key known only to the TURN server and sends the ticket in the STUN MOBILITY-TICKET attribute as part of Allocate success response.

The ticket is opaque to the client, so the structure is not subject to interoperability concerns, and implementations may diverge from this format. TURN Allocation state information is encrypted using 128-bit key for Advance Encryption Standard (AES) and 256-bit key for HMAC-SHA-256 for integrity protection.

4.1.3. Receiving an Allocate Success Response

In addition to the process described in Section 6.3 of [RFC5766], the client will store the MOBILITY-TICKET attribute, if present, from the response. This attribute will be presented by the client to the server during a subsequent Refresh request to aid mobility.

4.1.4. Receiving an Allocate Error Response

If the client receives an Allocate error response with error code TBD (Mobility Forbidden), the error is processed as follows:

o TBD (Mobility Forbidden): The request is valid, but the server is refusing to perform it, likely due to administrative restrictions. The client considers the current transaction as having failed. The client MAY notify the user or operator and SHOULD NOT retry the same request with this server until it believes the problem has been fixed.

All other error responses must be handled as described in [RFC5766].

4.2. Refreshing an Allocation

4.2.1. Sending a Refresh Request

If a client wants to refresh an existing allocation and update its time-to-expiry or delete an existing allocation, it will send a Refresh Request as described in Section 7.1 of [RFC5766]. If the client wants to retain the existing allocation in case of IP change, it will include the MOBILITY-TICKET attribute received in the Allocate Success response. If a Refresh transaction was previously made, the MOBILITY-TICKET attribute received in the Refresh Success response of the transaction must be used.

4.2.2. Receiving a Refresh Request

In addition to the process described in Section 7.2 of [RFC5766], the client does the following:

If the STUN MOBILITY-TICKET attribute is included in the Refresh Request then the server will not retrieve the 5-tuple from the packet to identify an associated allocation. Instead TURN server will decrypt the received ticket, verify the ticket's validity and retrieve the 5-tuple allocation from the contents of the ticket. If this 5-tuple obtained from the ticket does not identify an existing allocation then the server MUST reject the request with an error.

If the source IP address and port of the Refresh Request is different from the stored 5-tuple allocation, the TURN server proceeds with checks to see if NONCE in the Refresh request is the same as the one provided in the ticket. The TURN server also uses MESSAGE-INTEGRITY validation to identify that it is the same user which had previously created the TURN allocation. If the above checks are not successful then server MUST reject the request with a 441 (Wrong Credentials) error.

If all of the above checks pass, the TURN server understands that the client has moved to a new network and acquired a new IP address. The source IP address of the request could either be the host transport address or server-reflexive transport address. The server then updates its 5-tuple with the new client IP address and port. TURN server calculates the ticket with the new 5-tuple and sends the new ticket in the STUN MOBILITY-TICKET attribute as part of Refresh Success response.

4.2.3. Receiving a Refresh Response

In addition to the process described in Section 7.3 of [RFC5766], the client will store the MOBILITY-TICKET attribute, if present, from the

response. This attribute will be presented by the client to the server during a subsequent Refresh Request to aid mobility.

4.3. New STUN Attribute MOBILITY-TICKET

This attribute is used to retain an Allocation on the TURN server. It is exchanged between the client and server to aid mobility. The value is encrypted and identifies session state such as 5-tuple and NONCE. The value of MOBILITY-TICKET is a variable-length value.

4.4. New STUN Error Response Code

This document defines the following new error response code:

Mobility Forbidden: Mobility request was valid but cannot be performed due to administrative or similar restrictions.

5. IANA Considerations

IANA is requested to add the following attributes to the STUN attribute registry [iana-stun],

- o MOBILITY-TICKET (0x802E, in the comprehension-optional range)
- o MOBILITY-EVENT (0x802, in the comprehension-required range)
- o MOBILITY-SUPPORT (0x8000, in the comprehension-optional range)

and to add a new STUN error code "Mobility Forbidden" with the value 501 to the STUN Error Codes registry [iana-stun].

