**3GPP TSG-SA4 Meeting # 128 S4-241074**

**Jeju, KR, 20th - 24th May 2024**

**Source: Apple Inc.**

**Title: [VOPS] On codec string encoding for L-HEVC**

**Agenda item: 9.5**

**Document for: Agreement**

# 1 Overview

One of the possibilities to carry layered HEVC (L-HEVC) video in mp4 is by using the 'hvc1' or 'hev1' sample entry type as specified in clause 9 of ISO/IEC 14496-15 in a backwards compatible manner [3]. Existing products in the market are using this concept to carry stereoscopic content and alpha using the Multiview extensions of the HEVC standard as L-HEVC in mp4 [1][2]. Furthermore, such profiles with existing support in the mobile ecosystem are targeted for the ongoing SA4 VOPS work item.

However, when constructing the MIME types 'codecs' parameter, according to Annex E of [3], the MIME type specification only includes signalling of the base layer and does not provide necessary signalling for other layers. For other layers an additional mime type parameter is required, but at least one W3C API accepts a MIME type with no extra MIME parameters except for codecs. As per the W3C's [Media Capabilities API](https://www.w3.org/TR/media-capabilities/#http):

If the MIME type does not imply a codec, the string MUST also have one and only one parameter that is named codecs with a value describing a single media codec.

Furthermore, the signalling inside the codecs string does not expose other important information such as the types of auxiliary information that would allow us to obtain necessary information about rendering aspects of a stream.

At MPEG the File Format group currently studies several possible solutions to this problem [4]. One possible solution is based on an optional extension of the codecs parameter by attaching a new 4CC with additional parameters that will provide additional information. The table below summarizes the currently considered parameters that can be used in the extension.

|  |  |  |
| --- | --- | --- |
| Attribute Type(Upper case letters only) | Attribute values | Description |
| N | Decimal number | Decimal number of the nuh\_layer\_id |
| T | One of the following decimal numbers1 – texture2 - auxiliary s1 – alpha s2 – depth3 - other (e.g. 3D-HEVC depth) | The type of the layer.We need to signal if the layer is a primary, aux or 'other'. ('other' can be used to support 3D HEVC)If aux is signalled, then you indicate the actual type by the sub-type.E.g. T2s1 means auxiliary alpha |
| B | Decimal number  | Decimal number of bit depth minus 8 |
| S | One of the following decimal numbers0 – monochrome1 – 4:2:02 – 4:2:23 – 4:4:4 | Decimal number of the chroma\_format\_idc |
| C | The following lower-case letters followed by a decimal number signaling the CICP parameters:m[N] - matrix coefficientst[N] - transfer characteristicsc[N] - colour primaries | Colour properties signaled by decimal numbers as in CICP (or colr box).E.g.: Cm3t4c5 |
| X | Alphanumeric characters. | Profile, tier and level signaling as defined in Annex E.3 but where dots (".") are replaced by underscores ("\_"). |
| D | Dependency indication. does this layer depend on other layers?i - independentarray separated by 'L' followed by a decimal number of the layer id to indicate inter-layer references | Examples:Di - independent layerDL1L12L24 - this layer depends on layers 1, 12 and 24 |

Such a new element could be defined as also 4CC and attached after a comma (",") right after the signalled codec. In the example below the profile, tier, level signaling is simar to the signaling in HEVC but the dot separators are replaced with underscores since dots are used to separate layers.

codecs= "hvc1.1.6.L93.B0,also.N0T1B2S1X1\_6\_L93\_B0.N1T2s1B0S1X\_1\_6\_L93\_B0"

The also 4CC will need to be registered as a new sample entry type at MP4RA, to avoid possible collisions with future codecs. It should also be noted that this signalling allows the content author to add an additional description of the L-HEVC stream in a single track at the application level. One use case would be the ability to signal multiple codec strings in one manifest, and if the implementation cannot handle the extended codec string, it can fall back to the legacy codec string. Alternatively, such a legacy string can even be treated as a default codec signalling, and an extended version of the codec string can be used by players that are expected to support it, while legacy players ignore it.

# 2 Proposal

Given the challenges identified in layer signalling within MIME types for layered HEVC (L-HEVC) video, it is important that a standardized solution be developed to enhance interoperability, accuracy, and efficiency of multi-layer video stream handling. We propose that SA4 officially recognizes these challenges and articulate the need for a solution that provides comprehensive, backward-compatible signalling.

High level requirements for the proposed solution:

* **Comprehensive Layer Signaling**: Enable the signaling of multiple video layers in 3GPP-based services, including DASH MPDs (see TS 26.511), capability checks in the context of 5G Media Streaming (see TS 26.511), as well as for Media Messaging Services (see TS 26.143) . This should include but not be limited to the number of layers, types of each layer, their inter-dependencies, etc.
* **Backward Compatibility**: Ensure that the solution maintains backward compatibility, allowing existing players and systems to continue functioning without modifications while enabling enhanced capabilities for updated systems.
* **Adaptability and Extensibility**: Design the solution to be adaptable for future extensions and new types of layers or enhancements without requiring significant overhauls. Consider making the signalling codec agnostic.

Action Requested from SA4:

* **Acknowledgment of the Problem**: Confirm the need to address the identified gap in the signaling capability for layered HEVC.
* **Requirements Communication to MPEG**: Advocate for the development of a standardized solution at MPEG that meets the outlined high-level requirements. Propose to draft and send a Liaison Statement to MPEG expressing these needs and requesting their engagement in developing a comprehensive solution.
* **Potentially update TR 26**.966 with a CR to address the identified needs and requirements for consistent tracking of the problem.

By addressing these requirements, we aim to significantly enhance the capability of multimedia systems to handle complex video streams more effectively and efficiently, thus ensuring a robust multimedia experience across diverse devices and platforms.

# References

1. Apple Inc. "Apple HEVC Stereo Video Interoperability Profile (Beta),” Version 0.9 (Beta) June 21, 2023, [online]: https://developer.apple.com/av-foundation/HEVC-Stereo-Video-Profile.pdf
2. Apple Inc. "HEVC Video with Alpha: Interoperability Profile", [online] https://developer.apple.com/av-foundation/HEVC-Video-with-Alpha-Interoperability-Profile.pdf
3. ISO/IEC 14496-15: Information technology — Coding of audio-visual objects — Part 15: Carriage of network abstraction layer (NAL) unit structured video in the ISO base media file format
4. Technologies under Consideration for ISO/IEC 14496-15 Carriage of NAL unit structured video in ISOBMFF, [online] [WG03N1131\_23498](https://www.mpeg.org/wp-content/uploads/mpeg_meetings/145_OnLine/w23498.zip)