**3GPP TSG-SA4 Meeting #128 *S4-240910***

**Jeju, Korea, 20 - 24 May 2024**

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| *CR-Form-v12.2* | | | | | | | | |
| **Pseudo CHANGE REQUEST** | | | | | | | | |
|  | | | | | | | | |
|  | **26.265** | **CR** | **pseudo** | **rev** | 1 | **Current version:** | **0.1.0** |  |
|  | | | | | | | | |
| *For* [***HE******LP***](http://www.3gpp.org/3G_Specs/CRs.htm#_blank)*on using this form: comprehensive instructions can be found at* [*http://www.3gpp.org/Change-Requests*](http://www.3gpp.org/Change-Requests)*.* | | | | | | | | |
|  | | | | | | | | |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Proposed change affects:*** | UICC apps |  | ME | **X** | Radio Access Network |  | Core Network | **X** |

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | | | | | | | | | | |
| ***Title:*** | [VOPS] Introduction to TS 26.565 | | | | | | | | | |
|  |  | | | | | | | | | |
| ***Source to WG:*** | Qualcomm Incorporated, Tencent | | | | | | | | | |
| ***Source to TSG:*** |  | | | | | | | | | |
|  |  | | | | | | | | | |
| ***Work item code:*** | VOPS | | | | |  | ***Date:*** | | | 2024-05-14 |
|  |  | | | |  | |  | | |  |
| ***Category:*** | **B** |  | | | | | ***Release:*** | | | Rel-19 |
|  | *Use one of the following categories:* ***F*** *(correction)* ***A*** *(mirror corresponding to a change in an earlier release)* ***B*** *(addition of feature),* ***C*** *(functional modification of feature)* ***D*** *(editorial modification)*  Detailed explanations of the above categories can be found in 3GPP [TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm). | | | | | | | | *Use one of the following releases: Rel-8 (Release 8) Rel-9 (Release 9) Rel-10 (Release 10) Rel-11 (Release 11) … Rel-16 (Release 16) Rel-17 (Release 17) Rel-18 (Release 18) Rel-19 (Release 19)* | |
|  |  | | | | | | | | | |
| ***Reason for change:*** | | There were agreements in S4-240619, clause 5.1.1. This is implemented. | | | | | | | | |
|  | |  | | | | | | | | |
| ***Summary of change:*** | | Add Introduction | | | | | | | | |
|  | |  | | | | | | | | |
| ***Consequences if not approved:*** | |  | | | | | | | | |
|  | |  | | | | | | | | |
| ***Clauses affected:*** | | 2, 6 | | | | | | | | |
|  | |  | | | | | | | | |
|  | | **Y** | **N** |  | | | |  | | |
| ***Other specs*** | |  | **X** | Other core specifications | | | | TS/TR ... CR ... | | |
| ***affected:*** | |  | **X** | Test specifications | | | | TS/TR ... CR ... | | |
| ***(show related CRs)*** | |  | **X** | O&M Specifications | | | | TS/TR ... CR ... | | |
|  | |  | | | | | | | | |
| ***Other comments:*** | |  | | | | | | | | |
|  | |  | | | | | | | | |
| ***This CR's revision history:*** | |  | | | | | | | | |

## ===== CHANGE =====

# 2 References

The following documents contain provisions which, through reference in this text, constitute provisions of the present document.

- References are either specific (identified by date of publication, edition number, version number, etc.) or non‑specific.

- For a specific reference, subsequent revisions do not apply.

- For a non-specific reference, the latest version applies. In the case of a reference to a 3GPP document (including a GSM document), a non-specific reference implicitly refers to the latest version of that document *in the same Release as the present document*.

[1] 3GPP TR 21.905: "Vocabulary for 3GPP Specifications".

[h264] ITU-T Recommendation H.264 (08/2021): "Advanced video coding for generic audiovisual services".

[h265] ITU-T Recommendation H.265 (09/2023): "High efficiency video coding".

[CMAF] ISO/IEC 23000-19: "Information Technology Multimedia Application Format (MPEG-A) – Part 19: Common Media Application Format (CMAF) for segmented media".

[CENC] ISO/IEC 23001-7: "MPEG systems technologies - Part 7: Common encryption in ISO base media file format files".

[DPC] CTA-5003-A & Errata: "Web Application Video Ecosystem (WAVE): Device Playback Capabilities Specification", available at <https://cdn.cta.tech/cta/media/media/resources/standards/pdfs/cta-5003-final.pdf>.

[6381] IETF RFC 6381: The 'Codecs' and 'Profiles' Parameters for "Bucket" Media Types.

[MSE] 3GPP TR 26.857, "5G Media Service Enablers"

## ===== CHANGE =====

# 4 Context and Definitions

Editor’s Note from 619, clause 5.1

The principles of existing video capabilities are built around the following principles:

**Bitstream:** A media bitstream that conforms to a video encoding format and certain Operation Point.

**Operation Point:** A collection of discrete combinations of different content formats including spatial and temporal resolutions, colour mapping, transfer functions, etc. and the encoding format.