6. Security Considerations

6.1. Considerations for ICE mechanism

A mobility event only occurs after both ICE endpoints have exchanged their ICE information. Thus, both username fragments are already known to both endpoints. Each endpoint contributes at least 24 bits of randomness to the ice-ufrag (Section 15.4 of [RFC5245]), which provides 48 bits of randomness. An off-path attacker would have to guess those 48 bits to cause the endpoints to perform HMAC-SHA1 validation of the MESSAGE-INTEGRITY attribute.

An attacker on the path between the ICE endpoints will see both ice-ufrags, and can cause the endpoints to perform HMAC-SHA1 validation

by sending messages from any IP address.

6.2. Considerations for TURN mechanism

TURN server MUST use strong encryption and integrity protection for the ticket to prevent an attacker from using a brute force mechanism to obtain the ticket's contents or refreshing allocations.

Security considerations described in [RFC5766] are also applicable to this mechanism.

7. Acknowledgements

Thanks to Alfred Heggstad for his review and comments.

8. References

8.1. Normative References

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.
- [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.
- [RFC5766] Mahy, R., Matthews, P., and J. Rosenberg, "Traversal Using Relays around NAT (TURN): Relay Extensions to Session Traversal Utilities for NAT (STUN)", RFC 5766, April 2010.

8.2. Informative References

- [iana-stun] IANA, "IANA: STUN Attributes", April 2011, <<http://www.iana.org/assignments/stun-parameters/stun-parameters.xml>>.

Authors' Addresses

Dan Wing
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, California 95134
USA

Email: dwing@cisco.com

Prashanth Patil
Cisco Systems, Inc.
Cessna Business Park, Varthur Hobli
Sarjapur Marthalli Outer Ring Road
Bangalore, Karnataka 560103
India

Email: praspati@cisco.com

Tirumaleswar Reddy
Cisco Systems, Inc.
Cessna Business Park, Varthur Hobli
Sarjapur Marathalli Outer Ring Road
Bangalore, Karnataka 560103
India

Email: tiredddy@cisco.com

Multiparty Multimedia Session Control
Internet-Draft
Intended status: Standards Track
Expires: September 27, 2012

S. Zhou, Ed.
T. Tian
Z. Xie
ZTE Corporation
March 26, 2012

Security Descriptions Extension for Media Streams
draft-zhou-mmusic-sdes-keymod-01

Abstract

This document provides an extension to the cryptographic attribute (RFC 4568) defined for Session Description Protocol (RFC 4566) to enhance end-to-end communication security, so that some scenarios, e.g., forking and re-targeting can especially benefit from the extension. The usage of the provided extension in Secure Real-time Transport Protocol (SRTP, RFC3711) is also defined in this document.

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 27, 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
1.1. Terminology	3
2. Extension to SDES	3
3. Usage of keymod with Offer/Answer	4
3.1. Generating the Initial Offer - Unicast Streams	5
3.2. Generating the Initial Answer - Unicast Streams	5
3.3. Processing of the Initial Answer - Unicast Streams	6
4. Example	6
5. Applicability in Re-targeting Scenarios	7
5.1. Single CDIV instance	7
5.2. Multiple CDIV instances	8
5.3. Computation of K1'	9
6. Applicability in Forking Scenarios	9
7. IANA Considerations	10
8. Security Considerations	10
9. References	10
9.1. Normative References	10
9.2. Informative References	11
Authors' Addresses	11

1. Introduction

To ensure the media security established by Session Initiation Protocol (SIP), SDP Security Descriptions (SDS) is defined in RFC 4568 [RFC4568], where a cryptographic attribute and application in Secure Real-time Transport Protocol (SRTP, RFC 3711 [RFC3711]) unicast media streams are provided.

SDP Security Descriptions (SDS) is essentially a key transportation scheme in offer/answer model, in which keying material for the direction from offerer to answerer is chosen independently by the offerer and transported in clear text, the keying material for the reverse direction is also chosen independently by the answerer and transported in clear. Later the transported keying materials are provided to SRTP protocol to secure outgoing or incoming media communication. The protection of the transported keying materials obviously relies on the security of the signaling protocol which is beyond the scope of this document.

