Discarding Priority of RTP Video Packets

Presenter: Lijun Dong

lijun.dong@futurewei.com

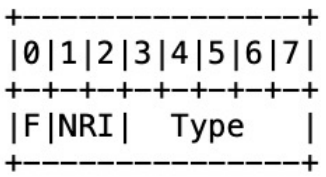
RTP Payload Format for Video

- RTP (Realtime transport protocol) is a protocol dedicated to the transport of real-time video and audio streams.

- RTP streams are typically delivered over UDP which is an unreliable transport mechanism. Hence, there is no guarantee of packet delivery.

- RTP payload formats for different video codecs have been specified:
  - RTP Payload Format for H.264/ Advanced Video Coding (AVC): RFC 6184
  - RTP Payload Format for Scalable Video Coding (SVC): RFC 6190
  - RTP Payload Format for H.265/ High Efficiency Video Coding (HEVC): RFC 7798
  - RTP Payload Format for H.266/ Versatile Video Coding (VVC): draft-ietf-avtcore-rtp-vvc-16

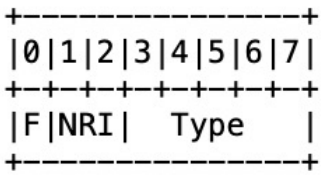
Packet Level Priority Difference in H.264 RTP Packets



- H.264 coded video data is organized into NAL units.
- The RTP packet for H.264 video inherits the same NAL unit header.
- The 2 bits NRI field indicates the relative importance/transport priority of the NAL unit determined by the encoder. A higher value of NRI shows higher importance.
- The 'Type' field indicates the payload format with three different basic payload structures:
  - Single NAL Unit Packet: Contains only a single NAL unit in the payload. The NRI field is associated with this single NAL unit.
  - Aggregation Packet (AP): Packet type used to aggregate multiple NAL units into a single RTP payload. The value of NRI is the maximum of all the NAL units carried in the aggregation packet.
  - Fragmentation Unit (FU): Used to fragment a single NAL unit over multiple RTP packets. All FU packets belong to the same NAL unit have the same NRI value.

More-important NAL units need to be better protected against transmission loss or packet dropping than less-important NAL units.

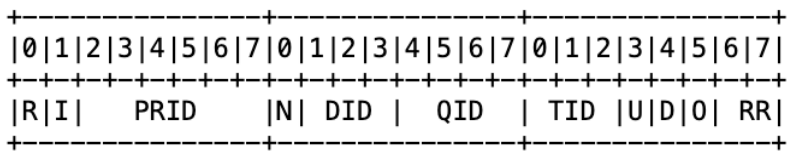
Packet Level Priority Difference in SVC RTP Packets



- SVC defines a coded video representation that supports different levels of scalability: spatial, quality, and temporal.

- Same as H.264, the 2 bits NRI field indicates the relative importance/transport priority of the NAL unit determined by the encoder.

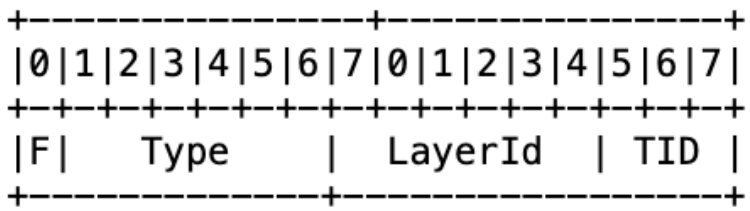
- In addition, there are three octets in the NAL unit header of SVC RTP packets.



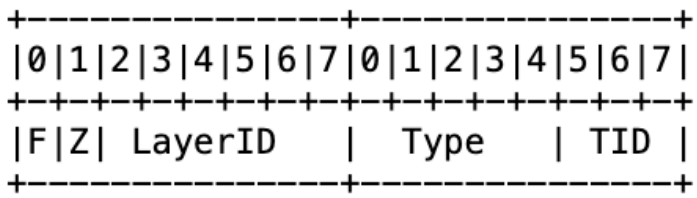
- The priority of a NAL unit in SVC video stream can be further specified by the priority_id field (PRID), which has 6 bits. A lower value of PRID indicates a higher priority.

Packet Level Priority Difference in H.265/HEVC RTP Packets

- HEVC includes an improved support of temporal scalability over H.264, by inclusion of the signaling of TemporalId in the NAL unit header.
- The TID value indicates (among other things) the relative importance of an RTP packet. A lower value of TID indicates a higher importance.



Packet Level Priority Difference in H.266/VVC RTP Packets



- VVC is reported to provide significant coding efficiency gains over H.265/HEVC, and other earlier video codecs.
- VVC uses a two-byte NAL unit header.
- Similar to H.265, the TID value indicates (among other things) the relative importance of an RTP packet. A lower value of TID indicates a higher importance.
- The LayerID field is used to identify the layer a NAL unit belongs to, wherein a layer may be, e.g., a spatial scalable layer, a quality scalable layer, a layer containing a different view, etc.
- The NAL unit with lower LayerID would be likely be used to predict the NAL units with higher LayerID, therefore likely to be more important.

