CLUE WG
Internet-Draft
Intended status: Standards Track
Expires: February 28, 2017
R. Even
Huawei Technologies
J. Lennox
Vidyo
August 27, 2016

Mapping RTP streams to CLUE Media Captures
draft-ietf-clue-rtp-mapping-08.txt

Abstract

This document describes how the Real Time transport Protocol (RTP) is used in the context of the CLUE protocol. It also describes the mechanisms and recommended practice for mapping RTP media streams defined in SDP to CLUE Media Captures.

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

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

SIP offer answer [RFC3264] uses SDP [RFC4566] to describe the RTP[RFC3550] media streams. Each RTP stream has a unique SSRC within its RTP session. The content of the RTP stream is created by an encoder in the endpoint. This may be an original content from a camera or a content created by an intermediary device like an MCU (Multipoint Control Unit).

This document makes recommendations, for the CLUE architecture, about how RTP and RTCP streams should be encoded and transmitted, and how their relation to CLUE Media Captures should be communicated. The proposed solution supports multiple RTP topologies [RFC7667].

With regards to the media (audio, video and timed text), systems that support CLUE use RTP for the media, SDP for codec and media transport negotiation (CLUE individual encodings) and the CLUE protocol for Media Capture description and selection. In order to associate the media in the different protocols there are three mapping that need to be specified:

1. CLUE individual encodings to SDP

2. RTP streams to SDP (this is not a CLUE specific mapping)

3. RTP streams to MC to map the received RTP steam to the current MC in the MCC.

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.

The definitions from the CLUE framework document [I-D.ietf-clue-framework] section 3 are used by this document as well.

3. RTP topologies for CLUE

The typical RTP topologies used by CLUE Telepresence systems specify different behaviors for RTP and RTCP distribution. A number of RTP topologies are described in [RFC7667]. For telepresence, the relevant topologies include Point-to-Point, as well as Media-Mixing mixers, Media- Switching mixers, and Selective Forwarding Middleboxs.

In the Point-to-Point topology, one peer communicates directly with a single peer over unicast. There can be one or more RTP sessions, each sent on a separate 5-tuple, and having a separate SSRC space, with each RTP session carrying multiple RTP streams identified by their SSRC. All SSRCs are recognized by the peers based on the information in the RTCP SDES report that includes the CNAME and SSRC of the sent RTP streams. There are different Point-to-Point use cases as specified in CLUE use case [RFC7205]. In some cases, a CLUE session which, at a high-level, is point-to-point may nonetheless have an RTP stream which is best described by one of the mixer topologies. For example, a CLUE endpoint can produce composite or switched captures for use by a receiving system with fewer displays than the sender has cameras. The Media Capture may be described using MCC.

For the Media Mixer topology [RFC7667], the peers communicate only with the mixer. The mixer provides mixed or composited media streams, using its own SSRC for the sent streams. The conference roster information including conference participants, endpoints, media and media-id (SSRC) can be determined using the conference event package [RFC4575] element.

In the Media-Switching Mixer topology [RFC7667], the peer to mixer communication is unicast with mixer RTCP feedback. It is conceptually similar to a compositing mixer as described in the previous paragraph, except that rather than compositing or mixing multiple sources, the mixer provides one or more conceptual sources selecting one source at a time from the original sources. The Mixer creates a conference-wide RTP session by sharing remote SSRC values as CSRCs to all conference participants, and forwarding RTCP reports.

In the Selective Forwarding Middlebox (SFM) [RFC7667] topology, the peer to middlebox communication is unicast with RTCP feedback. Every potential sender in the conference has a source which may be "projected" by the SFM into every other RTP session in the conference; thus, even though the SFM establishes a separate RTP session with each endpoint, every original source is maintained with an independent SSRC to every receiver, maintaining separate decoding state and its original RTCP SDES information.

4. Mapping CLUE Capture Encodings to RTP streams

The different topologies described in Section 3 create different SSRC distribution models and RTP stream multiplexing points.