When SDS is applied in some scenarios, e.g., forking and re-targeting, the intermediate users and devices besides the ultimate answerer also have knowledge of the keying material used for the outgoing media from the offerer, which is a security threat to the content of the end-to-end communication in the affected direction.

To resolve the problem, it is suggested exchanging a new pair of offer/answer with a new key between the offerer and the ultimate answerer, i.e., by using SIP UPDATE message [RFC3311], but it will require more round trip messages. In this document, a resolution is introduced based on the defined SDS extension.

1.1. 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].

2. Extension to SDS

Following the ABNF format in Security Descriptions, a new session parameter extension "keymod" is defined as follows:

```
srtp-session-extension = keymod
keymod                  = "keymod:" <keymod-info>
keymod-info             = <keymod-type> "|" <kdf-func> "|" <keymod-val>
keymod-type             = "rand"/"rand-salt"/keymod-type-ext
keymod-type-ext         = 1*(VCHAR)
kdf-func                = 1*(ALPHA / DIGIT / "_")
keymod-val              = *(base64);base64 encoded binary string
base64                  = ALPHA/DIGIT/"+"/" "/" "="
```

where base64 encoding follows RFC3548 [RFC3548], ALPHA, DIGIT, and VCHAR are defined in RFC4234 [RFC4234].

The defined "keymod" is a negotiated parameter, which indicates it does not apply to data sent from the answerer to the offerer, as defined in RFC 4568 [RFC4568].

An answer MAY contain keymod value indicating the answerer is asking for the offerer to refresh its keying material using the information following it.

If keymod-type is "rand", then only master key is requested to refresh according to specified function kdf-func;

If keymod-type is "rand-salt", then master key and master salt are both requested to refresh, the master key will be refreshed according to specified function kdf-func and the refresh method of master salt is simply replacement in this document.

The key derivation function kdf-func can be as simple as an assignment (defined as "is"), or an XOR between the old master key and the keymod-val value (defined as "xor"), or as complicated as any other key derivation functions based on cryptographic primitives, e.g., RFC 2104 [RFC2104].

In this document, only the two simple functions are defined: "is" and "xor", that is

```
kdf-func = "is"/"xor"/kdf-func-ext
```

```
kdf-func-ext= 1*(ALPHA / DIGIT / "_")
```

And if no kdf-func is indicated in keymod-info, the default kdf-func is "is".

3. Usage of keymod with Offer/Answer

3.1. Generating the Initial Offer - Unicast Streams

The generation of the initial offer for a unicast stream MUST follow that of the crypto attribute RFC4568 [RFC4568], and MAY

also include an additional "keymod" parameter with keymod-val being NULL. It indicates to the ultimate answerer that the offerer wants to employ the mechanism specified in

this document, a key agreement mechanism with a higher security level than the original SDES.

3.2. Generating the Initial Answer - Unicast Streams

The generation of the initial answer for a unicast stream MUST follow that of the crypto attribute RFC4568 [RFC4568], and if the offer message includes a "keymod" parameter, it SHOULD also include an additional "keymod" parameter. That is, when an offered crypto attribute is accepted, the crypto attribute in the answer MUST contain the following:

- o The tag and crypto-suite from the accepted crypto attribute in the offer (the same crypto-suite MUST be used in the send and receive direction).
- o The key(s) the answerer will be using for media sent to the offerer.

Additionally the answer MAY contain:

- o The keymod parameter for media sent from the offerer to the answerer.

The keymod parameter is constrained by the following limits:

- o If keymod type is "rand", the keymod-val value MUST be at the minimum length required by the specified crypto-suite for the master key.
- o If keymod type is "rand-salt", the keymod-val value length MUST be no less than the addition of the minimum lengths of master key and master salt required by the specified crypto-suite.

The keymod parameter and the master key retrieved from the offer message MAY be used together to derive a new master key used for the media from the offerer to the answerer.

