A Use Case of Packets' Significance Difference with Media Scalability

Presenter: Lijun Dong

lijun.dong@futurewei.com

IETF 112, Online, Nov 8-12, 2021

Feedback from Last Meeting

- The comments are mainly related to how to reveal the characteristics of microblocks contained in the packet to the network and implement the selective packet dropping at packet level.

- Very glad that the concept is very well received: it is beneficial to improve QoE of video streaming by reducing the packet dropping granularity from current traffic class level to the individual packet level, even portions within the packet.

- Yes, exactly, we need such information of whether this packet is significant or not to be revealed to the network nodes. For example, an API could be implemented to input such information or metadata from the application. which might be mapped to IPv6 extension header, IPv4 options or a dedicated metadata field in the IP header.
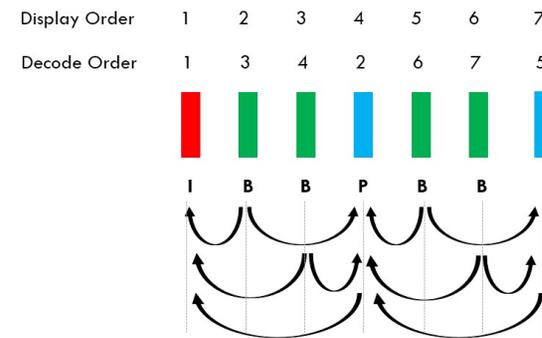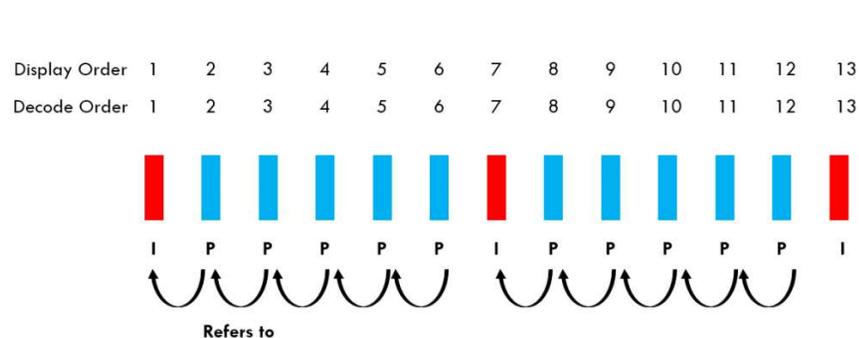
Important Characteristics of MPEG Video Packets

- ***Frame type***
  - I frame: An I-frame consists only of macroblocks that use Intra-prediction.
  - P-frame stands for Predicted Frame and allows macroblocks to be compressed using temporal prediction in addition to spatial prediction. For motion estimation, P-frames use frames that have been previously encoded. In essence, every macroblock in a P-frame can be,
    - temporally predicted, or
    - spatially predicted, or
    - skipped (i.e., the decoder copies the co-located block from the previous frame – a "zero" motion vector).
  - A B-frame is a frame that can refer to frames that occur both before and after it. The B stands for Bi-Directional for this reason. If the video codec uses macroblock-based compression (like H.264/AVC does), then each macroblock of a B-frame can
    - be predicted using backward prediction (using frames that occur in the future)
    - be predicted using forward prediction (using frames that occur in the past)
    - be predicted without inter-prediction – only Intra
    - can be skipped completely (with Intra or Inter prediction).

Important Characteristics of MPEG Video Packets (Continue…)

- ***Whether the packet contains frames that are referenced by other frames:***
  - I frame: I-frame does not refer to any other frame, is at least referenced by a P frame after it. Losing the first I frame in the GOP (Group of Pictures) would cause video picture even missing for few seconds, because P and B frames referencing to the I frame would not be decoded nor displayed either.
  - P frame: A P-frame refers to a picture in the past, might be referenced by a P frame after it, or a B frame before or after it.
  - B frame: A B-frame can act as a reference, and if so, it is termed as a reference B-frame. If a B-frame is not to be used as a reference, it is called a non-reference B-frame.

Important Characteristics of MPEG Video Packets (Continue...)

- ***Movement level of the video sample contained in the packet***
  - Video scenes with a low level of movement are less sensitive to both B-frame and P-frame packet loss, alternatively video scenes with a high level of movement are more sensitive to both B-frame and P-frame packet loss.

  - A lost P frame can impact the remaining part of the GOP.

  - A lost B frame has only local effects in a slowly moving content or with large static background.

  - In a scene of a dynamically moving content, losing B frame has more dramatic impact and its scale can be as far-reaching as a P frame loss.

Important Characteristics of MPEG Video Packets (Continue…)

- ***<u>Whether the video sample contained in the packet belongs to RoI or not</u>***
  - Video sample in RoI is more important than the video sample in non-RoI.

Requirements for Network and Applications

- The application shall reveal some information to the network to enable selective packet dropping. Some examples are listed below:
  - Receiving end user's preference on media quality, e.g., tolerable quality degradation regarding for example resolution.
  - Characteristics of media content contained in the packets, e.g., frame type, whether the packet contains frames that are referenced by other frames, movement level of the video sample contained in the packet.
  - Labeling of the packets or some parts of the packets that correspond to receiver's interested objects as RoI.

- Correspondingly, the network shall be able to leverage the above   information revealed by the application, and selectively drop packets   or parts of the packets from competing media streaming flows with   precedence order when network congestion happens. The retransmission could be maximumly eliminated. The receiving end user is able to consume the delivered packets as many as possible in-time with acceptable quality.

Thank You!

Contact： Lijun Dong (lijun.dong@futurewei.com)