Network Working Group E. Berger
Internet-Draft S. Nandakumar
Intended status: Standards Track Cisco Systems
Expires: April 24, 2014 October 21, 2013

Frame marking for RTP packets
draft-avtext-berger-framemarking-00

Abstract

This document describes a mechanisms to provide frame markings to allow RTP switches to perform stream operations on encrypted payload. The mechanisms support extensions to allow for codec specific information.

Table of Contents

1. Introduction

It is common practice in modern voice and video conferencing systems to implement a centralized component that acts as a RTP switch. It receives voice and video streams from each participant, which may be encrypted using SRTP [RFC3711]. The goal is to provide a set of streams back to the participants which enable them to render the right media content. In a simple video configuration, for example, the goal will be that each participant sees and hears just the active speaker. In that case, the goal of the switch is to receive the voice and video streams from each participant, determine the active speaker based on energy in the voice packets, and select the corresponding video stream for transmission to participants, see Figure 1

In this document, an "RTP switch" is used as a common short term for the terms switching RTP mixer", "source projecting middlebox", and "video switching MCU" as discussed in [I-D.ietf-avtcore-rtp-topologies-update].

```
         +---+      +------------+      +---+
         | A |<---->|            |<---->| B |
         +---+      |            |      +---+
                    |   RTP      |
         +---+      |  Switch    |      +---+
         | C |<---->|            |<---->| D |
         +---+      +------------+      +---+
```

Figure 1: RTP switch

In order to properly support switching of video streams, the RTP switch typically needs information in order to do a proper job:

- o Because of inter-frame dependencies, it should ideally switch video streams at a point where the first frame from the new speaker can be decoded by recipients without prior frames, e.g switch on an intra-frame.
- o In many cases, the switch may need to drop frames in order to realize congestion control techniques, and needs to know which frames can be dropped with minimal impact to video quality
- o Furthermore, it is highly desirable to do this in a way which is not specific to the video codec. Nearly all modern video codecs share common concepts around I, P, B frames.
- o It is also desirable to be able to do this for SRTP without requiring the video switch to decrypt the packets. SRTP will encrypt the RTP payload format contents and consequently this data is not usable for the switching function.

By providing meta-information about the RTP streams outside the encrypted media payload a RTP switch can do selective forwarding without decrypting the payload. This document provides a solution to this problem.

2. Solution

The solution uses RTP header extensions as defined in [RFC5285]. A subset of meta-information from the video stream is provided as an header extension to allow a RTP switch to do generic video switching handling of video streams encoded with different video codecs.

The following information are extracted from the media payload.

- o Discardable - The flag must be true for packets that can be dropped, and still provide a decodable media stream.
- o Switching point (1 bit) - The flag must be true for RTP packets in a frame that can be used as a switcing point. A switching point is the first packet where a new receiver can start decoding a video stream without prior frames, e.g an IDR frame from [RFC6184].
- o Temporal ID (3 bits) - The base temporal quality starts with 0, and increases with 1 for each temporal layer.
- o frame Type (3 bits) - Abstract frame type; P-frame=0, IDR=1, GDR=2. The abstracted frame types are:
  - * P-frame - a frame depending on a previous frame
  - * IDR - a frame without references to other frames
  - * GDR - a Gradual Decoder Refresh (GDR) packet includes both p-frame and frame information to allow a receiver to build of a IDR frame over a short period.

Video codec specific information can be provided as an extension.

2.1. RTP header extension

The values of frame information can be carried as RTP header extensions encoded using the one-byte header as described in [RFC5285]. Only the one-byte header version is listed with examples in the document.

```
0                   1
 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  ID=2 |  L=0  |D|S|TID  |Type |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

The frame marking can be extended with codec specific information

using a longer length value in the one-byte header. The codec specific information included in the header extension MUST match the SDP negotiated payload format for the RTP stream.

```
  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 2
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  ID=2 |  L=2 |D|S|TID  |Type | video codec specific information |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

In the following example the H265 LayerID is included as video codec specific information. The length field is 1 to add another 1 byte of data, the H265 LayerId is a 6-bit field and a 2-bit PADding at the end.

```
  0                   1                   2
 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
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  ID=2 |  L=1  |D|S|TID  |Type | H265-LayerId|PAD|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

2.2. Signaling information

The URI for declaring this header extension in an extmap attribute is "urn:ietf:params:rtp-hdrext:framemarkinginfo". It does not contain any extension attributes.

An example attribute line in SDP:

```
a=extmap:3 urn:ietf:params:rtp-hdrext:framemarkinginfo
```

2.3. Considerations on use

The header extension values MUST represent what is already in the RTP payload.

When a RTP switch needs to discard a received video frame due to congestion control considerations, it is RECOMMENDED that it preferably drop frames marked with the "discardable" bit.

When a RTP switch want to forward a new video stream to a receiver, its RECOMMENDED to forward the new video stream from the first switching point and forward. A RTP switch can request a media source to generate a switching point for H264 by sending Full Intra Request (RTCP FIR) as defined in [RFC5104].

3. Security Considerations

In the Secure Real-Time Transport Protocol (SRTP) [RFC3711], RTP header extensions are authenticated but not encrypted. When header extensions are used some of the payload type information are exposed and is visible to middle boxes. The encrypted media data is not exposed, so this is not seen as a high risk exposure.

4. IANA Considerations

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

Extension URI: urn:ietf:params:rtp-hdrext:framemarkinginfo
Description: Frame marking information for video streams
Contact: espeberg@cisco.com
Reference: RFC XXXX

Note to RFC Editor: please replace RFC XXXX with the number of this RFC.

5. References

5.1. Normative References

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

5.2. Informative References

[I-D.ietf-avtcore-rtp-topologies-update]
Westerlund, M. and S. Wenger, "RTP Topologies", draft-ietf-avtcore-rtp-topologies-update (work in progress), April 2013.

[RFC3550] Schulzrinne, H., Casner, S., Frederick, R., and V. Jacobson, "RTP: A Transport Protocol for Real-Time Applications", STD 64, RFC 3550, July 2003.

[RFC3711] Baugher, M., McGrew, D., Naslund, M., Carrara, E., and K. Norrman, "The Secure Real-time Transport Protocol (SRTP)", RFC 3711, March 2004.

[RFC5104] Wenger, S., Chandra, U., Westerlund, M., and B. Burman, "Codec Control Messages in the RTP Audio-Visual Profile with Feedback (AVPF)", RFC 5104, February 2008.

[RFC5285] Singer, D. and H. Desineni, "A General Mechanism for RTP Header Extensions", RFC 5285, July 2008.

[RFC6184] Wang, Y., Even, R., Kristensen, T., and R. Jesup, "RTP Payload Format for H.264 Video", RFC 6184, May 2011.

Authors' Addresses

Espen Berger
Cisco Systems

Phone: +47 98228179
Email: espeberg@cisco.com

Suhas Nandakumar
Cisco Systems
170 West Tasman Drive
San Jose, CA 95134
US

Email: snandaku@cisco.com