**Source: Qualcomm Inc., Lenovo**

**Title: Real-time metadata transport over RTP**

**Agenda Item: 10.8**

**Document for: Discussion and Agreement**

# Introduction

At the 3GPP-SA4#120-e, the Permanent Document for MeCAR v3.0.0 [1] was approved. The Permanent Document specifies AR/MR media type definition and real-time characteristics. The 5G\_RTP Permanent Document v0.0.2 [2] also specifies the requirements tentatively, including reliability, traffic and direction of AR media types.

Based on the metadata definition and the description specified in MeCAR PD, the AR/MR metadata may be classified into 3 categories.

1. **Device capability**: the metadata specifying the device capabilities and features such as camera information and projection information. The information is usually available at the beginning of an AR/MR session and may not change within an AR/MR session.
2. **Media description**: the metadata describing the space or media object content such as scene description, spatial description, and 3D visual mode. The data size may be large (>10MB) and the information usually does not change frequently and may be event driven.
3. **Interaction**: the metadata representing the user interaction such as FOV, user pose, viewport, gesture, body action, facial expression and AR anchor point. The interaction may trigger the event at the receiver side and the sender may expect low-latency response (50~1000ms). The interaction data size may be small, but the frequency could be high (>1KHz). The data transport may be continuously or may require burst transport. Depending on the applications, the interaction metadata may be synchronized to each other, or synchronized to other media streams, and the reliability may be strict or not strict.

# Real-time interaction metadata transport over RTP

For the applications that the media streams are exchanged using RTP, the interaction metadata may be carried in RTP header extensions [4] given the data size and low-latency requirements.

## RTP and RTP extension

RTP fixed header is specified in [3] and the format is shown in Figure 1.

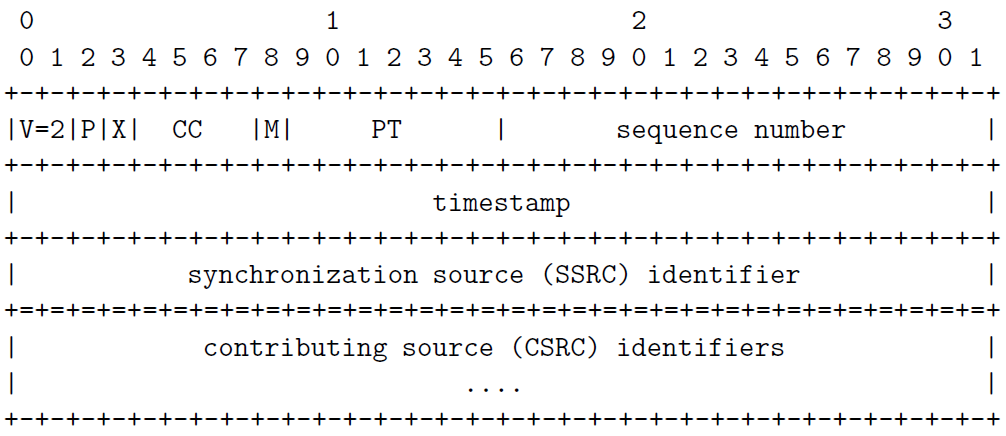


Figure 1 RTP fixed header format

When the extension bit (X) is set, a variable-length header extension must be appended to the RTP header, following the CSRC list if present. Figure 2 shows the general extension format. The RTP header extension may carry metadata in addition to the usual RTP header information as an optimization to lower latency.

Table

Description automatically generated with low confidence

Figure 2 RTP header extension

Two types of extension designs are specified in [4], one-byte header and two-byte header form of extension.

In the one-byte header form of extension, the 16-bit “defined by profile” must have the fixed bit pattern 0xBEDE. Each extension element MUST start with a byte containing an ID and a length. The 4-bit ID is the local identifier of this element in the range 1-14 inclusive. The 4-bit length is the number, minus one, of data bytes of this header extension element following the one-byte header. Figure 3 is an example of one-byte header extension.

