Video Frame Info
RTP Header Extension

draft-ietf-avtext-framemarking-05

Note Well: https://datatracker.ietf.org/ipr/2876/

Espen Berger, Suhas Nandakumar, Mo Zanaty

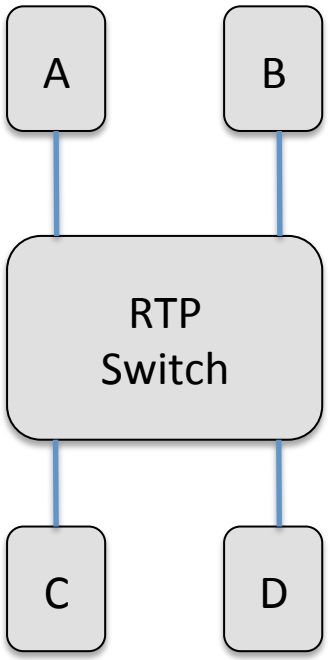
AVTCORE WG

IETF 99 – Prague, CZ – July 21, 2017

Review: Main Motivation

Payload-Agnostic RTP Switch

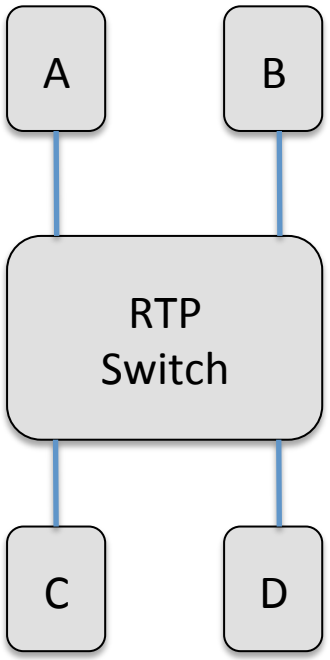
- Payload may be encrypted
  - Avoid decryption cost to improve switch scale and latency
- Payload may be encrypted end-to-end
  - Impossible to decrypt / inspect payload without end-to-end keys
- Payload may be unknown format
  - Codec-agnostic switching can support any format, old or new



Review: More Motivations

Smarter RTP Switch

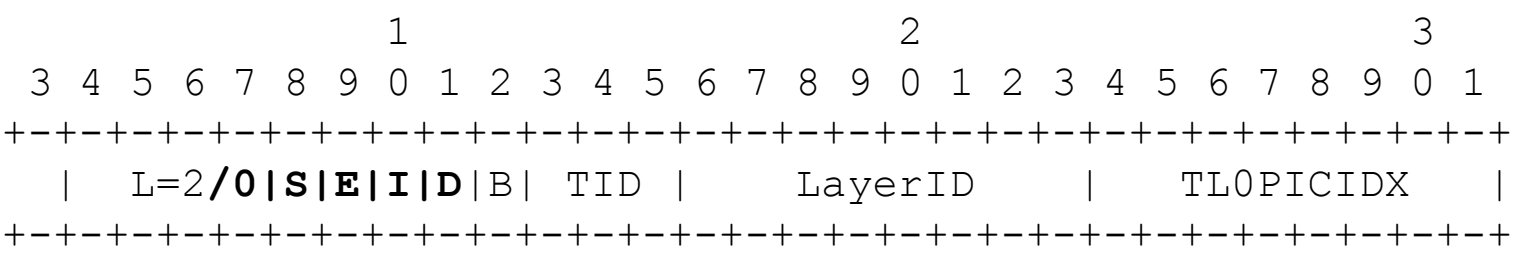
- Clean video switching at intra-frames
- Better recovery during packet loss
- Drop least important packets during congestion
- Drop scalable enhancement layers for constrained endpoints



Smarter Endpoints

- Better recovery during packet loss

Video Frame Info 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.
- **LID: Layer ID** (8 bits) - The spatial and quality layer ID defined by scalable codecs.
- **TL0PICIDX: Temporal Base Layer 0 Picture Index** (8 bits) - Running index of base temporal layer frames and dependencies on them.

Changes in version -05

3.2.1. Layer ID Mappings for Scalable Streams

- Removed VP9 LID mapping,
  moved to VP9 RTP payload draft.
- Editorial: Added references for
  VP8 [RFC7741] and H.264 [RFC6184].

3.4. Usage Considerations

- Discard highest TID/LID values first.
- NOT RECOMMENDED for complex or irregular scalability structures.

VP9 LID Mapping

- Removed VP9 LID mapping, moved to VP9 RTP payload draft.
- Added section on Future Codecs.

3.2.1.5. Future Codec LID Mapping

The RTP payload format specification for future video codecs SHOULD include a section describing the LID mapping and TID mapping for the codec. For example, the LID/TID mapping for the VP9 codec is described in the VP9 RTP Payload Format [I-D.ietf-payload-vp9].

Discard Priority

- Discard using “D” bit or highest TID/LID values.

3.4 Usage Considerations

...

When an 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 D (Discardable) bit set, or the highest values of TID and LID, which indicate the highest temporal and spatial/quality enhancement layers, since those typically have fewer dependenices on them than lower layers.

Complex Scalability Structures

- NOT RECOMMENDED for complex or irregular scalability structures.

3.4.2. Complex or Irregular Scalability Structures

The LID and TID information is most useful for simple, regular scalability structures such as hierarchical temporal or spatial/ quality layering structures. The LID and TID information is less useful, or even not useful at all, for complex, irregular scalability structures that do not conform to simple patterns of inter-layer dependencies and referencing structures. Therefore it is NOT RECOMMENDED to use LID and TID information for RTP switch forwarding decisions in the case of complex or irregular scalability structures.

Open Issues

- VP9 P/U bits vs. Frame Marking I/B bits.
  - Are the I/B bits sufficient?
  - What use cases are problematic if insufficient?
  - Do we need to alter the meaning of the B bit?

Next Steps

- Ready for WGLC after resolving open issues.
- Questions?
- Thank you!