Media Aware Network Element (MANE)

- Due to the explicit layering in the protocol stack, the upper layer data or headers are transparent to the network layer. The priority or importance associated with the NAL units encapsulated in RTP packets is invisible to intermediate routers.

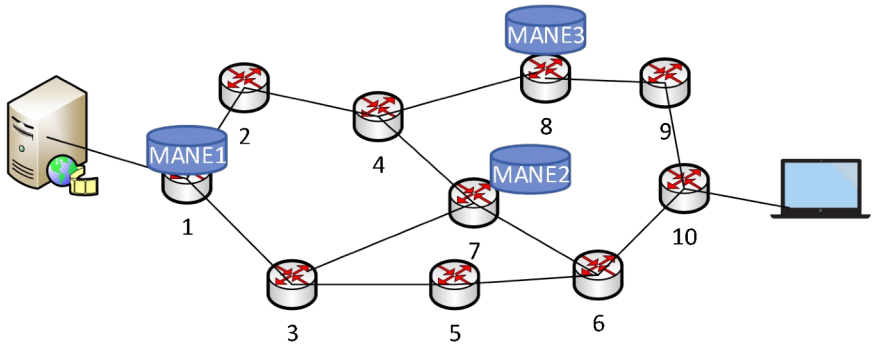
- Media-aware network element (MANE): A network element, such as a middlebox or application layer gateway that is capable of parsing certain aspects of the RTP payload headers or the RTP payload and reacting to the contents.

- The advantage of using MANEs is that they allow packets to be dropped according to the needs of the media coding and produces the least impact on the user experience.

Motivation

- For the congested network node that does not have MANE, the packets are simply dropped without media-awareness.

- In this draft, we intend to extend the MANE's media awareness to the routers, such that the priority based dropping due to network congestion could be applied at routers without MANE.



Realizing Discarding Priority of RTP Video Packets through DSCP

- We consider the two video types: interactive video and non-interactive video. The video stream from both types could be encoded according to H.264, SVC, H.265, H.266.

- For H.264 and SVC, the NAL units have the NRI field to indicate the discarding priority of the RTP packets.

- For H.265 and H.266, the NAL units have the TID field to indicate the discarding priority of the RTP packets.

- The NRI field is of 2 bits, and the TID field is of 3 bits, thus the DSCP value can be mapped according to either the NRI value or the TID value, as well as the video types.

- Either the video host or the MANE at the DiffServ domain edge can do the mapping and set up the DSCP value for each RTP packet. The discarding precedence of the RTP packets can be determined when link congestion happens.

Recommended DSCP Values for RTP Packets According to NRI Value and Video Type (with H.264 or SVC Encoder)

NRI Value	Interactive Video	Non-Interactive Video
11	AF41	AF42
10	AF42	AF43
01	AF31	AF32
00	AF32	AF33

Recommended DSCP Values for RTP Packets According to TID Value and Video Type (with H.265 or H.266 Encoder)

TID Value	Interactive Video	Non-Interactive Video
001	AF41	AF42
010	AF42	AF43
011	AF31	AF32
100	AF32	AF33
101	AF21	AF22
110	AF22	AF23
111	AF11	AF12

Discussions

- RFC 7657: In general, marking packets with different DSCPs results in different PHBs being applied at nodes in the network, making reordering very likely due to use of different pools of forwarding resource for each PHB.
  - Primary example where usage of multiple PHBs does not enable reordering within a single network 5-tuples use PHBs from a single AF class.
    - It would not be possible to map NRI, TID to a single AF class.
  - UDP is not sensitive to reordering in the network.
    - May use different DSCPs whose corresponding PHBs enable reordering within a single UDP tuple。

Discussions (continue)

- In the NAL extension header for SVC, there is an additional field (i.e., PRID) used to indicate the importance of the RTP packet at finer granularity. The PRID field occupies 6 bits additionally.

- In the NAL unit header for H.266, the LayerID is used to identify the layer a NAL unit belongs to. The LayerID field provides the importance information of the RTP packet at finer granularity as well. The LayerID field occupies 6 bits additionally.

- It is not feasible to use the DSCP mapping to indicate the additional discarding precedence provided by the 6 bits PRID, and the 6 bits LayerID.

- Many DSCP values are already used and could be unavailable. Are there any other ways to achieve the packet level dropping precedence within a flow?

- Other solutions need to explored if discarding precedence at finer granularity is considered to be supported.

Dispatch

- AVTCORE
  - Specify and maintain payload formats for use with RTP.

- Transport Area Working Group (tsvwg)

Thank You!

Contact : Lijun Dong (lijun.dong@futurewei.com)
Richard Li (richard.li@futurewei.com)