

# Video Frame Info RTP Header Extension

draft-ietf-avtext-framemarking-06

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

Espen Berger, Suhas Nandakumar, Mo Zanaty

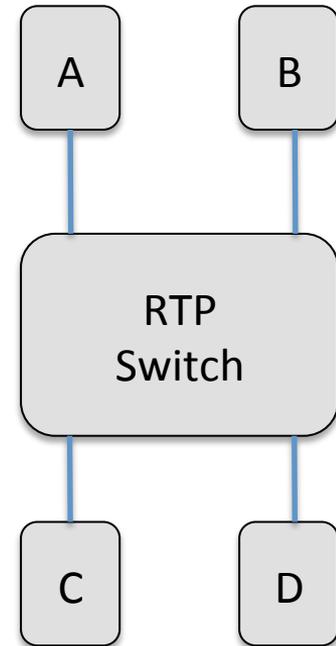
AVTCORE WG

IETF 100 – Singapore – November 17, 2017

# Review: Main Motivation

## Payload-Agnostic RTP Switch

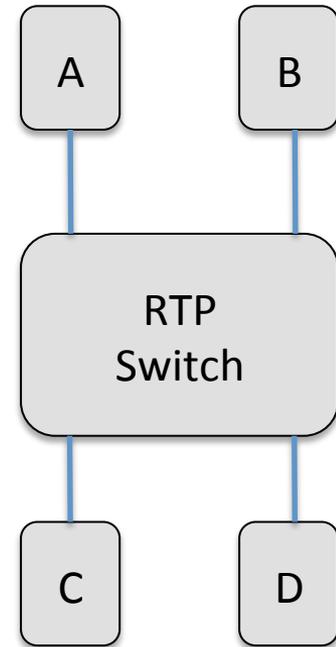
- Payload may be encrypted
  - Avoid decryption cost to improve switch scale and latency
- Payload may be encrypted end-to-end
  - Impossible to decrypt / inspect payload without end-to-end keys
- Payload may be unknown format
  - Codec-agnostic switching can support any format, old or new



# Review: More Motivations

## Smarter RTP Switch

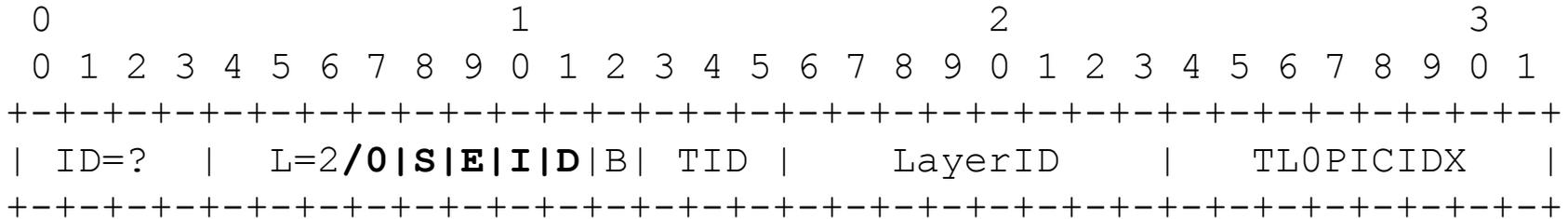
- Clean video switching at intra-frames
- Better recovery during packet loss
- Drop least important packets during congestion
- Drop scalable enhancement layers for constrained endpoints



## 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 temporally 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 -06

Motivated by VP9 SVC usage in SFUs, specifically in mapping VP9 P/U bits to Frame Marking I/B bits.

## 3.1. Extension for Non-Scalable Streams

## 3.2. Extension for Scalable Streams

- Independent (I) bit: temporal independence only
- To match (inverse of) VP9 P bit

## 3.4.2 Scalability Structures

- RECOMMENDED only for temporally nested scalability structures.
- NOT RECOMMENDED for other scalability structures.
- To infer VP9 U bit is always 1 for recommended/nested streams.

# Independence in Non-Scalable Streams

## 3.1. Extension for Non-Scalable Streams

- Independent (I) bit: temporal independence only

I: Independent Frame (1 bit) - MUST be 1 for frames that can be decoded independent of **temporally** prior frames, e.g. intra-frame, VPX keyframe, H.264 IDR [RFC6184], H.265 IDR/CRA/BLA/RAP [RFC7798]; otherwise MUST be 0.

# Independence in Scalable Streams

## 3.2. Extension for Scalable Streams

- Independent (I) bit: temporal independence only

I: Independent Frame (1 bit) - MUST be 1 for frames that can be decoded independent of **temporally** prior frames, e.g. intra-frame, VPX keyframe, H.264 IDR [RFC6184], H.265 IDR/CRA/BLA/RAP [RFC7798]; otherwise MUST be 0. **Note that this bit only signals temporal independence, so it can be 1 in spatial or quality enhancement layers that depend on temporally co-located layers but not temporally prior frames.**

# Scalability Structures

## 3.4.2. Scalability Structures

The LID and TID information is most useful for fixed scalability structures, such as nested hierarchical temporal layering structures, where each temporal layer only references lower temporal layers or the base temporal layer. The LID and TID information is less useful, or even not useful at all, for complex, irregular scalability structures that do not conform to common, fixed patterns of inter-layer dependencies and referencing structures. Therefore it is **RECOMMENDED** to use LID and TID information for RTP switch forwarding decisions **only in the case of temporally nested** scalability structures, and it is **NOT RECOMMENDED for other** (more complex or irregular) scalability structures.

# Discard Priority

- Discard using single “D” bit or highest TID/LID values.
- No changes for “D” bit in version -06.
- More granular priorities for non-scalable streams under consideration in:  
draft-even-avtcore-priority-markings

# Next Steps

- Ready for WGLC, all open issues resolved.
- Questions?
- Thank you!