CLUE WG
Internet-Draft
Intended status: Standards Track
Expires: July 18, 2017

R. Even
Huawei Technologies
J. Lennox
Vidyo
January 14, 2017

Mapping RTP streams to CLUE Media Captures
draft-ietf-clue-rtp-mapping-11.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 Session Description Protocol (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 July 18, 2017.

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 RTP processing in compliant CLUE 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 CLUE 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. If needed by CLUE endpoint, the conference roster information including conference participants, endpoints, media and media-id (SSRC) can be determined using the conference event package [RFC4575] element.

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; the video conferencing application can also have some knowledge about the purpose of each RTP session. For example, video conferencing applications that have main and slides video sources can send each media source in a separate RTP session with a content attribute [RFC4796] enabling different application behavior for each received RTP media source. Demultiplexing is straightforward because each media capture is sent as a single RTP stream, with each RTP stream being sent in a separate RTP session, on a distinct UDP 5-tuple. 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.

Each Capture Encoding MUST be sent as a separate RTP stream. CLUE endpoints MUST support sending each such RTP stream in a separate RTP session signalled by an SDP m= line. They MAY also support sending some or all of the RTP streams in a single RTP session, using the mechanism described in [I-D.ietf-mmusic-sdp-bundle-negotiation] to relate RTP streams to SDP m= lines.

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.

5. MCC Constituent CaptureID definition

For a MCC which can represent multiple switched MCs there is a need to know which MC is represented in the current RTP stream at any given time. This requires a mapping from the SSRC of the RTP stream conveying a particular MCC to the constituent MC. In order to address this mapping this document defines an RTP header extension

and SDES item that includes the captureID of the original MC, allowing the consumer to use the original source MC's attributes like the spatial information.

This mapping temporarily associates the SSRC of the RTP stream conveying a particular MCC with the captureID of the single original MC that is currently switched into the MCC. This mapping cannot be used for the composed case where more than one original MC is composed into the MCC simultaneously.

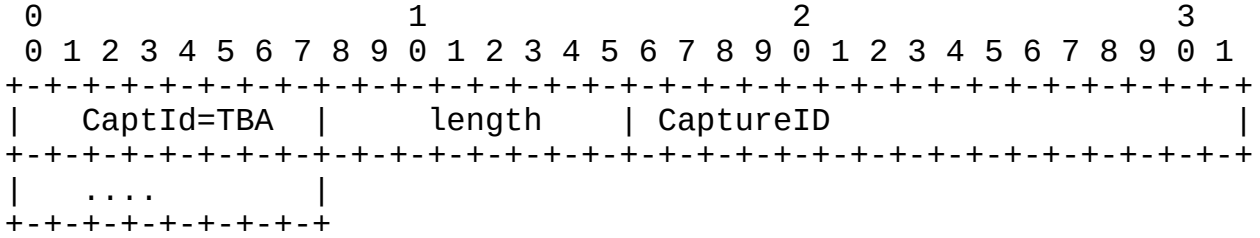
If there is only one MC in the MCC then the media provider MUST send the captureID of the current constituent MC in the RTP Header Extension and as a RTCP CaptureID SDES item. When the media provider switches the MC it sends within an MCC, it MUST send the captureID value for the MC just switched into the MCC.

If there is more than one MC composed into the MCC then the media provider MUST NOT send any of the MCs' captureIDs using this mechanism. However, if an MCC is sending contributing source (CSRC) information in the RTP header for a composed capture, it MAY send the captureID values in the RTCP SDES packets giving source information for the SSRC values sent as contributing sources (CSRCs).

If the media provider sends the captureID of a single MC switched into an MCC, then later sends a composed stream of multiple MCs in the same MCC, it MUST send the special value "-", a single dash character, as the captureID RTP Header Extension and RTCP CaptureID SDES item. The single dash character indicates there is no applicable value for the MCC constituent CaptureID. The media consumer interprets this as meaning that any previous CaptureID value associated with this SSRC no longer applies. As [I-D.ietf-clue-data-model-schema] defines the captureID syntax as "xs:ID", the single dash character is not a legal captureID value, so there is no possibility of confusing it with an actual captureID.

5.1. RTCP CaptureID SDES Item

This document specifies a new RTCP SDES item.



Note to the RFC Editor: Please replace TBA with the value assigned by IANA.

This CaptureIDis a variable-length UTF-8 string corresponding either to a CaptureID negotiated in the CLUE protocol, or the single character "-".

This SDES item MUST be sent in an SDES packet within a compound RTCP packet unless support for Reduced-size RTCP has been negotiated as specified in RFC 5506 [RFC5506], in which case it can be sent as an SDES packet in a non-compound RTCP packet.

5.2. RTP Header Extension

The CaptureIDis also carried in an RTP header extension [RFC5285], using the mechanism defined in [RFC7941].

Support is negotiated within SDP using the URN "urn:ietf:params:rtp-hdrext:sdes:CaptureID".

The CaptureID is sent in a RTP Header Extension because for switched captures, receivers need to know which original MC corresponds to the media being sent for an MCC, in order to correctly apply geometric adjustments to the received media.

As discussed in [RFC7941], there is no need to send the CaptId 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. See [RFC7941] for more discussion of this.

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 SDES 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:sdes:CaptId

Description: CLUE CaptId

Contact: ron.even.tlv@gmail.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 CaptId	[RFCXXXX]

Note to the RFC Editor: Please replace RFCXXXX with this RFC number.

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. CLUE endpoints MUST support RTP/SAVPF 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. CLUE endpoint MUST generate 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; this middleboxes are trusted to preserve the sessions' security. The middlebox SHOULD maintain the confidentiality, integrity and perform source authentication. The middlebox MAY 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-data-model-schema]
Presta, R. and S. Romano, "An XML Schema for the CLUE data model", draft-ietf-clue-data-model-schema-17 (work in progress), August 2016.

[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-36 (work in progress), October 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>.

[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, <http://www.rfc-editor.org/info/rfc7941>.

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-10 (work in progress), January 2017.

[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>.

[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>.

[RFC5506] Johansson, I. and M. Westerlund, "Support for Reduced-Size Real-Time Transport Control Protocol (RTCP): Opportunities and Consequences", RFC 5506, DOI 10.17487/RFC5506, April 2009, <http://www.rfc-editor.org/info/rfc5506>.

[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>.

[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@huawei.com

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

Email: jonathan@vidyo.com