

Video Frame Info RTP Header Extension

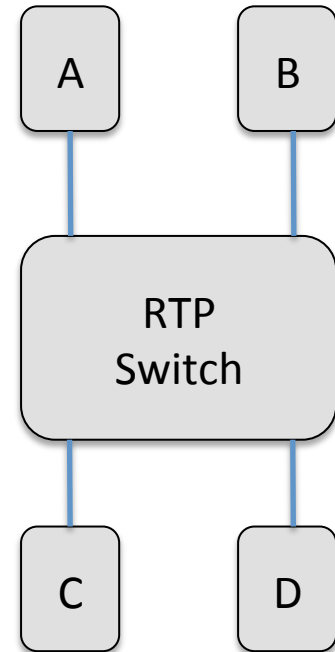
draft-berger-avtext-framemarking-00

Espen Berger, Suhas Nandakumar, Mo Zanaty
Cisco

Main Motivation

Payload-Agnostic RTP Switch

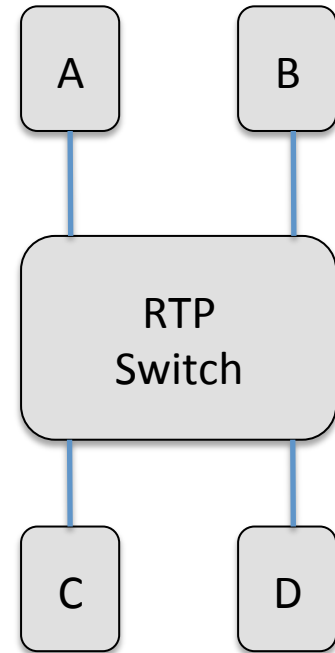
- Payload may be encrypted
 - Avoid decryption cost to improve switch scale and latency
- Payload may be encrypted end-to-end
 - Impossible to decrypt / inspect payload without end-to-end keys
- Payload may be unknown format
 - Codec-agnostic switching can support any format, old or new



More Motivations

Smarter RTP Switch

- Clean video switching at intra-frames
- Better recovery during packet loss
- Drop least important packets during congestion
- Drop scalable enhancement layers for constrained endpoints



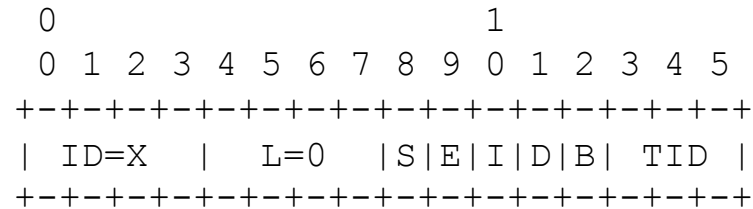
Smarter Endpoints

- Better recovery during packet loss

Proposed Solution

- Video Frame Info RTP Header Extension
- Fixed-length codec-independent part
- Optional, variable-length codec-specific part
- Uses RFC 5285 format, length determines presence and size of optional part
- Negotiated in SDP via extmap attribute:
a=extmap:X urn:ietf:params:rtp-hdext:frameinfo

Codec-Independent Extension



- S: Start of Frame - MUST be 1 in the first packet in a frame within a layer.
- E: End of Frame - MUST be 1 in the last packet in a frame within a layer.
- I: Independent Frame - MUST be 1 for frames that can be decoded independent of prior frames, e.g. key/intra-frame; otherwise MUST be 0.
- D: Discardable Frame - MUST be 1 for frames that can be dropped, and still provide a decodable media stream; otherwise MUST be 0.
- B: Base Layer Sync - MUST be 1 if this frame only depends on the base layer; otherwise MUST be 0.
- TID: Temporal ID (3 bits) - The base temporal quality starts with 0, and increases with 1 for each temporal layer/sub-layer.

Start/End of Frame/Layer

- Better recovery during packet loss
- Marker bit is unreliable for end of frame
- Layer start/end indicators are needed
- Newer payload formats are heading this way

Independent and Discardable

- Independent Frames
 - Clean video switching at intra-frames
- Discardable Frames
 - Drop least important packets during congestion

Temporal ID and Base Layer Sync

- Temporal Layer ID
 - Drop scalable enhancement layers for constrained endpoints
- Base Layer Sync
 - Clean layer switching at sync points

Codec-Specific Extension

```
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=X  |  L=N  |S|E|I|D|B| TID |  N-byte codec specific info  |
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
```

H.265:

```
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
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
| ID=X  |  L=1  |S|E|I|D|B| TID |H265LayerId|0 0|
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
```

VP8:

```
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=X  |  L=2  |S|E|I|D|B| TID |  PictureID  |0|
+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+--+
```