TSG CT/SA/RAN Meeting #XX S4-240402

DD - DD Mon. YYYY, City, Country

**Source: Huawei, Qualcomm**

**Title: Rel-18 Work Item Exception for IVAS\_Codec**

**Document for: Approval**

**Agenda Item:**

3GPP™ Work Item Exception

# Title : Rel-18 Work Item Exception for IVAS\_Codec

## Acronym : IVAS\_Codec

## Unique Identifier : 770024

**Release 18 Submission form**

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Feature / Item:** | **IVAS\_Codec** | | | | |
| **Affects:** | **UICC apps:**  **No** | **ME:**  **Yes** | **AN:**  **No** | **CN:**  **Yes** | **Others (specify):** |
| **Expected Completion Date:** | SA4#128 - May 2024 (SA#104 June 2024) | | | | |
| **Service(s) impacted:** | IVAS | | | | |
| **Specification(s) affected:** | TS 26.251, TR 26.997 | | | | |
| **Task(s) within work which are not complete:** | Finalize fixed-point C-code development, and carry out characterization tests to complete the specification and report. | | | | |
| **Consequences if not included in Release 18:** | No immersive audio service | | | | |

**Abstract of document:**

The overall objective of the IVAS\_Codec work item is to develop a single general-purpose audio codec for immersive 4G and 5G services and applications including the XR use cases envisioned in 3GPP TRs 26.918 and 26.928 and possibly relying on devices described in 26.998. In order to achieve such versatile goals, the selected IVAS codec is a collection of low-delay speech and audio coding techniques and rendering targeting services with interactive stereo or immersive audio communication. It comprises an encoder, a decoder, a renderer and several auxiliary functions associated with the support of stereo and immersive audio formats beyond EVS mono coding. It follows all requirements set forth by 3GPP, which include:

* The IVAS codec is an extension of the 3GPP Enhanced Voice Services (EVS) codec; it provides full and bit exact EVS codec functionality for mono speech/audio signal input.
* Encoding and decoding of stereo and immersive audio formats such as multi-channel audio, scene-based audio (Ambisonics), metadata-assisted spatial audio (MASA), object-based audio (ISM).
* VAD/DTX/CNG for rate efficient stereo and immersive conversational voice transmissions.
* Error concealment mechanisms to combat the effects of transmission errors and lost packets. Jitter buffer management is also provided.
* The IVAS codec operates on 20-ms audio frames. In addition, rendering is possible with 5ms granularity.
* Support for bit rate switching upon command.
* Stereo and immersive audio coding at the following discrete bit rates [kbps]: 13.2, 16.4, 24.4, 32, 48, 64, 80, 96, 128, 160, 192, 256, 384, and 512.

The encoder of IVAS expects mono, stereo, objects, multichannel, ambisonics, MASA, combination of objects and MASA, or combination of objects and SBA as input audio channels. In case of objects or MASA, also input metadata are expected. The encoder analyzes the scene, derives the spatial audio parameters, and downmixes the input channels to so called transport channels which are subsequently processed by the encoding tools. These tools comprise Single Channel Elements (SCE comprising one core coder), Channel Pair Elements (CPE comprising two core-coders), and Multichannel Coding Tool (MCT comprising a joint coding of multiple core-coders) while core-coder is inherited from the EVS codec with additional flexibility and variable bitrate.

Given the scope of IVAS\_Codec is significantly more extensive than any other speech/audio codecs developed by 3GPP in the past, the conversion of the selected floating-point C-code to fixed-point C-code is also much more challenging and would require more time to be completed. Therefore this exception is requested.

**Contentious Issues:**

None