

CLUE WG
Internet-Draft
Intended status: Standards Track
Expires: November 28, 2012

R. Even
Huawei Technologies
May 27, 2012

**Mapping RTP streams to CLUE media captures
draft-even-clue-rtp-mapping-02.txt**

Abstract

This document describes mechanisms and recommended practice for mapping RTP media streams defined in SDP to CLUE media captures.

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 28, 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.	Mapping CLUE Media Captures to RTP streams	4
4.	Acknowledgements	8
5.	IANA Considerations	8
6.	Security Considerations	8
7.	References	9
7.1.	Normative References	9
7.2.	Informative References	9
	Author's Address	9

Even

Expires November 28, 2012

[Page 2]

1. Introduction

Telepresence systems can send and receive multiple media streams. The CLUE framework [[I-D.ietf-clue-framework](#)] defines media captures as a source of Media, such as from one or more Capture Devices. A Media Capture (MC) may be the source of one or more Media streams. A Media Capture may also be constructed from other Media streams. A middle box can express Media Captures that it constructs from Media streams it receives.

SIP offer answer [[RFC3264](#)] uses SDP [[RFC4566](#)] to describe the RTP[RFC3550] media streams. Each RTP stream has a payload type number and SSRC. The content of the RTP stream is created by the encoder in the endpoint. This may be an original content from a camera or a content created by an intermediary device like an MCU.

The Telepresence systems MUST work in point to point calls and multipoint calls when there is a central multipoint Control unit (MCU). They should work in the RTP topologies defined in [[RFC5117](#)]. There may be some topologies that do not scale well with SIP offer answer like Topo-Translator. The assumption here is that when handling the RTP streams the MCU works using one of the other topologies Topo-Mixer, Topo-vide-switch-MCU or Topo-RTCP-Terminating-MCU. The major difference between these topologies is on how RTCP and CSRC are conveyed and the issue of identifying the original source of the RTP streams need to be discussed. Note that the Topo-RTCP-Terminating-MCU do not convey the CSRC information and needs some other means to identify the original source of the contributing RTP streams.

The relation between the SDP and CLUE descriptions is that the CLUE media capture defined adds some semantics describing the content of an RTP stream and the relations between the streams that is not specified in SDP like spatial relation between the media captures with regards to the conference room.

This document discusses the relation between the CLUE media captures and the RTP streams and makes recommendations on how to co-relate between the two. It tries to use SDP attributes when possible in order to not duplicate information.

2. Terminology

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

Even

Expires November 28, 2012

[Page 3]

3. Mapping CLUE Media Captures to RTP streams

The SDP description of the RTP streams provides among others the information about the encoding and decoding capabilities of the media as well as transport addresses for receiving the RTP and RTCP streams and some channels capabilities like the maximum bandwidth available for the data.

The CLUE MCs provide semantic information of the streams which may be its spatial information like left camera or on its content like loudest speaker.

An RTP stream can have different content based on the MC description. For example one video capture from a camera may capture a third of the room while another from the same camera may provide a zoom version of the whole room. These two media captures can be mapped to the same RTP stream and they are mutually exclusive using the same physical device.

Note: In this case the same SSRC will be used for two MCs since they originate from the same source. The consumer asking to switch between these two MCs may not be able to identify the switch based on the RTP information. This may require the provider to acknowledge the request in order to let the consumer know that its request was accepted.

Using the video capture example from the framework [\[I-D.ietf-clue-framework\]](#) document:

- o VC0- (the camera-left camera stream, purpose=main, switched:no
- o VC1- (the center camera stream, purpose=main, switched:no
- o VC2- (the camera-right camera stream), purpose=main, switched:no
- o VC3- (the loudest panel stream), purpose=main, switched:yes
- o VC4- (the loudest panel stream with PiPs), purpose=main, composed=true; switched:yes
- o VC5- (the zoomed out view of all people in the room), purpose=main, composed=no; switched:no
- o VC6- (presentation stream), purpose=presentation, switched:no

Where the physical simultaneity information is:

Even

Expires November 28, 2012

[Page 4]

{VC0, VC1, VC2, VC3, VC4, VC6}

{VC0, VC2, VC5, VC6}

To describe the above MCs Media Captures we need 6 RTP streams based on the first simultaneous entry. The RTP stream used for VC1 is also the RTP stream for VC5 coming from the same camera, it may even use the same RTP payload type number since the only difference between the two is the view port. This will mean that the mapping from a Media capture to an RTP stream is fixed. The number of RTP streams defined in the SDP depends on the number of non mutually exclusive captures in the capture scene.

The offer answer exchange should work with systems that do not support CLUE protocol or multiple streams. It should also allow both sides to learn if CLUE is supported. One approach is to have two stage negotiation offering a single audio and maybe also a single video that will allow a connection with media to be established. This exchange will also enable both sides to learn if CLUE is supported and start a second exchange that will list the available media streams.

Note: If CLUE decides to mandate RTP multiplexing it will make sense to negotiate in the initial offer at least all SDP media streams in order allow for ICE [[RFC5245](#)] negotiation.

Another approach is to use the same logic specified in the [[I-D.ietf-mmusic-sdp-bundle-negotiation](#)] which uses a new grouping attribute and allows offering of multiple media lines in the first offer.

The proposal bellow can be used as the initial offer if using the bundle approach or for the second offer/answer exchange if using two stage negotiation. It provides all the individual encoding information and specify the mapping between the SDP media lines and the different CLUE media captures.