3.3. Processing of the Initial Answer - Unicast Streams

When the offerer receives the answer, the offerer MUST do necessary verifications following RFC 4568 [RFC4568].

If the answer includes a "keymod" value in "crypto" attribute, the offerer MUST derive a new master key from the previous master key sent in the offer message and the keymod-info value received in the answer message.

Specifically, if the keymod type retrieved from the answer message is "rand", a new master key will be derived from the previous master key and the keymode-val value according to specified key derivation function kdf-func.

If the keymod type retrieved from the answer message is "rand-salt", a new master key will be derived from the previous master key and the keymode-val value according to specified key derivation function kdf-func, and the master salt will be replaced with the salt value contained in the keymode-val.

The derived new master key and new master salt will be used to protect the media from the offerer to the answerer.

4. Example

This example shows use of the keymod extension described in this document. The "a=crypto" line is actually a one long line, which is shown as two lines due to page formatting.

The following is an offer using crypto attribute indicating deploying keymod, asking the answerer to return a keymod value :

```
v=0
o=alice 2890844730 2890844731 IN IP4 host.example.com
s=
c=IN IP4 192.0.2.1
t=0 0
m=audio 20000 RTP/AVP 0
a=crypto:1 AES_CM_128_HMAC_SHA1_80
  inline:d0RmdmcmVCspeEc3QGZiNWpVLFJhQX1cfHAWJSoj|2^20|1:32
  keymod:rand|xor|
```

The following is an answer with the keymod extension where type "rand" is chosen and the refreshment of master key is "xor":

```
v=0
o=Bob 2890844725 2890844725 IN IP4 host.example.org
s=
c=IN IP4 192.0.2.2
t=0 0
m=audio 30000 RTP/AVP 0
a=crypto:1 AES_CM_128_HMAC_SHA1_32
  inline:NzB4d1BINUAvLEw6UzF3WSJ+PSdFcGdUJShpX1Zj|2^20|1:32;
  keymod:rand|xor|WVNfX19zZW1jdGwgKCKgew==
```

The following is an answer with the keymod extension where type "rand-salt" is chosen and the refreshments of master key and master salt are both "is":

```
v=0
o=Bob 2890844725 2890844725 IN IP4 host.example.org
s=
c=IN IP4 192.0.2.2
t=0 0
m=audio 30000 RTP/AVP 0
a=crypto:1 AES_CM_128_HMAC_SHA1_32
  inline:NzB4d1BINUAvLEw6UzF3WSJ+PSdFcGdUJShpX1Zj|2^20|1:32;
  keymod:rand-salt|WVNfX19zZW1jdGwgKCKgewkyMjA7fQp9CnVubGVz
```

5. Applicability in Re-targeting Scenarios

In this section, applicability of the defined keymod parameter in re-targeting scenarios is provided.

Re-targeting, or Communications Diversion (CDIV) service is a widely used communication service which enables a served user to divert the communications addressed to the served user's address to another destination according to the specified service type. As define in RFC 4458 [RFC4458] and 3GPP TS 24.604 [TS], there are several conditions that may incur a CDIV service, e.g., when the served user is at the statuses of "Not reachable" , "User busy", "No reply", or the served user has registered with the CDIV Agent Server (AS) to redirect the call unconditionally. The redirected destination may be another call number or a voice mailbox of the same user. CDIV may happen multiple times consecutively till the last destination, see the example below.

5.1. Single CDIV instance

See Figure 1, A initiates a call to B by including a crypto attribute with a key parameter K1 and an empty KEYMOD1 in the SIP message. B has subscribed a CDIV service to divert calls to C. When the

diversion condition is met, the call is re-invited by the Proxy or CDIV AS to C. Proxy sends re-invite SIP message which includes K1, KEYMOD1 and an additional "cause" value to C (the usage and the specification of the CAUSE parameter refers to RFC 4458 [RFC4458] , then C determines it a CVID call and responds with a SIP message with a key parameter K2 and a keymod parameter KEYMOD2. When A receives the SIP message including K2 and KEYMOD2, A will derive a new key parameter K1' from K1 and KEYMOD2 the same way as C. Thus the communication between A and C is protected by K2 and K1', i.e., A uses K1' to protect the media sent from A to C, and C uses K2 to protect the media sent from C to A.