Most video conferencing systems today can separate multiple RTP sources by placing them into RTP sessions using, the SDP description. For example, main and slides video sources are separated into separate RTP sessions based on the content attribute [RFC4796]. This solution is straightforward if the multiplexing point is at the UDP transport level, where each RTP stream uses a separate RTP session. This will also be true for mapping the RTP streams to Media Captures Encodings if each Media Capture Encodings uses a separate RTP session, and the consumer can identify it based on the receiving RTP port. In this case, SDP only needs to label the RTP session with an identifier that can be used to identify the Media Capture in the CLUE description. The SDP label attribute serves as this identifier. In this case, the mapping does not change even if the RTP session is switched using same or different SSRC.

Even though Session multiplexing is supported by CLUE, for scaling reasons, CLUE indicates that SSRC multiplexing in a single or multiple sessions using [I-D.ietf-mmusic-sdp-bundle-negotiation]may be used. When SSRC multiplexing is used, the mapping of RTP streams to Captures Encodings needs to be considered.

MCCs bring another mapping issue, in that an MCC represents multiple Media Captures that can be sent as part of this MCC if configured by the consumer. When receiving an RTP stream which is mapped to the MCC, the consumer needs to know which original MC it is in order to get the MC parameters from the advertisement. If a consumer requested a MCC, the original MC does not have a capture encoding, so it cannot be associated with an m-line using a label as described in CLUE signaling [I-D.ietf-clue-signaling]. This is important, for example, to get correct scaling information for the original MC, which may be different for the various MCs that are contributing to the MCC.

4.1. Review of RTP related documents relevant to CLUE work.

This section provides an overview of the RFCs and drafts that can be used in a CLUE system and as a base for a mapping solution. This section is for information only; the normative behavior is given in the cited documents. Tools for SSRC multiplexing support are defined for general conferencing applications; CLUE systems use the same tools.

When looking at the available tools based on current work in MMUSIC, AVTcore and AVText Working Groups for supporting SSRC multiplexing the following documents are considered to be relevant.

Negotiating Media Multiplexing Using the Session Description Protocol in [I-D.ietf-mmusic-sdp-bundle-negotiation] defines a "bundle" SDP grouping extension that can be used with SDP Offer/Answer mechanism to negotiate the usage of a single 5-tuple for sending and receiving media associated with multiple SDP media descriptions ("m="). [I-D.ietf-mmusic-sdp-bundle-negotiation] specifies how to associate a received RTP stream with the m-line describing it. The assumption in Bundle is that each SDP m-line represents a single media source. [I-D.ietf-mmusic-sdp-bundle-negotiation] specifies using the SDP mid value and sending it as RTCP SDES and an RTP header extension in order to be able to map the RTP stream to the SDP m-line. This is relevant when there are multiple RTP streams with the same payload subtype number.

SDP Source attribute [RFC5576] provides mechanisms to describe specific attributes of RTP sources based on their SSRC.

Negotiation of generic image attributes in SDP [RFC6236] provides the means to negotiate the image size. The image attribute can be used to offer different image parameters like size. Offering multiple RTP streams with different resolutions is done using separate RTP session for each image option. ([I-D.ietf-mmusic-sdp-bundle-negotiation] provides the support of a single RTP session but each image option will need a separate SDP m-line).

The recommended support of the simulcast case is to use [I-D.ietf-mmusic-sdp-simulcast].

4.2. Recommendations

The recommendation is that CLUE endpoints using SSRC multiplexing MUST support [I-D.ietf-mmusic-sdp-bundle-negotiation].

5. CaptureID definition

For MCC which can represent multiple switched MCs there is a need to know which MC represents the current RTP stream. This requires a mapping from an RTP stream to an MC. In order to address this mapping this document defines an RTP header extension that includes the CaptureID in order to map to the original MC allowing the consumer to use the original source MC attributes like the spatial information. The media provider MUST send for MCC Capture Encoding the captureID of the current MC in the RTP header and as a RTCP SDES message.

