

RTP Payload for APV

draft-ietf-avtcore-rtp-apv-01

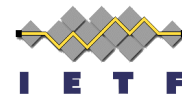
Y. Lim, M. Park, M. Budagavi, R. Joshi, K. Choi

Advanced Professional Video

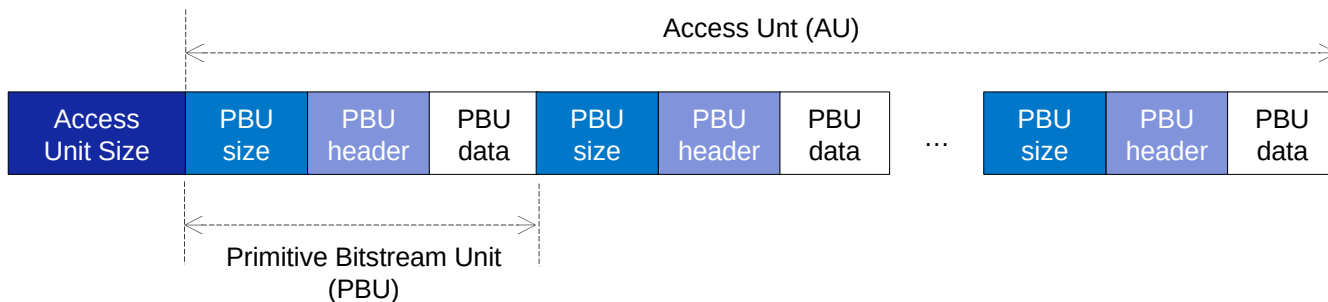


- Informational RFC on APV has been published as [RFC9924](#)
 - Perceptually lossless video quality that is close to the original, uncompressed quality;
 - Low complexity and high throughput intra frame only coding without inter frame coding;
 - Intra frame coding without prediction between pixel values but with prediction between transformed values for low delay encoding;
 - High bit rates of up to a few Gbps for 2K, 4K, and 8K resolution content, enabled by a lightweight entropy coding scheme;
 - Frame tiling for immersive content and for enabling parallel encoding and decoding;
 - Various chroma sampling formats from 4:0:0 to 4:4:4:4, and bit depths from 10 to 16 (Note: Only the profiles supporting 10 bits and 12 bits are currently defined);
 - The ability to decode and re-encode multiple times without severe visual quality degradation; and
 - Various metadata including HDR10/10+ and user-defined formats.

Bitstream structure of APV (1/2)



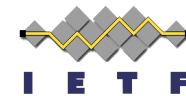
● New bitstream structure



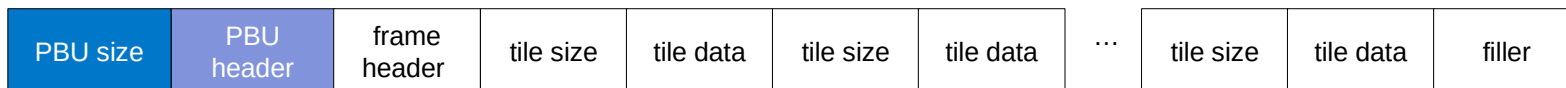
```
-----  
syntax code  
-----  
pbu_header() {  
    pbu_type  
    group_id  
    reserved_zero_8bits  
}
```

- Two layer : An access unit is composed of one or more primitive bitstream unit(s)
 - Access Unit Size can be added for storage and delivery when necessary
- Type of PBUs
 - Coded frame: primary frame, non-primary frame, preview frame, depth frame, alpha frame
 - Metadata: access unit information, metadata
 - Filler: filler
- PBU header provides type and group_id for association with access unit information
 - Simple and fixed structure of PBU size and PBU header enables fast scanning and skip to the target data
- Filler can be added anywhere for fixed size AU or to avoid rewriting of entire bitstream while editing

Bitstream structure of APV (2/2)



