3GPP TSG-SA4 Meeting S4-211529

10th-19th November 2021

**Agenda item** 11.5

**Source:** Tencent

**Title:** Audio mixing of multiple streams in ITT4RT

**Document for** Agreement

# Introduction

The previous version of this contribution was discussed during SA#115. In that meeting two solutions were provided:

1. The use of RTCP header extensions for signaling audio gains, which was agreed at that meeting.
2. The use of SDP for signaling audio again for which questions were raised about the management of the SDP bandwidth during the session if this option is used.

With further investigation, we concluded that even the SDP signalling of the audio gain updates are occasional (e.g. at most every 5 minutes on average), it is difficult to characterize their impact on the additional SDP bandwidth. Therefore, we removed this option from this contribution.

Therefore, we propose the RTP header extension solution to be included in a separate dCR. As we discussed in the previous meeting, if and when IVAS provides a solution for signalling the audio gains, then this solution is not needed and the dCR doesn’t need to go to CR. Otherwise, we propose the dCR to be submitted as a Release 17 CR.

# Y.X Audio mixing

When multiple audio streams are transmitted from the ITT4RT-Tx client, the ITT4RT-Rx client decodes and downmixes these audio streams before rendering. The recommended mixing gain of different audio streams is defined by the ITT4RT-Tx client and updated during the session. This setting is used by the ITT4RT-Rx client to mix the audio streams if it chooses to accept ITT4RT-Tx’s recommendation. The ITT4RT-Rx client can overwrite this setting at any time.

For downmixing multiple audio streams, the ITT4RT-Tx client sends the recommended mixing gain for each audio source of that sender: r0, r1, .., rN for the 360 video (a0) and overlay videos a1, a2, .., aN respectively and therefore recommend the mix of r0\*a0+r1\*a1+……+rn\*an.

The ITT4RT-Tx client indicates the audio mixing gain value to the ITT4RT-Rx client via RTP header-extension based (see Y.X.1).

## Y.X.1 RTP based solution

For RTP based solution, the packet of each RTP audio stream can indicate, in an RTP header extension, the mixing level of that audio sample, carried in the RTP packet with the following RTP extension.

0 1

0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

| ID | len=0 |audio-mixing-gain|

+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+

Audio mixing gain using One-Byte Header Format

The 4-bit ID is the local identifier of this element. The 4-bit length is the number, minus one, of data bytes of this header extension element following the one-byte header. The URI for declaring the audio mixing gain header extension in a Session Description Protocol (SDP) extmap attribute [95] and mapping it to a local extension header identifier is

urn:3gpp:audio-mixing-gain

The header extension is present only in the first/ first few packets of the RTP audio stream and is repeated when the mixing gain needs to be updated for optimality. The audio mixing gain is expressed in dB via signed integer with values between "-127" and "0" and linear weighting between the extreme ends (hence the numerical values directly represent the gain in dB). The -128 value is used for muting the channel. The meaning of positive values other than 0 is undefined and shall be ignored if received.

# O.6 urn:3gpp:audio-mixing-gain

The desired extension naming URI:

urn:3gpp:audio-mixing-gain

A formal reference to the publicly available specification:

3GPP TS 26.114

A short phrase describing the function of the extension:

Signalling of the audio mixing gain header extension for the sent audio, see clause Y.X.1

Contact information for the organization or person making the registration

3GPP Specifications Manager

3gppContact@etsi.org

+33 (0)492944200

# Proposal

The proposal is to include the above section (Y.X.1 and O.6 ) into a separate dCR from the current ITT4RT dCR.