5.1. RTCP CaptureId SDES Item

This document specifies a new RTCP SDES message

```
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
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   | CaptureId = XX |     length    |CaptureId
   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
   |   ....
```

This CaptureID is the same as in the CLUE MC and is also used in the RTP header extension.

This SDES message MAY be sent in a compound RTCP packet based on the application need.

5.2. RTP Header Extension

The CaptureId is carried within the RTP header extension field, using [RFC5285] two bytes header extension.

Support is negotiated within the SDP, i.e.

a=extmap:1 urn:ietf:params:rtp-hdrext:CaptureId

Packets tagged by the sender with the CaptureId then contain a header extension as shown below

```
 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
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      |  ID           |   Len-1       |    CaptureId
      +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
      | ....             |
      +-+-+-+-+-+-+-+-+-+
```

There is no need to send the CaptureId header extension with all RTP packets. Senders MAY choose to send it only when a new MC is sent. If such a mode is being used, the header extension SHOULD be sent in the first few RTP packets to reduce the risk of losing it due to packet loss.

6. Examples

In this partial advertisement the Media Provider advertises a composed capture VC7 made by a big picture representing the current speaker (VC3) and two picture-in-picture boxes representing the previous speakers (the previous one -VC5- and the oldest one -VC6).

```
<ns2:mediaCapture xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance"
      xsi:type="ns2:videoCaptureType" captureID="VC7" mediaType="video">
          <ns2:captureSceneIDREF>CS1</ns2:captureSceneIDREF>
          <ns2:nonSpatiallyDefinable>true</ns2:nonSpatiallyDefinable>
          <ns2:content>
                <ns2:captureIDREF>VC3</ns2:captureIDREF>
                <ns2:captureIDREF>VC5</ns2:captureIDREF>
                <ns2:captureIDREF>VC6</ns2:captureIDREF>
          </ns2:content>
                  <ns2:maxCaptures>3</ns2:maxCaptures>
            <ns2:allowSubsetChoice>false</ns2:allowSubsetChoice>
          <ns2:description lang="en">big picture of the current speaker
            pips about previous speakers</ns2:description>
            <ns2:priority>1</ns2:priority>
            <ns2:lang>it</ns2:lang>
            <ns2:mobility>static</ns2:mobility>
            <ns2:view>individual</ns2:view>
        </ns2:mediaCapture>
```

In this case the media provider will send capture IDs VC3, VC5 or VC6 as an RTP header extension and RTCP SDES message for the RTP stream associated with the MC.

7. Acknowledgements

The authors would like to thanks Allyn Romanow and Paul Witty for contributing text to this work.

8. IANA Considerations

This document defines a new extension URI in the RTP Compact Header Extensions subregistry of the Real-Time Transport Protocol (RTP) Parameters registry, according to the following data:

Extension URI: urn:ietf:params:rtp-hdrext:CaptureId

Description: CLUE CaptureId

Contact: roni.even@mail01.huawei.com

Reference: RFC XXXX

The IANA is requested to register one new RTCP SDES items in the "RTCP SDES Item Types" registry, as follows:

Value	Abbrev	Name	Reference
TBA	CCID	CLUE CaptureId	[RFCXXXX]

9. Security Considerations

The security considerations of the RTP specification, the RTP/SAVPF profile, and the various RTP/RTCP extensions and RTP payload formats that form the complete protocol suite described in this memo apply. It is not believed there are any new security considerations resulting from the combination of these various protocol extensions.

The Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback [RFC5124] (RTP/SAVPF) provides handling of fundamental issues by offering confidentiality, integrity and partial source authentication. A mandatory to support media security solution is created by combining this secured RTP profile and DTLS-SRTP keying [RFC5764]

RTCP packets convey a Canonical Name (CNAME) identifier that is used to associate RTP packet streams that need to be synchronised across related RTP sessions. Inappropriate choice of CNAME values can be a privacy concern, since long-term persistent CNAME identifiers can be used to track users across multiple calls. This memo mandates generation of short-term persistent RTCP CNAMES, as specified in RFC7022 [RFC7022], resulting in untraceable CNAME values that alleviate this risk.

