**3GPP TSG-SA WG4 RTC SWG Ad-hoc Post SA4#133-eS4aR250167**

**Paris, France, 3 – 5 August 2025** *(revision of S4aR250155)*

**Source: InterDigital Canada**

**Title: [AvCall-MED] Media Configuration for Avatars**

**Agenda item: 7**

**Document for: Agreement**

# Introduction

In TS 26.264, media configuration requirements for AR-MTSI clients in terminals are defined in clause 7. This clause currently only describes network media rendering configuration. In order to establish an avatar-based IMS call, an AR-MTSI client in terminal needs to be capable of signalling support for AR Avatar calls.

This contribution proposes changes to clause 7 in TS 26.264 to support media configuration for Avatars.

# Proposed Changes

\* \* \* Begin Change \* \* \*

7 Media configurations

7.1 General

The media configuration requirements for MTSI clients in terminals specified in TS 26.114 [2], clause 6, also apply for AR-MTSI client in terminal.

An SDP framework for AR data exchange for AR communication is presented to negotiate codec support for AR media, AR metadata, as well as RTP/RTCP signalling necessary for AR media rendering processing.

AR-MTSI client in terminal shall use RTP for the real-time transport of AR media for AR communication. Any AR media as an overlay may refer to the overlay configuration described in clause Y.6.4.3 of TS 26.114 [2]*.*

AR-MTSI client in terminal shall use data channels for exchange of AR metadata and rendering negotiation. The SDP attribute “*3gpp\_armetadata\_types*”should be used to indicate the types of AR metadata which defined in clause 6 (e.g. pose, action and scene description) within the data channel.

The syntax for the SDP attribute is:

att-field = "3gpp\_armetadata\_types"

att-value = message\_type \*("," message\_type)

message\_type = char-val

; URN identifying the message type of AR metadata

; char-val is defined in RFC 7405

Poses as part of AR metadata may be transmitted via RTP session in a RTP header extension as specified in clause 4.3 of TS 26.522 [8].

The data channel requirements for an MTSI client in terminal specified in TS 26.114 [2], clause 6.210, also apply for AR-MTSI client in terminal that supports Avatar communication.

AR-MTSI client in terminal that supports Avatar communication shall use application data channel for transport of animation data.

7.2 Network media rendering configuration

The AR-MTSI client in terminal shall indicate its support for AR calls by including the “webrtc-datachannel” in the “+sip.sub-type” Contact header field.

A new Contact header field parameter, “+sip.3gpp-ar-support” is used to indicate the level of support for AR calls. The possible values for the “3gpp-ar-support” parameter are:

- “**ar-capable**”: indicates that the terminal is fully capable of receiving and rendering AR media as described by the capabilities in [2] clause 9.2.

- “**ar-assisted**”: indicates that the terminal is capable of transmitting AR metadata on the uplink. However, the UE has no support for processing and rendering a 3D scene. The participation in an AR call requires the deployment of network rendering. The rendered view(s) are controlled by the pose information that is shared by the terminal.

In the absence of the “+sip.3gpp-ar-support”, it shall be assumed that the terminal has no support for AR calls. In this case, the MTSI client can only participate in an AR call if network rendering is offered.

An AR-MTSI terminal that intends to participate in an AR call shall register with the “**ar-capable**” value for the “+sip.3gpp-ar-support” parameter and shall offer/answer an SDP that includes a data channel with the sub-protocol “mpeg-sd”. Any updates that the AR-MTSI terminal intends to share, including pose updates, will be sent as scene updates to the AR AS. An AR-MTSI terminal that intends to participate in an AR call with the support for network rendering shall register with the “**ar-assisted**” value for the “+sip.3gpp-ar-support” parameter and shall offer/answer an SDP that includes a data channel with the sub-protocol “3gpp-sr-metadata” as defined in [6]. Pose updates that are to be used for the rendering are shared as pose predictions with the MF.

As specified in Annex AC.9 of TS 23.228 [4], the AR application server may provide network assisted rendering. An AR-MTSI client in terminal can decide to request network media rendering based on user selection and its status such as power, signal, computing power, internal storage, etc. The AR-MTSI client in terminal shall complete an AR media rendering negotiation with the AR AS before it initiates subsequent procedures to activate the network media rendering. The data channel should be established for rendering negotiation with SDP offer/answer between AR-MTSI client in terminal and MF with the sub-protocol “3gpp-sr-conf”, and continue to be used for rendering re-negotiation until the end of the AR communication.

