**3GPP TSG- Meeting #**

**, , -  *revision of S4-240696***

|  |
| --- |
| *CR-Form-v12.3* |
| **CHANGE REQUEST** |
|  |
|  |  | **CR** |  | **rev** | **1** | **Current version:** |  |  |
|  |
| *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:***  |  |
|  |  |
| ***Source to WG:*** |  |
| ***Source to TSG:*** | S4 |
|  |  |
| ***Work item code:*** |  |  | ***Date:*** |  |
|  |  |  |  |  |
| ***Category:*** |  |  | ***Release:*** |  |
|  | *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 canbe 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-17 (Release 17)Rel-18 (Release 18)Rel-19 (Release 19) Rel-20 (Release 20)* |
|  |  |
| ***Reason for change:*** | IVAS CMAF/DASH support was incomplete |
|  |  |
| ***Summary of change:*** | Corrections to fix support for IVAS- Corection to reference the IVAS Operation Point- Implementation of IVAS Media Profile |
|  |  |
| ***Consequences if not approved:*** | Lack of support for IVAS for 5GMS |
|  |  |
| ***Clauses affected:*** | 5.2/5.3, 6.1, 6.2.1, 6.3.5, 7.5, A.1 |
|  |  |
|  | **Y** | **N** |  |  |
| ***Other specs*** | **X** |  |  Other core specifications  | TS 26.511 CR 0012 |
| ***affected:*** |  | **X** |  Test specifications | TS 26.143 CR 0001 |
| ***(show related CRs)*** |  | **X** |  O&M Specifications |  |
|  |  |
| ***Other comments:*** |  |
|  |  |
| ***This CR's revision history:*** |  |

\* \* \* First Change \* \* \* \*

## 5.2 Decoding Capabilities

The following speech media decoding capabilities are defined:

- **AMR**: All decoding requirements for the AMR speech codec as specified in 3GPP TS 26.071 [3], 3GPP TS 26.090 [4], 3GPP TS 26.073 [5] and 3GPP TS 26.104 [6]) including all 8 modes and source-controlled rate operation ‎3GPP TS 26.093 [7].

- **AMR-WB**: All decoding requirements for the AMR-WB codec as specified in 3GPP TS 26.171 ‎‎[8], 3GPP TS 26.190 ‎[9], 3GPP TS 26.173 ‎[10] and 3GPP TS 26.204 [11] including all 9 modes and source-controlled rate operation ‎3GPP TS 26.193 [12].

- **EVS**: All decoding requirements for the EVS codec as specified in 3GPP TS 26.441 [13], 3GPP TS 26.445 [16], 3GPP TS 26.442 [14] and 3GPP TS 26.443 [15] as described below including functions for backwards compatibility with AMR-WB (3GPP TS 26.446 [17]) and discontinuous transmission (3GPP TS 26.450 [18]).

NOTE: Speech media decoding capabilities with the IVAS codec are identical to audio media decoding capabilities with IVAS as defined below.

The following audio media decoding capabilities are defined:

- **eAAC***+*: All decoding requirements for the eAAC+ audio codec as specified in 3GPP TS 26.401 [19], 3GPP TS 26.402 [20], 3GPP TS 26.410 [24] and 3GPP TS 26.411 [25].

- **AMR-WB***+*: All decoding requirements for the AMR-WB+ audio codec as specified in 3GPP TS 26.290 ‎‎[26], 3GPP TS 26.304 ‎[27] and 3GPP TS 26.273 [28].

- **xHE-AAC stereo**: All decoding requirements for the xHE-AAC stereo audio codec as specified in the MPEG-D USAC "Extended high efficiency AAC profile" as defined in ISO/IEC 23003-3 [37] as well as all processing requirements applicable to the MPEG-D DRC loudness control profile and to the dynamic range control profile, level 1 or higher, as specified in ISO/IEC 23003-4 [38].

NOTE: xHE-AAC® is a registered trademark of Fraunhofer in Germany and other countries and is used with Fraunhofer’s permission.

* **IVAS**: All decoding and rendering requirements for the IVAS codec as specified in 3GPP TS 26.250 [x1], TS 26.252 [42], TS 26.253 [43], TS 26.254 [44], TS 26.255 [45], TS 26.256 [46], and one of TS 26.251 (fixed-point) [47] or TS 26.258 (floating-point) [48].

