RTP Payload Format for SVC Video – draft-ietf-avt-rtp-svc-08

Stephan Wenger, Nokia
stephan.wenger@nokia.com

Ye-Kui Wang, Nokia
ye-kui.wang@nokia.com

Thomas Schierl, HHI
thomas.schierl@hhi.fraunhofer.de

Status of SVC standardization

- Standard approved by ITU-T on 22-Nov-2007:

  Annex G of ITU-T recommendation H.264: 2008, Advanced video coding for generic audiovisual services (4th Edition)

  Pointer to pre-published text:

  http://www.itu.int/rec/T-REC-H.264-200711-P/en

  Pointer to HHI page:

  http://ip.hhi.de/imagecom_G1/savce/downloads/H.264-MPEG4-AVC-Version8-FinalDraft.pdf

- SVC includes the following new scalable profiles:
  - Scalable Baseline
  - Scalable High
  - Scalable High Intra

Relative changes to -03

- Rewritten introduction and terminology (**Thanks to Alex! Who will become an additional author of this draft since the next version!**)
  - new definition of LAYER, is now more inline with JVT
  - lots of others

- Improved signaling / SDP example section

- Less technical changes...but a lot of discussion on decoding order recovery modes (see slides later)

Mailing list traffic, a selective list

- There was a lot of traffics in the mailing list mainly on the decoding order recovery processes, and many private email exchanges between the current authors and Alex mainly on terminology.

- Comments from Alex and Mike already went into the draft.

- Decoding order recovery processes:

  A lot of comments especially on classical RTP mode

  Desirable: Generic method as RTP header extensions
  Colin: Removing header extensions shall not break system

  see summary of other replies on next slides…

Decoding Order Recovery:
with RTP Session Multiplexing

- Two methods are in the draft

- Different PROs / CONs for both methods

- Demands:
  - Backward compatible with RFC3984
  - Full support for existing packetization modes
  - Allow delivering received packets in decoding order even in case of packet loss

  → Neither of the processes can fulfill all of the demands

Decoding Order Recovery
with CL-DON

- By: ○ CL-DON in PACSI (for packetization mode 1)

    ○ extending the DON semantics already present to CL-DON semantics (for packetization mode 2)

- PROs / CONs:

    + No need to discard received NAL units in some cases of loss

    + No need for insertion of additional NAL units

    + Decoding order recovery always possible, if presentation order == decoding order

    – Does not support packetization mode 0 (AhG: Unclear whether mode 0 can be supported)

    – Does not support Fragmentation Units of NON-INTERLEAVED mode (Unclear whether possible)

Decoding Order Recovery
with classical RTP mechanisms

- By: ○ Timestamps(RTP/NTP) and RTP Sequence Numbers
      ○ Dependency Grouping

- PROs / CONs:

  + Full backward compatibility with SINGLE NAL, NON-INTERLEAVED
                              and INTERLEAVED packetization mode

  + Requires no additional buffering parameters / capabilities

  + Decoding order recovery always possible, if presentation order ==
    decoding order

  - If presentation order != decoding order & a certain
    packet loss pattern, CL-DON is more error resilient

  - Requires NTP timestamps for inter session synchronization

  - May require dummy NAL unit insertion or re-ordering of type 14 NAL
    units

Decoding Order Recovery: Possible solutions

- Both modes stay in the draft with detailed description (no pro/con description, but detailed description)

- Negotiation for CL-DON with NON-INTERLEAVED mode: No interleaving allowed!

- POSSIBLE SOLUTION:

  Allow ignoring CL-DON with NON-INTERLEAVED mode and always support classical RTP mechanisms as well – **NOT AGREED / STILL OPEN**

Open issues

1) Cross layer decoding order dependency - two suggested solutions on the table.

2) Backward compatibility to H.264, enabling H.264 (RFC 3984 single NAL unit mode) to interoperate with SVC using base layer. Need more definition.

3) Clarify the PACSI packet since there were changes between the draft revision, TL0PicIdx is now optional.

4) Review the SDP parameters.

Open issue (cont.)

5) Changed semantics between RFC 3984 and SVC like sprop-deint-buf-req - probably will need new parameters.

6) What to do with bugs in RFC 3984.

7) Clarify the usage of the new parameters like sprop-scalability-info, relation to SEI and usage in offer/answer.

8) The text should be clear enough to allow an implementer to use it for creating the payload without having to read the H.264 SVC document.

Open issue (cont.)

9) Non-VCL NAL units, e.g. SEI messages and parameter sets, may be needed by an enhancement layer but not the base layer. However, according to SVC, within an access unit, these non-VCL NAL units must precede VCL NAL units in decoding order. In session multiplexing, should non-VCL NAL units be transported in the same session as the layer that requires the non-VCL NAL unit, or should they be always transported in the base session? It may be impossible to find out without parsing details which session respectively SPS/subset SPS a picture parameter set belongs to. It may make sense for simplicity to allow a MANE to include all of the non-VCL NAL units within all the sessions.

Open issue (cont.)

10) sprop-spatial-resolution: in this draft or a more generic draft?

11) Further to the comment 9 above, if different sessions carry different temporal enhancement layers, then who should get, e.g., the subset SPS. It is actually possible that none does, if transmitted out-of-band. We should enumerate the possibilities and leave no doubt about how it is supposed to work. This can be done in the definition of the RTP sessions in 5.1.2, but even better if specific text is added (after discussion/approval).

12) Do we need to describe the additional NAL unit insertion or re-ordering of Prefix NAL unit (type 14)?

Our to-do list

- Go fast ahead with the other open issues, once decoding order recovery has been cleared!