**3GPP TSG SA#104 SP-xxxxxx**

**Budapest, June 2022**

**Source: 3GPP SA4**

**Title: EVS Codec Extension for Immersive Voice and Audio Services Ph2**

**Document for: Approval**

**SA Agenda Item: x.x**

3GPP WG-SA4 Meeting #128 S4-24xxxx

Jeju, Korea, 20 – 24 May, 2024

**Source: Dolby Laboratories Inc., Ericsson LM, Fraunhofer IIS, Huawei Technologies Co Ltd., Nokia Corporation, NTT, Orange, Panasonic Corporation, Philips International B.V., Qualcomm Incorporated, VoiceAge Corporation, Xiaomi**

**Title: WID on EVS Codec Extension for Immersive Voice and Audio Services Phase 2**

**Document for: Approval**

**Agenda Item: x.x**

3GPP™ Work Item Description

For guidance, see [3GPP Working Procedures](http://www.3gpp.org/About/WP.htm), article 39; and [3GPP TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm).
Comprehensive instructions can be found at <http://www.3gpp.org/Work-Items>

# Title: WID on **EVS Codec Extension for Immersive Voice and Audio Services Phase 2**

## Acronym: IVAS\_Codec\_Ph2

## Unique identifier: 770024

## 1 Impacts

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| **Affects:** | UICC apps | ME | AN | CN | Others (specify) |
| **Yes** |  | X |  | X |  |
| **No** | X |  | X |  |  |
| **Don't know** |  |  |  |  |  |

## 2 Classification of the Work Item and linked work items

### 2.1 Primary classification

This work item is a …

|  |  |
| --- | --- |
| x | Feature |
|  | Building Block |
|  | *Work Task* |
|  | Study Item |

### 2.3 Other related Work Items and dependencies

|  |
| --- |
| Other related Work Items (if any) |
| Unique ID | Title | Nature of relationship |
| 770024 | IVAS\_CODEC | *Same functionality codec based on floating point C-code*  |
|  |  |  |
|  |  |  |
| 470030 | EVS\_Codec | *Basis for the Codec* |
| 830005 | ATIAS | *ATIAS interfaces with IVAS on capture and on renderer sides*  |

## 3 Justification

## 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.

## Support optional split rendering (ISAR) that benefits resources constrained lightweight UEs.

## 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. In addition, this conversion task is being handled by a 3rd party, and would require more time to be completed, debugged and verified. Therefore initiating a Phase 2 development for IVAS\_Codec is necessary.

## 4 Objective

The overall objective of this work item is to develop additional features to complement IVAS standard. The following objectives should be achieved with the work item:

* A fixed-point C-code that is expected to deliver the same functionalities and similar performances as the floating point C-code in TS 26.258.
	+ Verified interworking performances between floating point and fixed-point C-codes, i.e. fl-fl, fl-fx, fx-fl, fx-fx all should have comparable performances.
* Conduct the characterization tests on the fixed-point C-code and complement the TR 26.997.
* Complete the fixed point code conformance test procedure and requirements
* Define a tiered grouping allowing easier operating mode negotiation and guarantee of interoperability.
* Extensions or enhancements to RTP payload formats and SDP negotiation.
* Split rendering integration into RTP PF and 26.114 and 26.119

## 5 Expected Output and Time scale

|  |
| --- |
| **New specifications** *{One line per specification. Create/delete lines as needed}* |
| Type  | Series | Title | For info at TSG#  | For approval at TSG# | Remarks |
|  |  |  |  |  |  |
| TS | 26.251 | Codec for Immersive Voice and Audio Services - ANSI C code (fixed-point) | SA#107 (March 2025) | SA#108 (June 2025) |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
|  |  |  |  |  |  |
| CR | 26.997 | IVAS Codec Performance Characterization | SA#107 (March 2025) | SA#108 (June 2025) |  |
|  |  |  |  |  |  |

|  |
| --- |
| **Impacted existing TS/TR** *{One line per specification. Create/delete lines as needed}* |
| TS/TR No. | Description of change  | Target completion plenary# |
| *26.114* | *Additional support for the IVAS Codec* | SA#108 (June 2025) |
| *26.119* | *Additional support for the IVAS Codec* | SA#108 (June 2025) |

## 6 Work item Rapporteur(s)

Su, Huan-yu, Huawei Technologies Co Ltd.,su.huanyu@huawei.com

Varga, Imre, Qualcomm Incorporated,ivarga@qti.qualcomm.com

## 7 Work item leadership

*SA4*

## 8 Aspects that involve other WGs

*None*

## 9 Supporting Individual Members

|  |
| --- |
| Supporting IM name |
| Ericsson LM |
| Qualcomm Incorporated |
| Huawei Technologies Co Ltd |
| LG Electronics Inc. |
| Dolby Laboratories Inc. |
| Nokia Corporation |
| Fraunhofer IIS |
| VoiceAge Corporation |
| Orange |
| Samsung Electronics Co., Ltd |
| ZTE Corporation |
| Philips International B.V. |
| Xiaomi |
| Panasonic |
| NTT |