NOTE: The IVAS decoder supports decoding of streams encoded with EVS. Therefore, support of IVAS media decoding capabilities implies support of EVS media decoding capabilities.

- **AAC-ELDv2**: the capability to decode MPEG-4 Low Delay AAC v2 Profile Level 2 bitstreams [x9] and to output it as 2-channel audio. Note that this profile contains the audio object types 23 (ER AAC LD), 39 (ER AAC ELD) and 44 (LD MPEG Surround).

## 5.3 Encoding Capabilities

The following speech media encoding capabilities are defined:

- **AMR**: The encoding requirements for the AMR speech codec as specified in 3GPP TS 26.401 [19], clause 7, as well as 3GPP TS 26.403 [21], 3GPP TS 26.404 [22] and 3GPP TS 26.405 [23].

- **AMR-WB**: The encoding requirements for the AMR-WB by one of the following methods:

- according to 3GPP TS 26.173 ‎[10]

- according to 3GPP TS 26.204 [11];

- the AMR-WB IO mode according to TS 26.442 [14] and TS 26.443 [15],

- the AMR-WB IO mode according to TS 26.452 [34].

- **EVS**: The encoding requirements for the EVS codec by one of the following methods:

- TS 26.442 [14] and TS 26.443 [15] encoding functions; or

- TS 26.452 [34] encoding functions.

NOTE: Speech media encoding capabilities with the IVAS codec are identical to audio media encoding capabilities for IVAS as defined below.

The following audio media encoding capabilities are defined:

- **eAAC***+*: The encoding requirements for the AAC+ audio codec as specified 3GPP TS 26.401 [19], clause 7, as well as 3GPP TS 26.403 [21], 3GPP TS 26.404 [22] and 3GPP TS 26.405 [23].

- **AMR-WB***+*: The encoding requirements for the AMR-WB+ audio codec by one of the following methods

- according to 3GPP TS 26.273 [28]; or

- according to 3GPP TS 26.304 [27].

- **xHE-AAC stereo**: All encoding requirements for the xHE-AAC stereo audio codec as specified in the MPEG-D USAC "Baseline USAC" profile as defined in ISO/IEC 23003-3 [37] and with the additional requirements that all encoded media contains the required metadata sets conforming to the MPEG-D DRC loudness control profile or to the dynamic range control profile, level 1 or higher, as specified in ISO/IEC 23003-4 [38].

**- IVAS**: All encoding requirements for the IVAS codec as specified in 3GPP TS 26.250 [41], TS 26.252 [42], TS 26.253 [43], and one of TS 26.251 (fixed-point) [47] or TS 26.258 (floating-point) [48].

NOTE: The IVAS encoder supports EVS encoding. Therefore support of IVAS media encoding capabilities implies support of EVS media encoding capabilities.

- **AAC-ELDv2**: the capability to encode MPEG-4 Low Delay AAC v2 Profile Level 2 according to ISO/IEC 14496-3 [x9]. Note that this profile contains the audio object types 23 (ER AAC LD), 39 (ER AAC ELD) and 44 (LD MPEG Surround).

\* \* \* Next Change \* \* \* \*

## 6.1 Introduction

The speech and audio Operation Points defined in this clause are primarily introduced in order to be used as content format in the context of 5G Media Streaming, but not restricted to this use case.

An operation point is a combination of rendering formats and media decoding capabilities.

For each Operation Point, Bitstream and Receiver requirements are detailed in the remainder of clause 6.

Table 6.1 provides an overview of the Operation Points defined in the present document.

Table 6.1: Speech and Audio Operation Points

|  |  |  |  |
| --- | --- | --- | --- |
| Operation Point name | Format Properties  | Decoding and Encoding Capabilities | Reference |
| AMR speech | Sampling frequency: 8 kHz | *AMR* | 6.2.2 |
| AMR-WB speech | Sampling frequency: 16 kHz | *AMR-WB* | 6.2.3 |
| EVS mono | Sampling frequency: 8, 16, 32, 48 kHz | *EVS* | 6.2.4 |
| eAAC+ stereo | Sampling frequency: 32, 44.1, 48 kHz | *eAAC+* | 6.3.2 |
| AMR-WB+ | Sampling frequency: 8, 16, 32, 48 kHz | *AMR-WB+* | 6.3.3 |
| xHE-AAC stereo | Sampling frequency: 32, 44.1, 48 kHz | *xHE-AAC stereo* | 6.3.4 |
| IVAS | Sampling frequency: 16, 32, 48kHz | *IVAS* | 6.3.5 |

\* \* \* Next Change \* \* \* \*

### 6.2.1 Introduction

This clause defines speech operation points. For each operation point, the requirements for the bitstream as well as for the receiver are defined.