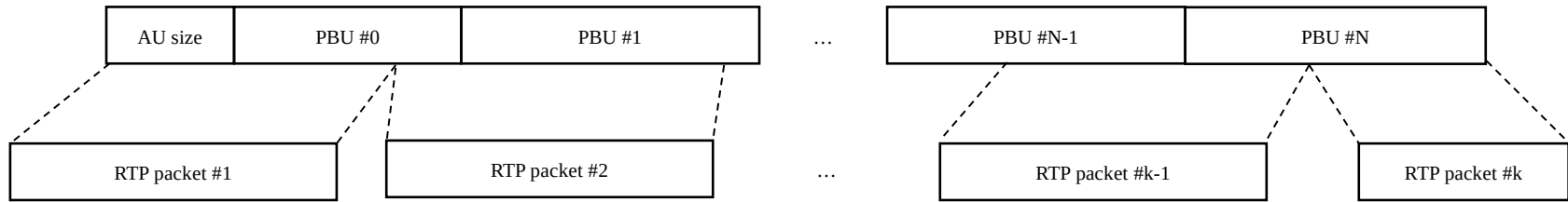
● PBU with coded frame data



- Each frame is self-contained
 - Frame header provide all configuration information for a frame
- Frame is composed with one or more tiles
 - Information about the structure of tiles is provided by frame header
- Filler can be added optionally

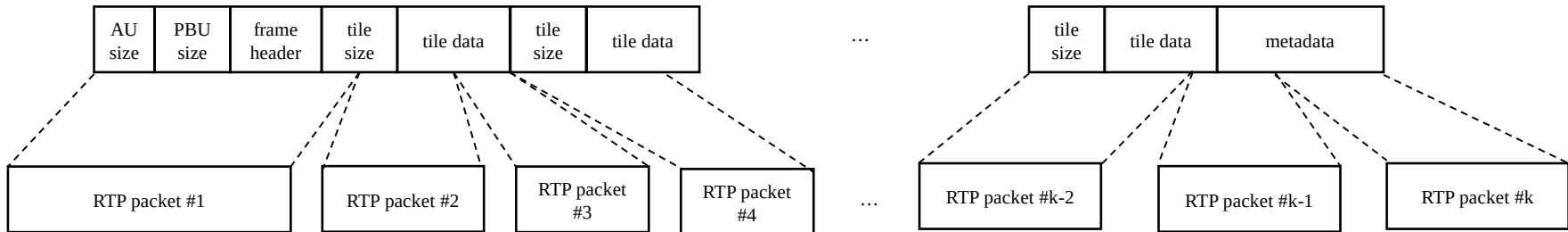
Two packetization modes

- AU size MUST be the first byte of a RTP packet payloads
- Simple mode
 - No alignment with internal data structure of APV bitstream



- Low delay mode

- Beginning of tile size field MUST be aligned with beginning of RTP packet payloads.



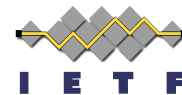
Usage of RTP packet header



- Marker bit (M): 1 bit
 - set to 1 for the last packet of a frame.

- Timestamp: 32 bits
 - The RTP timestamp is set to the sampling time of a frame.
 - A 90 kHz clock rate **MUST** be used.

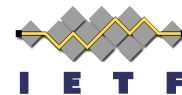
Design of payload header (1/3)



- Version (V) : 2 bits
 - This field indicates the version of the payload header.

- Operation Mode (OM) : 2 bits
 - This field indicates which operation mode is used
 - 00b : reserved
 - 01b : simple mode as defined in Section 5.2
 - 10b : low-delay mode as defined in Section 5.3
 - 11b : reserved

Design of payload header (3/3)



- Frame Header repeated (H) : 1 bit
 - This field indicates that the frame header is repeated in this payload.
 - only used in low delay mode
 - a copy of frame header is added to the end of the payload
- Static Frame Header (S) : 1 bit
 - the values of frame header data is identical except the value of capture_time_distance field with the immediately preceding frame data sent.
- Fragment Counter (FC) : 16 bit
 - the number of remaining payload excluding the current one carrying the current frame or tile

Payload format parameters



- Media type registration
 - Type name: video
 - Subtype name: apv
 - Required parameters: N/A
 - Optional parameters: profile-id, level-id, band-id
- SDP parameters
 - generally one way APV session is offered over RTP using SDP in a declarative style.

Next step

Please review and comment!