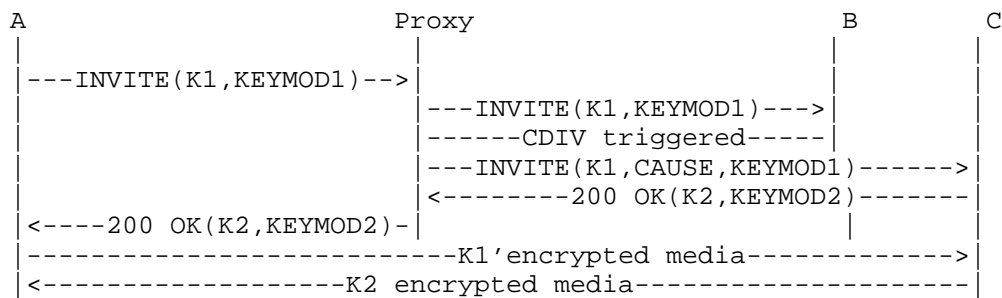


Figure 1

5.2. Multiple CDIV instances

See Figure 2, A initiates a call to B by including a crypto attribute with a key parameter K1 and an empty KEYMOD1 in the SIP message. B has subscribed a CDIV service to divert calls to C. When the diversion condition for B is met, the call is re-invited by the CDIV AS to C. C has also subscribed a CDIV service to divert calls to D. When the diversion condition for C is met, the call is re-invited by the Proxy or CDIV AS to D. Proxy sends re-invite SIP message which includes K1, KEYMOD1 and an additional "cause" value to D (the usage and the specification of the CAUSE parameter refers to RFC 4458 [RFC4458], then D determines it a CVID call and responds with a SIP message with a key parameter K2 and a keymod parameter KEYMOD2. When A receives the SIP message including K2 and KEYMOD2, A will derive a new key parameter K1' from K1 and KEYMOD2 the same way as D. Thus the communication between A and D is protected by K2 and K1', i.e., A uses K1' to protect the media sent from A to D, and D uses K2 to protect the media sent from D to A.

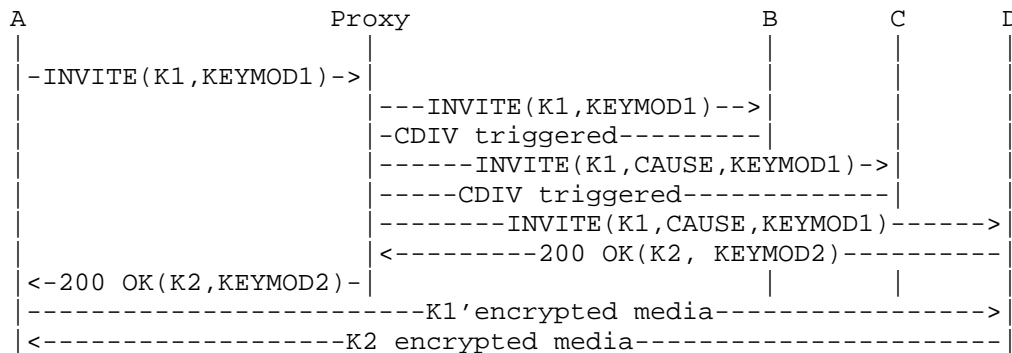


Figure 2

5.3. Computation of K1'

In the above examples, if key method "inline" is used in key parameter. K1 consists of a master key msk1 and a master salt mss1, K2 consists of a master key msk2 and a master salt mss2.

If keymod type is "rand", the keymod-val contained in KEYMOD2 is used to calculate the new master key:

```
msk1'=kdf-func(keymod-val, msk1)
```

If keymod type is "rand-salt", the keymod-val contained in KEYMOD2 can be divided into two parts, key and salt, a new master key will be calculated as:

```
msk1'=kdf-func(keymod-val(key), msk1)
```

and a new master salt will be:

```
mss1'=keymod-val(salt).
```

6. Applicability in Forking Scenarios

In this section, applicability of the defined keymod parameter in forking scenarios is provided, see the example below.