Note: Speech operation relying on IVAS codec falls back to the audio operation point for IVAS codec as defined below under section 6.3.5.

\* \* \* Next Change \* \* \* \*

### 6.3.5 IVAS

#### 6.3.5.1 Bitstream Encoding Requirements

The following requirements apply to the **IVAS** Operation Point.

- The input audio format shall be either mono, stereo, binaural, multi-channel (5.1, 5.1.2, 5.1.4, 7.1, 7.1.4), scene-based (Ambisonics up to 3rd order), metadata-assisted spatial audio (MASA), object-based, a combined format of objects with scene-based (OSBA), or a combined format of objects with metadata-assisted spatial audio (OMASA).

- The sampling frequency shall be either 8 kHz (only EVS interoperable coding), 16 kHz, 32 kHz and 48 kHz (fullband audio content).

- The bitstream shall be encoded according to 3GPP TS 26.250 [41], TS 26.252 [42], TS 26.253 [43], and one of TS 26.251 (fixed-point) [47] or TS 26.258 (floating-point) [48].

NOTE: IVAS codec level setting is TBD.

#### 6.3.5.2 Receiver Requirements

Receivers conforming to the **IVAS** Operation Point shall support the *IVAS*media decoding capability according to clause 5.2 and shall support rendering and playback of the decoded signal.

NOTE: The IVASdecoder supports decoding of streams encoded with EVS. Therefore support of *IVAS* media decoding capabilities implies support of *EVS* media decoding capabilities.

NOTE: IVAS codec level setting is TBD.

#### 6.3.5.3 Sender Requirements

Senders conforming to the **IVAS** Operation Point shall support the **IVAS-Enc** media encoding capability according to clause 5.3 in real-time for the audio formats according to the supported IVAS codec level 1, 2 or 3 as either mono, stereo, binaural, multi-channel (5.1, 5.1.2, 5.1.4, 7.1, 7.1.4), scene-based (Ambisonics up to 3rd order), metadata-assisted spatial audio (MASA), and object-based with sampling frequency 8 kHz (only EVS interoperable coding), 16 kHz, 32 kHz and 48 kHz (fullband audio content).

NOTE: IVAS codec level setting is TBD.

\* \* \* Next Change \* \* \* \*

## 7.5 IVAS Media Profile

### 7.5.1 Mapping to ISO BMFF

If media is provided following the operation point **IVAS** and is encapsulated in the ISO BMFF, then the file format track shall contain the IVASSampleEntryBox with box\_type and conform to the requirements of the sample entry 'sivs' as defined in TS 26.244 [29].

### 7.5.2 Media Profile Definition

#### 7.5.2.1 CMAF Track Definition

If media is provided following the operation point **IVAS** and is encapsulated in a CMAF track, then the CMAF track shall conform to clause 7.5.1, and conform to the general CMAF Track constraints in ISO/IEC 23000-19 [30], clause 7 as well as the general audio track constraints defined in ISO/IEC 23000-19 [30], clause 10.

#### 7.5.2.2 CMAF Switching Set and Media Profile Definition

If media is provided following the operation point **IVAS** and is provided in a CMAF Switching Set, then every CMAF track in the CMAF Switching Set shall conform to clause 7.5.2.1, and shall conform to the general CMAF Switching Set constraints in ISO/IEC 23000-19 [30], clause 7. A CMAF Switching Set that follows these requirements conforms to the CMAF IVAS media profile 'civs' defined in this clause.