Some potential denial of service attacks exist if the RTCP reporting interval is configured to an inappropriate value. This could be done by configuring the RTCP bandwidth fraction to an excessively large or small value using the SDP "b=RR:" or "b=RS:" lines [RFC3556], or some similar mechanism, or by choosing an excessively large or small value for the RTP/AVPF minimal receiver report interval (if using SDP, this is the "a=rtcp-fb:... trr-int" parameter) [RFC4585] The risks are as follows:

1. the RTCP bandwidth could be configured to make the regular reporting interval so large that effective congestion control cannot be maintained, potentially leading to denial of service due to congestion caused by the media traffic;

2. the RTCP interval could be configured to a very small value, causing endpoints to generate high rate RTCP traffic, potentially leading to denial of service due to the non-congestion controlled RTCP traffic; and

3. RTCP parameters could be configured differently for each endpoint, with some of the endpoints using a large reporting interval and some using a smaller interval, leading to denial of service due to premature participant timeouts due to mismatched timeout periods which are based on the reporting interval (this is a particular concern if endpoints use a small but non-zero value for the RTP/AVPF minimal receiver report interval (trr-int) [RFC4585], as discussed in [I-D.ietf-avtcore-rtp-multi-stream]).

Premature participant timeout can be avoided by using the fixed (non-reduced) minimum interval when calculating the participant timeout ([I-D.ietf-avtcore-rtp-multi-stream]). To address the other concerns, endpoints SHOULD ignore parameters that configure the RTCP reporting interval to be significantly longer than the default five second interval specified in [RFC3550] (unless the media data rate is so low that the longer reporting interval roughly corresponds to 5% of the media data rate), or that configure the RTCP reporting interval small enough that the RTCP bandwidth would exceed the media bandwidth.

The guidelines in [RFC6562] apply when using variable bit rate (VBR) audio codecs such as Opus. The use of the encryption of the header extensions are RECOMMENDED, unless there are known reasons, like RTP middleboxes performing voice activity based source selection or third party monitoring that will greatly benefit from the information, and this has been expressed using API or signalling. If further evidence are produced to show that information leakage is significant from audio level indications, then use of encryption needs to be mandated at that time.

In multi-party communication scenarios using RTP Middleboxes, a lot of trust is placed on these middleboxes to preserve the sessions security. The middlebox needs to maintain the confidentiality, integrity and perform source authentication. The middlebox can perform checks that prevents any endpoint participating in a conference to impersonate another. Some additional security considerations regarding multi-party topologies can be found in [RFC7667]

10. References

10.1. Normative References

[I-D.ietf-clue-framework]
Duckworth, M., Pepperell, A., and S. Wenger, "Framework for Telepresence Multi-Streams", draft-ietf-clue-framework-25 (work in progress), January 2016.

[I-D.ietf-mmusic-sdp-bundle-negotiation]
Holmberg, C., Alvestrand, H., and C. Jennings, "Negotiating Media Multiplexing Using the Session Description Protocol (SDP)", draft-ietf-mmusic-sdp-bundle-negotiation-32 (work in progress), August 2016.

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

10.2. Informative References

[I-D.ietf-avtcore-rtp-multi-stream]
Lennox, J., Westerlund, M., Wu, W., and C. Perkins, "Sending Multiple Media Streams in a Single RTP Session", draft-ietf-avtcore-rtp-multi-stream-11 (work in progress), December 2015.

[I-D.ietf-clue-signaling]
Kyzivat, P., Xiao, L., Groves, C., and R. Hansen, "CLUE Signaling", draft-ietf-clue-signaling-09 (work in progress), March 2016.