See Figure 3, A initiates a call to a user U by including a crypto attribute with a key parameter K1, an empty KEYMOD1 in the SIP message. And U has multiple devices, e.g., B,C,D, then the call is forked to all the devices till user U answers the call from D. D responds with a SIP message with a key parameter K2 and a keymod parameter KEYMOD2. When A receives the SIP message including K2 and

KEYMOD2, A will derive a new key parameter K1' from K1 and KEYMOD2 the same way as D. Thus the communication between A and D is protected by K2 and K1', i.e., A uses K1' to protect the media sent from A to D, and D uses K2 to protect the media sent from D to A. The computation of K1' is exactly the same as in Section 5.3

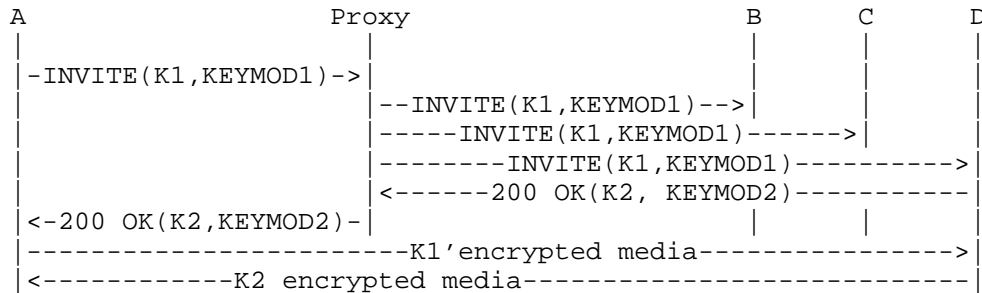


Figure 3

7. IANA Considerations

This document includes no request to IANA.

8. Security Considerations

This document includes an extension to the crypto attribute defined in RFC 4568 [RFC4568], so the security considerations are mostly the same, except that the described solution improves a security drawback when RFC 4568 [RFC4568] is applied in some specific scenarios, i.e., forking and re-targeting.

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.
- [RFC3548] Josefsson, S., "The Base16, Base32, and Base64 Data Encodings", RFC 3548, 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.

- [RFC4234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 4234, October 2005.
- [RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, July 2006.
- [RFC4568] Andreasen, F., Baugher, M., and D. Wing, "Session Description Protocol (SDP) Security Descriptions for Media Streams", RFC 4568, July 2006.

9.2. Informative References

- [RFC2104] Krawczyk, H., Bellare, M., and R. Canetti, "HMAC: Keyed-Hashing for Message Authentication", RFC 2104, February 1997.
- [RFC3311] Rosenberg, J., "The Session Initiation Protocol (SIP) UPDATE Method", RFC 3311, October 2002.
- [RFC4458] Jennings, C., Audet, F., and J. Elwell, "Session Initiation Protocol (SIP) URIs for Applications such as Voicemail and Interactive Voice Response (IVR)", RFC 4458, April 2006.
- [TS] "3GPP TS 24.604 Communication Diversion (CDIV) using IP Multimedia (IM) Core Network (CN) subsystem; Protocol specification".

Authors' Addresses

Sujing Zhou (editor)
ZTE Corporation
No.68 Zijinghua Rd. Yuhuatai District
Nanjing, Jiang Su 210012
R.R.China

Email: zhou.sujing@zte.com.cn

Tian Tian
ZTE Corporation
No.68 Zijinghua Rd. Yuhuatai District
Nanjing, Jiang Su 210012
P.R.China

Phone: +86-025-5287-7867
Email: tian.tian1@zte.com.cn

Zhenhua Xie
ZTE Corporation
No.68 Zijinghua Rd. Yuhuatai District
Nanjing, Jiang Su 210012
P.R.China

Phone: +86-25-52871287
Fax: +86-25-52871000
Email: xie.zhenhua@zte.com.cn