#### 7.5.2.3 Mapping to DASH Adaptation Set

If media is provided following the operation point **IVAS** and is provided in a DASH Media Presentation in an Adaptation Set, a switching set conforming to clause 7.5.2.2 may be provided in a DASH Media Presentation Description in an Adaptation Set. In that case, the Adaptation Set shall conform to the Adaptation set constraints of the DASH profile for CMAF as defined in ISO/IEC 23009-1 [31]. The following parameters shall be present on Adaptation Set level:

- @codecs is set to 'sivs'

- @mimeType is set to be compatible with "audio/mp4 profiles='civs'"

- @audioSamplingRate is set to one of the following: '16000', '32000', '48000'

If the Adaptation Set conforms to the constraints for the **IVAS** Operation Point as defined in this clause, then the @profiles parameter in the Adaptation Set may signal conformance to this Media Profile by using "urn:3GPP:audio:mp:ivas.

#### 7.5.2.4 Playback Requirements

For a receiver supporting the IVAS media profile the following applies:

- It shall support the receiver requirements as documented in clause 6.3.5.2 for any CMAF Track conforming to the CMAF IVAS media profile 'civs' as defined in clause 7.5.2.2.

- It shall support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF IVAS media profile 'civs' as defined in clause 7.5.2.2, namely:

- 8.2 Sequential Track Playback

- 8.3 Random Access to Fragment

- 8.4 Random Access to Time

- 8.5 Switching Set Playback

- 8.6 Regular Playback of Chunked Content

- 8.7 Regular Playback of Chunked Content, non-aligned append

- It should support the following playback requirements as documented in clause 8 of CTA-WAVE 5003 [32] for any content conforming to a CMAF Switching Set according to CMAF IVAS media profile 'civs' as defined in clause 7.5.2.2, namely:

- 8.9 Out-Of-Order Loading

- 8.10 Overlapping Fragments

- 8.12 Playback of Encrypted Content

#### 7.5.2.5 Content Generation Requirements

For a transmitter supporting the IVAS media profile the following applies:

- It shall support all media encoding capabilities for IVAS as defined in clause 5.3.

- It shall support the sender requirements for IVAS as defined in clause 6.3.5.3.

- It shall support the generation of a CMAF Track as defined in clause 7.5.2.1 that conforms to the CMAF Media Profile 'civs' as defined in clause 7.5.2.2.

- If used for Adaptive Bit Rate (ABR) distribution, it shall support the generation of a CMAF Switching Set as defined in clause 7.5.2.2.

\* \* \* Next Change \* \* \* \*

# A.1 3GPP Registered URIs

The clause documents the registered URIs in the present document following the process in <http://www.3gpp.org/specifications-groups/34-uniform-resource-name-urn-list>

Table A-1 lists all registered URN values as well as:

- a brief description of its functionality;

- a reference to the specification or other publicly available document (if any) containing the definition;

- the name and email address of the person making the application; and

- any supplementary information considered necessary to support the application.

Table A-1: 3GPP Registered URNs

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| URN | Description | Reference | Contact | Remarks |
| urn:3GPP:audio:mp:amr | AMR Media Profile | TS 26.117, clause 7.2.2.3  | Thomas Stockhammertsto@qti.qualcomm.com | none |
| urn:3GPP:audio:mp:amr-wb | AMR-WB Media Profile | TS 26.117, clause 7.3.2.3 | Thomas Stockhammertsto@qti.qualcomm.com | none |
| urn:3GPP:audio:mp:evs | EVS Media Profile | TS 26.117, clause 7.4.2.3 | Thomas Stockhammertsto@qti.qualcomm.com | none |
| urn:3GPP:audio:mp:ivas | IVAS Media Profile | TS 26.117 clause 7.5.2.3 | tbd | none |
| urn:3GPP:audio:mp:eAAC+ | eAAC+ stereo Media Profile | TS 26.117, clause 7.6.2.3 | Thomas Stockhammertsto@qti.qualcomm.com | none |
| urn:3GPP:audio:mp:amr-wb+ | AMR-WB+ Media Profile | TS 26.117, clause 7.7.2.3 | Thomas Stockhammertsto@qti.qualcomm.com | none |
| urn:3GPP:audio:mp:xHE-AAC | xHE-AAC Media Profile | TS 26.117, clause 7.8.3 | Frédéric Gabinfrederic.gabin@dolby.com | none |

\* \* \* End of Changes \* \* \* \*