RTP Payload Format for High-Efficiency Video Coding
draft-ietf-payload-rtp-h265-02.txt

Ye-Kui Wang
yekuiw@qti.qualcomm.com

Yago Sanchez
yago.sanchez@hhi.fraunhofer.de

Thomas Schierl
thomas.schierl@hhi.fraunhofer.de

Stephan Wenger
stewe@stewe.org

Miska Hannuksela
miska.hannuksela@nokia.com

High Efficiency Video Coding (H.265/HEVC)

- ITU-T Rec. H.265 (Apr. 2013), ISO/IEC 23008-2 (Nov. 2013)
- Adoption of RTP-based H.265 into application standards in progress
  - 3GPP SA4 agreed the support of HEVC in all 3GPP multimedia services in Jan 2014
    - MTSI video telephony service
    - PSS streaming service
  - Thus 3GPP has a normative reference dependency on this draft
  - DVB also considers support of RTP-based HEVC in DVB specs
- Tons of HEVC producs out: Galaxy S4, TVs, chipsets, …

- draft-ietf-payload-rtp-h265 became a WG draft on July 1st, 2013
- Current Version -02

Lots of reviews
Thanks so much

- Version -01 had extensive review by at least Bernard, Tickets #2-#8
  - Some tickets addressed architectural issues and related terminology well beyond the payload format scope, on those we recommended informative text (included in v2)
  - Authors propose those tickets be closed—disposition was already on mailing list; we can post a summary after the meeting
- Version (-02) had recently extensive review by Ericsson folks
  - Magnus: 44 (!) comments on scalability, PACI, editorial/small issues
  - Editorials and small issues are/to be addressed by Ye-Kui on list
  - PACI: we tentatively agree with Magnus analysis. Proposal for bug fix in next revision
  - Tickets #9 through #12 by Jonatan; Tickets #9, #11 resolved in principle, #10, #12 active discussions on list

Big Issue: MST in multisource mux environment 1/2

- Both Magnus and Bernard identified this issue, thanks to both for constructive discussions
- When transmitting layers in their own respective RTP packet flow (identified by SSRC) AND other flows are present in the same RTP Session (defined by SSRC numbering space), then we cannot associate layers to a layered bitstream with the signaling currently available
- Same problem exists in RFC 6190, and will come up in any new payload format allowing scalability, multiple description, multiview, … -- any technology that makes it advisable to prune the complete bitstream by a sub-bitstream in middleboxes
- Generic solution is architecturally desirable

Big Issue: MST in multisource mux environment 1/2

- The final solution in the payload format would require to normatively reference a generic solution; something like draft-westerlund-avtcore-rtp-simulcast-03, something based on appid, … we had this discussion in both mmusic and avtcore; it's early in the process and we don't really know where we are going
- Authors (and industry!) don't want to wait for the IETF to sort this out
- However, informative references to draft-ietf-avtext-rtp-grouping-taxonomy-01 are probably OK

- **Carve out MST multisource by declaring out of scope**
- **Explain using terminology from taxonomy document**

- **Is it OK with WG to proceed like this?**

Other Open Issues and Authors' to do list 1/2

- PACI extension mechanism likely broken -> proposal on list shortly

- Ticket #10, marker bit semantics in certain use cases involving scalability
  - Potentially a thorny problem, but use case not clear to Authors
  - In our view, the payload format is NOT required to support any conceivable trick that would be allowed by the H.265 syntax. It's, of course, desirable, but we would be worried to tinker with something as fundamental as the functionality of the marker bit for an unclear use case.
  - Let's wait for results of ongoing list discussion

- Ticket #12, max_fps and the more general issue of out-of-level signaling
  - Under active discussion by people not present at the meeting
  - Wait for results of mailing list discussion

Other Open Issues and Authors' to do list 2/2

- Handling of RFC 5104 payload specific feedback messages
  - RFC 5104 contains some advise for older codecs (up to H.264)
  - Draft currently describes RPSI
  - Add corresponding text for FIR and SLI as well.
- Other RFC 5104 feedback messages don't seem to need an H.265-specific definition, or, for VBCM, that definition would need to be defined by the ITU-T.

Thanks

Back up slides

- Differences relative to RFC 6184 (two slides)
- More details about the multiplexing issue

Compared to RFC 6184 (1/2)

- No multiple packetization modes – just one now
- No support of multi-time aggregation packets (MTAPs)
- Single NAL unit packet still supported
  - With optional inclusion of decoding order number (DON)
- (Single-time) aggregation packets (APs) still supported
  - With optional inclusion of DON
- Fragmentation unit (FU) still supported
  - Only one type
  - With optional inclusion of DON
- Introduced the support of PACI (PAyload Content Information) packets

Compared to RFC 6184 (2/2)

- Multi-stream transmission (MST) is now supported as HEVC has a complete design for temporal scalability support
- A single de-packetization process applies regardless of whether MST or single-stream transmission (SST) is in use, and regardless of whether interleaved packetization is in use
  - Expected to work also for HEVC scalable and 3D extensions
- Includes the HEVC specific use with some feedback messages as specified in AVPF (RFC 4585)
- Included media type parameters "segmentation-id", and "spatial-segmentation-idc“ for parallel processing

Backup Slide: Details re multisrc MST

- The scalability issue
  - Magnus indicated that this issue is significant
  - When multiple RTP streams (or packet streams) are used to carry different (sub-)layers of one video bitstream, the encoded and dependent streams are sent over different RTP packet streams (SSRCs) in the same RTP session, and the SSRCs can't be associated on RTP level for a particular media encoder instance. A single media source and media encoding per endpoint can be determined based on the CNAME associated with the SSRC. But, that fails to work when there are multiple media sources or media encodings from an endpoint in the same RTP session. In this case some solution for identifying the media encoder, and media source is required to enable the RTP receiver to bind these together.
  - Currently there are multiple solutions in discussion. SRCNAME, APPID, etc and they are getting tied into the full complexities of the most advanced use cases for simulcast and telepresence. It is likely that the IETF won't conclude at one or more solutions soon.
- Possible options to proceed (pre discussions with Magnus)
  - To clarify that the scope of this draft only deals with payload format and the de-packetization process, while the grouping signalling is not specified. Thus, an application that utilizes this draft and uses multi-session transmission must specify its own way of grouping signalling, e.g. by referring to whatever solution IETF recommends in the future.
  - To require that all RTP packet stream carrying encoded or dependent streams from the same media encoder instance SHALL have the same SSRC in their respective RTP sessions. But this solution is not generic and not preferred.
  - To clarify in the draft that some grouping signalling needed for multi-session transmission is currently missing, and the RFC will need to be updated once one or more solutions are specified by the IETF.

- We request that the group make a decision to take one of the above options to proceed.