**Source: Dolby Sweden AB, Fraunhofer IIS, Nokia**

**Title: Proposal on IVAS levels**

**Agenda item: 7.5**

**Document for: Discussion and Agreement**

# **Introduction**

During SA4 #128 meeting held at Jeju, Korea, several contributions on IVAS levels negotiation were noted.

[1] defines a core set and further proposes that a UE capable of offering more than the core set features should explicitly signal to the other UE the additional features.

[2] defines levels based on bitrate and defines separate expectations for encoder and decoder. Furthermore, it allows for more flexibility on decoder such that a bitstream of any level can be decoded to provide basic experience.

[3] defines levels based on UE capabilities and further proposes UEs exchanging the list of operating points and bitrates that are supported and then selecting a configuration that enables highest level of immersive experience.

The source considered all the inputs on levels and proposes a bitrate-based level negotiation approach.

# **Discussion**

IVAS is a major upgrade on top of EVS and unlocks very high-quality immersive audio experience by supporting a wide range of coded formats and bitrates. However, supporting all features of IVAS can be a challenging on less capable devices. Moreover, it may require major architectural changes which might discourage the implementers and may potentially delay the deployment of IVAS. To facilitate the deployment of IVAS, it is critical that IVAS levels are defined such that the negotiations are simple and can be easily implemented on top of EVS.

Furthermore, in the current architecture, device configurations (e.g., capture and playback format) may not be available to modem processor and hence the negotiation of level may solely depend on clock/processing power available. IVAS complexity scales with bitrate (given the number of transport channel increase with bitrate) as also suggested in [2] and hence it is desired to define levels solely based on bitrate.

# **IVAS Complexity**

A relation between IVAS complexity and levels was defined in IVAS-4 design constraints [5] :

|  |
| --- |
| Complexity/memory limits are defined in levels. The following level-dependent limits apply for IVAS codec operations (encoder/decoder/renderer total) excluding JBM and other supplementary operations:* Level 1 (if supported):
* Complexity <= 3 \* EVS
* RAM <= 3 \* EVS
* Level 2 (if supported):
* Complexity <= 6 \* EVS
* RAM <= 6 \* EVS
* Level 3:
* Complexity <= 10 \* EVS
* RAM <= 10 \* EVS

Full functionality shall be provided at the highest level. The support of the lower levels with reduced functionality is recommended.In addition, the EVS interoperability mode should not require substantially increased complexity or memory compared to standard EVS.The following level-independent ROM and PROM constraints apply:* ROM, PROM <= 10 \* EVS

The complexity/memory shall be evaluated using the WMC automated tool based on ITU-T G.191 for both CuT and reference in a consistent way for worst case. To account for measurement inaccuracies, the limits must not be exceeded with a tolerance of 10%. Complexity level shall be provided to encoder / decoder / renderer during codec initialization.The decoder/renderer at all levels shall be able to decode any IVAS bitstream and render it to an output format that may be level dependent.As part of the selection deliverables, proponents shall provide a detailed documentation how and with which specific operation modes their IVAS candidate meets the complexity constraints of the different levels. |

A detailed report on IVAS floating point code complexity at various formats was generated by Fraunhofer IIS in March 2024, [20240323\_Complexity\_Overview\_v1.pdf](https://forge.3gpp.org/rep/ivas-codec-pc/ivas-codec/-/wikis/uploads/8057f6f4dc4e009def89f7923d52e6d8/20240323_Complexity_Overview_v1.pdf). More recent complexity numbers on the latest IVAS floating point code are present here: [ivas-codec-pc.3gpp.org/ivas-codec/](https://ivas-codec-pc.3gpp.org/ivas-codec/). These reports show how complexity scales with bitrate and it can be seen that upto 80 kbps, for a variety of input formats, the encoder complexity does not depend on number of input channels, for e.g., the encoder complexity numbers are very similar for FOA, HOA2, HOA3 and OSBA upto 80 kbps. Furthermore, it can also be seen that for a given coded format, the decoder complexity depends on bitrate and output format and while some output formats have higher complexity than others, it might be desired to not put major limitations on the output format in the level definition and give the flexibility to choose the output format as per receiver’s capability.

# **Level definition**

Based on the above mentioned considerations, the source proposes following definition of IVAS levels:

* Define levels based on bitrate as follows.
	+ Level 1: All bitrates upto 80 kbps, all formats
	+ Level 2: All bitrates upto 192 kbps, all formats
	+ Level 3: All bitrates upto 512 kbps, all formats

Requirements

* A compliant IVAS encoder of a given level shall support all IVAS coding modes/coded formats at all bit rates of that level.
* A compliant IVAS decoder of a given level shall support decoding of all IVAS coding modes/coded formats at all bit rates of that level to an immersive output format and shall support:
	+ Integrated rendering supporting all the output formats as specified in Table 4.4-1 of TS 26.253, and
	+ Pass-through operation as external (EXT) processing output (as described in clause 4.3 of TS 26.253) for further consumption by an external renderer
* A compliant IVAS decoder of any level shall support decoding of any IVAS bitstream.

NOTE 1: Immersive output format implies any output format other than mono.

NOTE 2: While the definition of IVAS levels ensures support for immersive output formats at the decoder output, specific device requirements to ensure immersive audio presentation are present in other specifications.

NOTE 3: Compliant IVAS encoder means that an encoder implementation meets the encoder compliance criteria defined in TS 26.250. Compliant IVAS decoder means that a decoder implementation meets the decoder compliance criteria defined in TS 26.250.

# **Conclusion**

A proposal on IVAS levels was presented. It follows the complexity level definition in [5] and provides a rather simple definition of IVAS levels that is only based on operating bitrate while following the onion principal. The proposal enables the implementers to support a variety of capture formats with the same device and get an immersive experience even at the lowest level. It further enables a decoder implementation that is interoperable and allows for flexibility on the output format based on receiver’s capabilities.

It is proposed to agree to this level definition. Corresponding CR to the IVAS general overview specification [4] is provided in [6].

References

1. Tdoc S4-240981: “Proposal for IVAS service options and operating mode negotiation”, Huawei Technologies Co., Ltd.
2. Tdoc S4-241024: “Definition of IVAS complexity levels”, Fraunhofer IIS
3. Tdoc S4-241071: “On IVAS Levels”, Nokia Corporation
4. 3GPP TS 26.250: “Codec for Immersive Voice and Audio Services (IVAS); General overview”
5. Tdoc S4-221619: “IVAS Design Constraints (IVAS-4)”, v1.0.0
6. Tdoc S4-241838: “Change Request to TS 26.250 on IVAS codec levels”