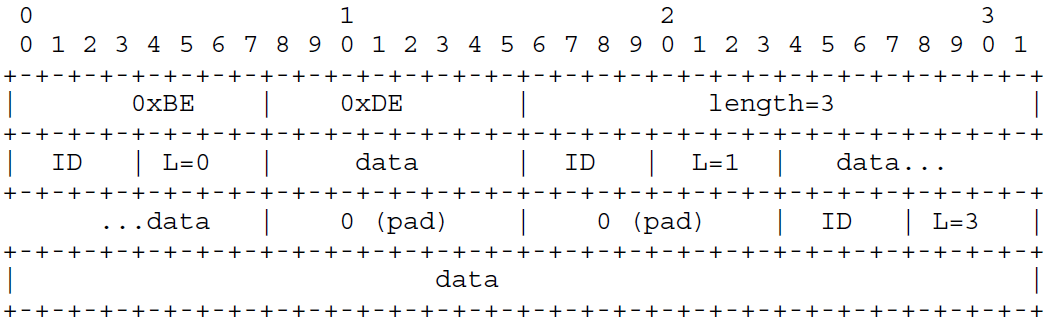


Figure 3 One-byte header extension example

In the two-byte header form of extension, the 16-bit “defined by profile” has 12-bit 0x100 and 4-bit appbits. The appbits field is 4 bits that are application dependent and may be defined to be any value or meaning. Each extension element starts with a byte containing an ID and a byte containing a length. The 8-bit length field is the length of extension data in bytes, not including the ID and length fields. The value zero (0) indicates that there is no subsequent data. Figure 4 is an example of two-byte header extension.

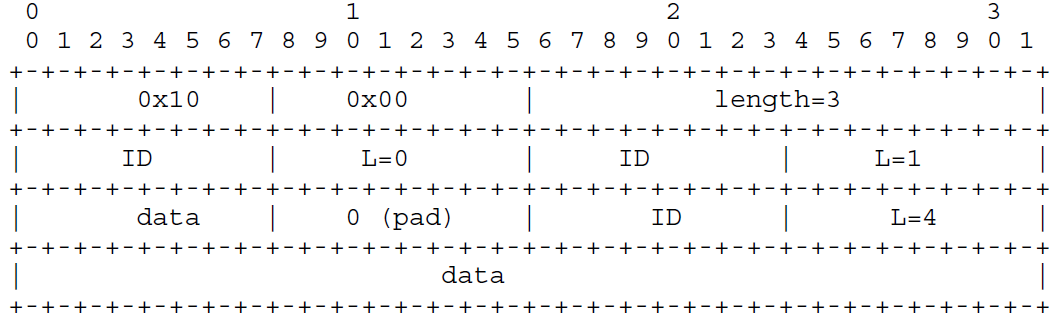


Figure 4 Two-byte header extension example

## Potential Solution for using RTP header extension for interaction metadata

The RTP header extension may be used to carry metadata while the media content is carried in the RTP payload data. Depending on the data length, either one-byte or two-byte header extension may be used; a single metadata type or multiple metadata types may be carried in the extension in a RTP packet.

[Editor’s Note: carrying interaction metadata over the RTP header extension using this approach may be useful but further study of use cases and other types of interaction metadata will be investigated to identify other potential solutions]

Figure 5 illustrates an RTP header extension design concept to carry the real-time interaction metadata.

Subprotocol payload ID is a fixed length field indicating the subprotocol or specifications used for the metadata format, such as OpenXR.

Metadata type is a fixed length field indicating the metadata type specified in the sub-protocol.

Metadata attributes is a fixed length field indicating the metadata attributes such as time synchronization.

Metadata length is a fixed length field indicating the length of metadata payload in bytes.

Table

Description automatically generated

Figure 5 RTP header extension for the metadata

[Editor’s note: the data length of Metadata type field will be investigated to accommodate the potential metadata type indication such as URN]

## Security considerations

The interaction category real-time metadata can contain sensitive information tracking the interactions of an end user, e.g., elements of pose, tracking information of palm, hand, or face, as well as controller inputs. Therefore, the integrity and confidentiality of metadata in transit is in some scenarios, depending on the application requirements, necessary.

The transport of the real-time interaction class metadata over RTP header extensions introduced in clause 2.2 can ensure such necessity for integrity and confidentiality using the secure extension protocol of RTP, i.e., SRTP [6] and its extension RFC6904 [7]. SRTP protects the integrity of the RTP extension headers by signing the RTP PDU contents (including any RTP header extensions), whereas RFC6904 ensures confidentiality by encryption of selected RTP header extensions.

# Proposal

It is requested to include clause 2 of this document in the 5G\_RTP permanent document and take them into account in the discussion of 5G\_RTP and related topics.

# References

1. 3GPP TSG SA WG4 S4-221150, “MeCAR Permanent Document v3.0”, August 2022
2. 3GPP TSG SA WG4 S4-221209, “5G\_RTP Permanent Document v0.0.2”, August 2022
3. IETF RFC3550, "RTP: A Transport Protocol for Real-Time Applications", July 2003
4. IETF RFC8285, “A General Mechanism for RTP Header Extensions”, Oct. 2017
5. IETF RFC3711, “The Secure Real-time Transport Protocol (SRTP)”, March 2004
6. IETF RFC6904, “Encryption of Header Extensions in the Secure Real-time Transport Protocol (SRTP)”, April 2013