The example bellow uses a separate UDP port for each m-line but multiplexing grouping as specified in [[I-D.ietf-mmusic-sdp-bundle-negotiation](#)] can be applied here.

This mapping can be done by defining a new [[RFC5888](#)] grouping attribute CaptureId and a new CLUE MC attribute RTP-id.

The above example will have the following SDP

Even

Expires November 28, 2012

[Page 5]

b=AS:10000

a=group:captureId 1 2 3 4 5 6

m=video 49170 RTP/AVP 96

a=rtpmap:96 H264/90000

a=fmtp:96 profile-level-id=42A01E; //Baseline profile, Level 3.0

mid=1

b=TIAS:20000000

m=video 49172 RTP/AVP 96

a=rtpmap:96 H264/90000

a=fmtp:96 profile-level-id=42A01E; //Baseline profile, Level 3.0

mid=2

b=TIAS:20000000

m=video 49174 RTP/AVP 96

a=rtpmap:96 H264/90000

a=fmtp:96 profile-level-id=42A01E; //Baseline profile, Level 3.0

mid=3

b=TIAS:20000000

m=video 49176 RTP/AVP 96

a=rtpmap:96 H264/90000

a=fmtp:96 profile-level-id=42A01E; //Baseline profile, Level 3.0

mid=4

b=TIAS:20000000

Even

Expires November 28, 2012

[Page 6]

```
m=video 49178 RTP/AVP 96

a=rtpmap:96 H264/900000

a=fmtp:96 profile-level-id=42A01E; //Baseline profile, Level
3.0

mid=5

b=TIAS:20000000

m=video 49180 RTP/AVP 96

a=rtpmap:96 H264/900000

a=fmtp:96 profile-level-id=42A01E; //Baseline profile, Level
3.0

mid=6

b=TIAS:20000000
```

There is a need for a new MC attribute RTPid which will have the mid of the related RTP stream

- o VC0- (the camera-left camera stream, purpose=main, switched:no, RTPid=1
- o VC1- (the center camera stream, purpose=main, switched:no, RTPid=2
- o VC2- (the camera-right camera stream), purpose=main, switched:no, RTPid=3
- o VC3- (the loudest panel stream), purpose=main, switched:yes, RTPid=4
- o VC4- (the loudest panel stream with PiPs), purpose=main, composed=true; switched:yes, RTPid=5
- o VC5- (the zoomed out view of all people in the room), purpose=main, composed=no; switched:no, RTPid=2
- o VC6- (presentation stream), purpose=presentation, switched:no, RTPid=6

There was a concern about the size of SDP or CLUE messages if the MCU will advertise the roster information. This is not the current approach defined in IETF work. The assumption is that the MCU

Even

Expires November 28, 2012

[Page 7]

advertises a set of virtual MCs and provide the content of the MC based on the multipoint application logic. The roster information is provided by other means like XCON conference event package enabling the conference participants to use this information to select a new source from a new participant. Still the assumption is that the MCU acting will create the RTP stream using his own SSRC.

Note that for the video switch MCU the SDP SSRC attribute can be used to provide the information about the different sources, an example is in [\[RFC6184\] section 8.3](#), this solution does not scale when there are many participants since the SDP may grow big and it will be better in this case to use the XCON conference event package that supports partial updates.

The offer is for H.264 [\[RFC6184\]](#) profile level 3. The level ID specifies the maximum encoding and decoding capabilities of the H.264 codec like the max macroblocks process rate, max frame size. The max-mbps, max-smbps, max-fs, max-cpb, max-dpb, and max-br parameters are used to signal the receiver implementation. The level ID is used to convey also the maximum encoding capability. [\[RFC6184\]](#) does not provide the means to override specific level parameters for the encoder side; it only allows the decoder to ask for a specific configuration.

[RFC6236] a generic session setup attribute that make it possible to negotiate the image attributes like frame size and aspect ratio.

The bandwidth parameters can be used to specify the maximum session bandwidth as well as the maximum bandwidth for individual streams.

[4.](#) Acknowledgements

place holder

[5.](#) IANA Considerations

TBD

[6.](#) Security Considerations

TBD.

[7.](#) References

Even

Expires November 28, 2012

[Page 8]

7.1. Normative References

- [I-D.ietf-clue-framework]
Romanow, A., Duckworth, M., Pepperell, A., and B. Baldino,
"Framework for Telepresence Multi-Streams",
[draft-ietf-clue-framework-05](#) (work in progress),
February 2012.
- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.

7.2. 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.
- [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.
- [RFC5117] Westerlund, M. and S. Wenger, "RTP Topologies", [RFC 5117](#),
January 2008.
- [RFC5888] Camarillo, G. and H. Schulzrinne, "The Session Description
Protocol (SDP) Grouping Framework", [RFC 5888](#), June 2010.
- [RFC6184] Wang, Y., Even, R., Kristensen, T., and R. Jesup, "RTP
Payload Format for H.264 Video", [RFC 6184](#), May 2011.
- [RFC6236] Johansson, I. and K. Jung, "Negotiation of Generic Image
Attributes in the Session Description Protocol (SDP)",
[RFC 6236](#), May 2011.

Even

Expires November 28, 2012

[Page 9]

Author's Address

Roni Even
Huawei Technologies
Tel Aviv,
Israel

Email: even.roni@huawei.com