[I-D.ietf-mmusic-sdp-simulcast]
Westerlund, M., Nandakumar, S., and M. Zanaty, "Using Simulcast in SDP and RTP Sessions", draft-ietf-mmusic-sdp-simulcast-05 (work in progress), June 2016.

[RFC3264] Rosenberg, J. and H. Schulzrinne, "An Offer/Answer Model with Session Description Protocol (SDP)", RFC 3264, DOI 10.17487/RFC3264, June 2002, <http://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, <http://www.rfc-editor.org/info/rfc3550>.

[RFC3556] Casner, S., "Session Description Protocol (SDP) Bandwidth Modifiers for RTP Control Protocol (RTCP) Bandwidth", RFC 3556, DOI 10.17487/RFC3556, July 2003, <http://www.rfc-editor.org/info/rfc3556>.

[RFC4566] Handley, M., Jacobson, V., and C. Perkins, "SDP: Session Description Protocol", RFC 4566, DOI 10.17487/RFC4566, July 2006, <http://www.rfc-editor.org/info/rfc4566>.

[RFC4575] Rosenberg, J., Schulzrinne, H., and O. Levin, Ed., "A Session Initiation Protocol (SIP) Event Package for Conference State", RFC 4575, DOI 10.17487/RFC4575, August 2006, <http://www.rfc-editor.org/info/rfc4575>.

[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, <http://www.rfc-editor.org/info/rfc4585>.

[RFC4796] Hautakorpi, J. and G. Camarillo, "The Session Description Protocol (SDP) Content Attribute", RFC 4796, DOI 10.17487/RFC4796, February 2007, <http://www.rfc-editor.org/info/rfc4796>.

[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, <http://www.rfc-editor.org/info/rfc5104>.

[RFC5124] Ott, J. and E. Carrara, "Extended Secure RTP Profile for Real-time Transport Control Protocol (RTCP)-Based Feedback (RTP/SAVPF)", RFC 5124, DOI 10.17487/RFC5124, February 2008, <http://www.rfc-editor.org/info/rfc5124>.

[RFC5285] Singer, D. and H. Desineni, "A General Mechanism for RTP Header Extensions", RFC 5285, DOI 10.17487/RFC5285, July 2008, <http://www.rfc-editor.org/info/rfc5285>.

[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, <http://www.rfc-editor.org/info/rfc5576>.

[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, <http://www.rfc-editor.org/info/rfc5764>.

[RFC6236] Johansson, I. and K. Jung, "Negotiation of Generic Image Attributes in the Session Description Protocol (SDP)", RFC 6236, DOI 10.17487/RFC6236, May 2011, <http://www.rfc-editor.org/info/rfc6236>.

[RFC6562] Perkins, C. and JM. Valin, "Guidelines for the Use of Variable Bit Rate Audio with Secure RTP", RFC 6562, DOI 10.17487/RFC6562, March 2012, <http://www.rfc-editor.org/info/rfc6562>.

[RFC7022] Begen, A., Perkins, C., Wing, D., and E. Rescorla, "Guidelines for Choosing RTP Control Protocol (RTCP) Canonical Names (CNAMEs)", RFC 7022, DOI 10.17487/RFC7022, September 2013, <http://www.rfc-editor.org/info/rfc7022>.

[RFC7205] Romanow, A., Botzko, S., Duckworth, M., and R. Even, Ed., "Use Cases for Telepresence Multistreams", RFC 7205, DOI 10.17487/RFC7205, April 2014, <http://www.rfc-editor.org/info/rfc7205>.

[RFC7667] Westerlund, M. and S. Wenger, "RTP Topologies", RFC 7667, DOI 10.17487/RFC7667, November 2015, <http://www.rfc-editor.org/info/rfc7667>.

Authors' Addresses

Roni Even
Huawei Technologies
Tel Aviv
Israel

Email: roni.even@mail01.huawei.com

Jonathan Lennox
Vidyo, Inc.
433 Hackensack Avenue
Seventh Floor
Hackensack, NJ 07601
US

Email: jonathan@vidyo.com