**3GPP TSG-WG SA4 Meeting #118E e-meeting *S4-220*492**

**Elbonia, April 6th – 14th, 2022**

Title: [DRAFT] LS Reply on QoS support with PDU Set granularity

Response to: LS (S2-2201803/S4-220460)

Release: Rel-18, Rel-17

Work Item: FS\_XRM, FS\_XRTraffic

Source: SA4

To: SA2

Cc: RAN1, RAN2, RAN3

**Contact Person:**

Name: Qi Pan, Thomas Stockhammer (Rapporteur FS\_XRTraffic)

Tel. Number:

E-mail Address: [panqi8@huawei.com](mailto:panqi8@huawei.com), tsto@qti.qualcomm.com

**Send any reply LS to: 3GPP Liaisons Coordinator,** [**mailto:3GPPLiaison@etsi.org**](mailto:3GPPLiaison@etsi.org)

**1. Overall Description:**

3GPP TSG SA WG4 (SA4) would like to thank 3GPP TSG SA WG2 (SA2) on the LS on QoS support with PDU Set granularity. We believe this is an important topic and we are generally very interested in exchanging information with SA2 on this matter.

3GPP SA4 is continuously working on issues related to the LS. In particular, the Rel-17 study on XR Traffic, documented in TR26.926, is addressing some concrete aspects mentioned in the below questions. Unfortunately, the XR Traffic work has been down-prioritized due to some urgent work item completion in SA4. Nevertheless, we expect that the information on TR26.926 v1.1.0 and the accompanied Permanent Document (PD) supports some of the answers below, and more comprehensive information will be available later this year. However, we also note that TR 26.926 was primarily generated to develop typical traffic patterns for RAN1 simulations but may not serve as a universally applicable framework. As an example, TR 26.926 assumes a specific video codec, assumes a subset of applications, and defines trace formats, that includes advanced metadata that may not be available on system interfaces as of today. In particular, the concept of *slices* was introduced as it is the way how H.265/HEVC provides video data in Network Abstraction Layer (NAL) units to the network. A more generic term is application data unit, also introduced in TR26.926 which is independent of the specific codec and media type.

For your questions, SA4 would like to provide the answers as following.

1. ***Any feedback/guidance on the definition of PDU Sets and video slices in the attached document S2-2201851.***

For definition of a PDU Set, SA4 understands the PDU Set represents the set of packets transmitted within the 5G system and defined to facilitate SA2 study on QoS enhancement. Primarily, we view the definition an SA2 concept and SA4 can only comment to the extent if we can possibly map media data to the PDU Set concept.

As one example, SA4 has analyzed to what extent RTP packetized H.265/HEVC video as defined in IETF RFC 7798 may be mapped to the PDU Set concept. RFC7798 provides detailed information on how to map video frames and their corresponding NAL units (the data unit in H.265/HEVC interfacing with the network) to a PDU Set concept. We believe for example, that Fragmentation Unit Packets (defined as a fragment of a NAL unit consisting of an integer number of consecutive octets of that NAL unit, for example a video slice) may be considered as a PDU Set according to the first part of the PDU Set definition (*A PDU Set is composed of one or more PDUs carrying the payload of one unit of information generated at the application level (e.g., a frame or video slice for XRM Services)*, …). In some implementations (note that neither the video codec specifications, nor the IETF RFC, nor 3GPP specifications up to today provide any requirements or recommendation on implementations), the loss of one fragmentation packet of the NAL Unit may result in discarding the entire NAL unit and hence the second part of the PDU definition (*which are of same importance requirement at application layer. All PDUs in a PDU Set are needed by the application layer to use the corresponding unit of information.*) applies. In other implementations, receivers may use the data up to the first lost fragmentation unit to recover at least parts of the video data included in the NAL unit and apply error concealment afterwards. In this case, the third part of the PDU Set definition (*the application layer can still recover parts of the information unit, when some PDUs are missing*) applies, but in this case the equal importance part of the PDU Set definition (*which are of same importance requirement at application layer*) may be misleading (Note that in this operation mode, as an example if the first packet of the PDU Set is lost, all other packets of the fragmentation units are useless, whereas of the last packet is lost, all packets except the last one can be used by the decoder. Obviously, in this example scenario the first packet is “more important” that the last one). In addition, for video, data included in NAL Units is typically spatially and/or temporarily predicted from video data in other NAL Units, so some sort of cross-dependency of NAL units exists. Based on this and other potential scenarios, even mapping multiple NAL units to a single PDU Set may not be excluded to be considered as a viable setup. Generally speaking, packet losses in video applications typically result in some sort of impacted video quality.