An **AR-assisted** terminal that intends to deploy network rendering for AR media rendering, shall use the negotiation processes between the AR-MTSI client in terminal and the AR AS to determine the split rendering configuration. The split rendering configuration shall be in JSON format as specified in clause 8.4.2 of TS 26.565 [6]. The exchange of the configuration information shall take place using the established application data channel. The split rendering configuration message shall be formatted according to clause 8.4.2.2 of TS 26.565 [6] and shall have the type: “**urn:3gpp:split-rendering:v1:configuration**”. The output description message shall be formatted according to clause C.1.4 of TS 26.565 [6] and shall have the type: “**urn:3gpp:split-rendering:v1:output**”.

For a terminal that does not support AR calls, the IMS AS may trigger network rendering on behalf of the terminal upon receiving an (re)INVITE for an AR call. The output format for the rendered media shall be conformant to clauses 10.4.3 and 10.4.4 of TS 26.119 [3]. The MF that performs the remote rendering shall select a suitable rendering viewpoint for the session, e.g. a selected viewpoint in the scene or the initial viewpoint for the participant as assigned by the AR AS in the scene description. In case no network rendering can be setup, the IMS AS should reject the call.

7.3 Avatar animation and rendering configuration

7.3.1 Avatar capability configuration

The AR-MTSI client in terminal shall indicate its support for avatar calls by including the “webrtc-datachannel” in the “+sip.sub-type” Contact header field.

A new Contact header field parameter, “+sip.3gpp-avatar-support” is used to indicate the level of support for Avatar calls. The possible values for the “3gpp-avatar-support” parameter are:

- “**avatar-capable**”: indicates that the terminal is fully capable of receiving, animating and rendering of avatar.

- “**avatar-assisted**”: indicates that the UE has no support for animating or rendering an avatar

NOTE: The SIP register message for avatar calls is FFS. It may be needed for indicating which UEs require avatar assistance and which ones are avatar capable.

7.3.2 Network Animation and Rendering

When an AR-MTSI client initiates or receives a call with avatar media, the IMS Application Server (IMS AS) shall evaluate whether network-based avatar animation and rendering is required. The IMS AS forwards avatar-related INVITE requests to the Avatar-capable AR AS for capability assessment and media routing decisions.

The AR AS shall invoke network-based animation and rendering through an MF when:

* the receiving MTSI client has not registered the "+sip.3gpp.avatar-support" feature tag,
* the receiving MTSI client has registered the "+sip.3gpp.avatar-support" feature tag with “avatar-assisted” value,
* the receiving MTSI client's registered capabilities indicate insufficient resources for avatar animation,
* the offered animation frameworks are not supported by the receiving MTSI client.

When network animation and rendering is invoked, the sending AR-MTSI client shall use the negotiation processes with the AR AS to determine the avatar animation and rendering configuration. The exchange of the configuration information shall take place using an established application data channel. The application data channel established for animation and rendering negotiation of the avatar with SDP offer/answer between the sending AR-MTSI client in terminal and the MF with the sub-protocol “3gpp-avatar-negotiation” and continue to be used for animation and rendering re-negotiation until the end of the Avatar communication. In this case, the MF shall establish a video stream with the receiving MTSI client using standard video codecs as specified in [2].

When network animation and rendering is invoked, the AR AS shall allocate an MF capable of real-time avatar rendering and configure it with the appropriate rendering parameters based on the receiving UE's video capabilities. The IMS AS shall modify the SDP to route avatar animation data to the MF instead of the receiving UE, effectively inserting the MF into the media path between the sending and receiving MTSI clients.

The MF performing network-based avatar animation and rendering shall fetch the ARF container from the BAR using the reference provided in the SDP and load the avatar model for rendering. The MF shall receive animation streams through the data channel from the sending AR-MTSI client and apply these animation samples to the avatar model in real-time. The animated avatar shall be rendered as a 2D video stream or stereoscopic 3D video stream based on the receiving UE's capabilities.

For the receiving MTSI client, the avatar data channel media description shall be replaced with a video media description including standard video codec negotiation as specified in [2], and avatar-specific attributes that are not applicable to video streams shall be removed.

The MF should ensure lip-sync between avatar animation and associated audio streams.

\* \* \* End of Change \* \* \*

# Proposal

It is proposed to update the draft base CR with the changes described in section 2 of this contribution.