**Receiver:** A receiver that can decode and render any bitstream that is conforming to a certain Operation Point.

Decoding capabilities are defined which are a combination of

* The capability to decode a bitstream conforming to a certain profile and level
* The bitstream being restricted in terms of flags and settings

An illustration of an operation points and decoding capabilities is provided below.

A diagram of a diagram with Crust in the background

Description automatically generated

*Receivers* are a combination of decoding capabilities and the ability to rendering the formats included in an operation point.

The timing and the properties of the format may be signaled in the bitstream, or may be signaled by external means, for example on packaging level, i.e. on ISO BMFF or RTP level.

At the receiving end conformance always refers to real-time decoding and rendering.

Bitstreams can either conform to any of the above “circles”

* Codec & Profile
* Level
* Decoding capabilites
* Operation Point

Concurrent decoding capabilities are defined as the ability to decode several bitstreams in parallel.

Encoding capabilities are defined by the ability to encode a *video signal* with certain boundary parameters to a bitstream that is decodable (and possibly can be rendered). Typically, specifications would require real-time encoding.

Decoding capabilities can be shared across many different applications.

Operation Points are more specific towards applications and may not or only partially be defined in a new spec.

## 4.1 Motivation

Video codecs, encoders and decoders are core components of 3GPP services. At the same time, video encoders and decoders residing on 3GPP UEs and defined in 3GPP specifications also provide interoperability points for third-party services. Video capabilities are predominantly independent of the service in use. This specification addresses the definition of video capabilities and operating points such that 3GPP service specifications as well as third-party service providers can refer to the interoperability points defined in this specification.

The present specification makes use some of the concepts recommended in TR 26.857 [2], i.e. the concept of Media Service Enablers.

## 4.2 Reference architectures and definitions

In order to define the normative aspects of this specification, reference architectures are defined. The core architecture is provided in Figure 4.2-1. The workflow addresses the generation of a *video bitstream* from a video signal using a *video encoder* as well as the decoding of a video bitstream by a *video decoder* and providing the resulting decoded video as well as associated metadata to a rendering and display process. The video encoder as well as the video decoder may be configured to certain operations indicated by APIs in Figure 4.2-1. These APIs are not normatively specified but serve as an example reference to configure encoders and decoders as documented in Annex [A].



Figure 4.2-1 Reference architecture for video operating points and capabilities

A more system-centric architecture is provided in Figure 4.2-2. The workflow addresses the generation of a *transport stream* from a video signal using a *video encoder* and a *packager*. The package may include for example timing and metadata information. The de-packaging and decoding of the *transport stream* by a de-packager and a *video decoder*, respectively, allows for providing the resulting video signal as well as associated metadata to a rendering and display process. Again, the packager/encoder as well as the de-packager/decoder may be configured to certain operations indicated by APIs in Figure 4.2-2.



Figure 4.2-2 Reference architecture for system operating points and capabilities

Editor’s Note: A reference architecture for multiple decoders still needs to be defined.

Based on this introduction, the following terms are defined

**Operating Point:** A collection of different possible video formats including spatial and temporal resolutions, colour mapping, transfer functions, etc. and a video encoding format.

**Bitstream:** A compressed media representation presented as a sequence of bits that conforms to a particular video coding specification/format and one or more Operating Points.

**Receiver:** A device that can ingest and decode any bitstream that is conforming to a particular video coding specification and Operating Point, and optionally render it.

In addition, on system level the following terms are defined:

**System Operating Point:** A collection of different possible video formats including spatial and temporal resolutions, colour mapping, transfer functions, etc., a video encoding and a packaging format.

**Transport Stream:** A packaged media bitstream that conforms to a particular video coding and packaging specification/format and one or more Operating Points.

**System Receiver:** A receiver that can de-package and decode any system bitstream that is conforming to a particular System Operating Point, and optionally render it.

## 4.3 Specification

This specification defines the following capabilities:

- Video Decoding capability: The capability to decode any video bitstream that conforms to an operating point and provides a conforming output video signal and possibly associated metadata.

- System Receiver capability: The capability to un-package and decode any transport stream that conforms to a system operating point and provides a conforming output video signal and possibly associated metadata

- Video Encoding capability: The capability to encode any video signal included in the operating point to a bitstream that is decodable by decoder that conforms to the same operating point.

- System Transmitter capability: The capability to encode and package any video signal included in the operating point to a system bitstream that can be unpacked and decoded by a system receiver that conforms to the same operating point.

While not explicitly stated in the capabilities, it is a requirement for decoders and receivers to process the data in real-time. For encoder, real-time encoding is typically also a requirement.

This specification primarily focuses on video encoding and decoding capabilities.

## 4.4 Video representation formats

Editor’s Note: Need to define video signal parameters, SDR, HDR, etc.

## 4.5 Reference API parameters

Editor’s Note: The following parameters just summarize parameters available.

Decoding:

* Codec string
* Metadata processing

Encoding:

* Codec string
* Bitrate and bitrate modes
* displayWidth/Height
* framerate
* latency modes
* Codec specific parameters
* Metadata

Packaging