Note that H.264/AVC has similar interfaces as H.265/HEVC based on NAL Units and based on IETF RFC6184, the same principles explained above for H.265/HEVC likely also apply for H.264/AVC.

In yet another example, a PDU Set may be mapped to all source and repair packets of an Application Layer FEC source block. Application Layer FEC is for example used in multicast/broadcast (for details refer to 3GPP TS 26.346) or in conversational applications (see TS 26.114). Typically, a for an applicational layer source block packets from 0 to K-1 identify the source symbols of a source block in sequential order, where K is the number of source symbols in the source block. Encoding Symbol IDs K onwards identify repair symbols generated from the source symbols using an FEC encoder, e,g, Raptor. Typically, N >= K packets are sent, each carrying an FEC source or repair symbols. Typically, the decoder requires only any K or only a small amount more than K packet of the N packets to recover the source packets. Based on this, the definition of a PDU Set applies to all packets of a source block (*A PDU Set is composed of one or more PDUs carrying the payload of one unit of information generated at the application level (e.g., a frame or video slice for XRM Services), …*) and any K packets are sufficient to recover, i.e. all packets are of same importance (*which are of same importance requirement at application laye*r). As only K out N are required, the first definition does not hold(*All PDUs in a PDU Set are needed by the application layer to use the corresponding unit of information.*), but more the second (*In some cases, the application layer can still recover parts of the information unit, when some PDUs are missing*). Actually, note that in this example of an Application Layer FEC not not only *parts* of the information unit but the *full* information unit can be recovered.

SA4 is also considering how other examples such as Audio application data units, Video application data units beyond H.265/HEVC, DASH and CMAF Segments, etc. may map to the PDU Set concept, but has not yet completed a full analysis. However, for this analysis not only the conceptual PDU Set definition is relevant, but also potential constraints and requirements on the timing relationship of PDUs within one PDU Set, the bitrate constraints, and also the anticipated content delivery protocol (congestion based, retransmission based, FEC-based, rate adaptation, long-term prediction, etc.). In addition, third-party services and receiver implementations may not exactly follow the specifications in IETF or SA4, and hence, a comprehensive analysis of the SA2’s PDU Set concept may be difficult to achieve.

Based on this analysis, we believe that the current PDU Set definition and the anticipated use for an application may map to and benefit some use cases and operation modes for media delivery. In particular, it should be possible that the media application layer can make use of the PDU Set definition and define the appropriate information units that are mapped to a PDU set. Detailed implementation of such mapping, e.g. signalling of information units to be mapped to PDU Sets from the application to the 5G Systems, should be for further study. For other cases, the current definition and concept around PDU Sets may not fully apply and other QoS frameworks may be consider. In general, SA4 is willing to continuously work with SA2 on this matter.

On the definition of a video slice, no general definition exists, but some video coding specification such as H.264/AVC and H.265/HEVC make use of this concept and map video slices to NAL units. However, the current definition of SA2 in itself may be misleading for example (i) as a slice is typically not fully self-contained lacking parameter sets and/or relying on data sent in earlier video slices for spatial and/or temporal prediction, or (ii) the processing of a video slice in encoder and decoder is application dependent. In general, it seems unnecessary to understand the media codec specific details of video slices from SA2 perspective. SA2 should rather refrain from define video slices as it is not a network concept, but instead established well-defined and generic network-centric concepts. A *slice* as understood by SA4 may only be used an example.

1. ***Any input on traffic characteristics of typical media services that rely on e.g. PDU Sets, as necessary.***

Based on the response to question 1 above, it is clear that the PDU Set concept when map to SA4 information units results in some open questions and neither a comprehensive nor a definitive answer can be provided. In addition, **i**t is difficult, and likely impossible to identify common “traffic characteristics”, since the traffic characteristics heavily depend on the application choices, such the application, the codec in use, the data formats, the encoding operation (bitrate control, usage of slices, error resilience such as intra frames, GDR, or long-term prediction, etc.). In particular, low-latency XR and cloud gaming video services such as Split-Rendering or Cloud Gaming typically would not use the traditional coding structure with a fixed GOP structures. In addition, the field of low-latency video delivery is undergoing heavy innovation and new coding method may be established frequently. Nevertheless, SA4 documents a comprehensive set of encoding and packaging options in TR 26.926 (and the PD) and provide some typical traffic characteristics. We encourage SA2 to use those as a baseline, if specific examples for XR traffic characteristics are needed. However, we encourage SA2 to not imply that the traffic characteristics are universally applicable for all XR applications.

1. ***Clarify what, if any, dependency there is between the IP packets that make up PDU Sets, e.g. a frame/”slice”.***

Again**,** based onthe response to question 1, no single comprehensive answer can be provided. The examples above provide some insight into possible dependencies, from none, to a dependency that the information unit is prefix dependent to the case that if any piece is lost of an information unit, the entire unit may be discarded. We believe that different application media layer mappings and receiver implementations should be addressed by the PDU Set concept and the media/application layer should be able to configure the appropriate handling.

1. ***Clarify what, if any, dependency there is between PDU Sets that may carry, e.g. different frames/”slices”.***

Again, referring to the response in question 1, no single comprehensive answer can be provided. Referring to the example, spatial and or temporal prediction of slices/frames across NAL units in the case of the video coding typically applies. In motion-compensated predicted video decoding, some frames/slices refer to other frames (typically entire frames and not restricted to slices) based on the video encoding configuration, but also based on dynamic operational decisions. As consequence, a PDU Set may “depend” on previously received PDU Sets. However, such dependencies are not necessarily resulting in discarding dependent information units, but the user experience may be degraded. We repeat again, that due to it heavy-compression and spatial-temporal prediction, any packet losses in video result in degradation and hence PDU loss avoidance should be the primary objectives of a video and XR friendly network – unless application layer retransmission or FEC is in use.

1. ***Provide feedback about the detailed traffic characteristics of XR media services in “Annex X (informative): Traffic characteristics of XR and media services” in the attached document S2-2201807.***

SA4 has reviewed the Annex and has some doubts on the usefulness of the it. Please be aware of the following aspects:

* The concept of I, P and B frames is an ancient concept from MPEG-2 video and broadcast and should not be carried forward. In modern video codecs, complex prediction structures are used that take into account application constraints, encoding complexity, latency and dynamic decisions in the encoding. This may result in irregularities, for example based on sequence properties. In particular for low-latency delivery with error resiliency, different flavours of enoding operations exists such as GDR, multi-hypothesis prediction, inloop filters, or long-term prediction, including synchronized reference buffers between the encoder and decoder. For low-latency as needed in XR, classical B-frames would not exist. Hence, clause X.1 is consider of little value.
* On X2 (“Other traffic characteristics”):,
  + having a generic statement on frame periodicity and jitter seems not applicable. Video can be variable framerate, for example there are systems that grab a frame from the camera when the uplink buffer is cleared, a video frame is rendered based on a trigger, etc.. There are surely cases for constant frame rate applications as well. So if anything on X2 would remain, it is relevant that frame periodicity is an application configuration option, as would be a variable framerate application. In addition, on “5GS should take both periodicity and jitter into consideration for XR and media service” we believe it is an incorrect assumption, since typically for XR traffic there are playout buffers that mitigate/cancel the impact of jitter.
  + Implying importance to data units may be misleading. As explained in answer 1 as well above on Annex X.1, the concept of a “frame” having assigned a single type, and each type getting assigned an importance is not taking into account the complex nature of modern video codecs and its operation. The explanation may serve as an example, but surely not as a universally applicable concept.
  + On PDU set signaling, we agree that there should be an agreement between the application and 5GS on PDU set definition and how the signaling is provided.
  + Traffic correlation is highly speculative and there is no evidence on this. We would refrain to say that this is correct

We believe that generally SA2 should focus on defining abstract QoS concepts w/o implying any specifics on video. SA4 would then, as started in this LS, identify if the concepts are applicable to SA4 defined or third-party XR services. We also encourage to review the findings in TR 26.926 which we attempt to complete by SA#96. SA4 will keep SA2 informed if any issues would be identified in the future SA4 study. Findings will be documented in TR 26.926.

**2. Actions:**

**To SA2 group.**

**ACTION:** SA4 asks SA2 group to take the above reply into account and provide feedback if any.

**3. Date of Next TSG SA WG4 Meetings:**

SA4#119-e 11 – 20 May 2022 E-Meeting

SA4#120 22 – 26 August 2022 Malaga