**3GPP TSG-SA WG4 Meeting #128S4-241100**

**South Korea, Jeju, 20 – 24 May 2024**

**Source: Nokia Corporation (Rapporteur)**

**Title: Rapporteur's summary of inputs about LS to SA2 on FS\_XRM\_Ph2 Study Item**

**Agenda item: 10.8**

**Document for: Agreement**

# Introduction

This document is the 5G\_RTP\_Ph2 Rapporteur’s view on the contributions related to the liaison statements from SA2 S4-240873 and S4-240874 of competence of the abovementioned SI.

The idea of this document is to summarize the inputs and facilitate a possible way forward. It is anyway recommended to discuss comprehensively all issues before formulating replies to SA2.

# Summary of the SA2 liaison statements content

# The input liaison statements to this SA4 meeting are the following:

# S4-240873 (on Application Layer FEC)

Solutions have been proposed to provide information about the presence of application layer forward error correction (AL-FEC) to NG-RAN to enable NG-RAN to discard obsolete AL-FEC PDUs. Obsolete AL-FEC PDUs refers to PDUs that are not needed at the UE because enough PDUs to reconstruct the actual content have already been successfully sent to the UE.

Questions to SA4:

* SA2 understands that different AL-FEC mechanisms exist (e.g., maximum-distance separable (MDS) schemes like RaptorQ and Reed-Solomon, FlexFEC, etc.) and is discussing for which AL-FEC mechanisms to enable AL-FEC awareness at RAN. **Can SA4 identify commonly used AL-FEC mechanisms (not necessarily 3GPP defined),** which should be supported for AL-FEC awareness at RAN from SA4's perspective?
* Does SA4 see a need (from a general application perspective) to support both **static and dynamic redundancy** ratios (i.e., the ratio of AL-FEC information) for AL-FEC awareness at RAN?
* Does SA4 see a **need for the application layer to distinguish RAN's intentionally dropped obsolete FEC packets from congestion related drops, and related to this, the need for specific application behaviour, e.g., to reduce the sending rate**? The background to this question is the following:
  + Some companies in SA2 commented that transport protocols or applications need to reduce their sending rate in response to packet losses.
  + Other companies argued that there is no need for reducing the sending rate when NG-RAN discards obsolete AL-FEC PDUs as long as NG-RAN can still meet the QoS characteristics of the other QoS flows in the same cell (i.e., because there is no fairness issue in this case).
* **One solution** (solution #3 in TR 23.700-70) proposed that an application may signal the required content ratio for a PDU Set (i.e., the required ratio of PDUs of a PDU Set needed by the receiver to reconstruct the original content) by first providing a mapping between content ratio levels and PDU Set Importance (PSI) values in the control plane to 5GS and by then using the PSI in the GTP-U header and the mapping received to determine the content ratio per PDU Set at NG-RAN. Does SA4 consider this a feasible option?

# S4-240874 (on FS\_XRM\_Ph2 topics)

Questions to SA4:

* **(1):** PDU Set correlation information (Sol#23) provides the dependency relationship among PDU Sets. **Does SA4, RAN2 and RAN3 see any improvement with adding inter-PDU set correlation information to assist RAN making PDU set discarding decision as comparing to the existing (R18) PDU Set information that is already provided by the AS?**
* **(2):** In Sol#29, PDU Set QoS or ordinary per packet based QoS (e.g. PER, PDB) can be applied for different media streams multiplexed in an IP flow, **SA2 would like to ask SA4 whether a media stream (e.g. a video RTP stream) can include packet which is not related to PDU Set?**
* **(4):** In Sol#30, the PSA UPF may identify the size of incoming burst based on N6 protocol, and send it to NG-RAN to assist RAN scheduling. **is it possible that the application server provides the burst size in the first packet of the burst via N6?**
* **(5):** Some of the solutions support only QUIC-based media delivery. **Can SA4 provide feedback on choosing only solutions for PDU Set identification for encrypted traffic that only support QUIC as transport protocol?**

# Summary of the input contributions

This is the list of contributions submitted at this meeting and the related discussions occurred over the RTC email reflector.

**On Application Layer FEC**

Huawei – S4-240973

Meta – S4-241023

Qualcomm – S4-241086

Lenovo – S4-241088

Ericsson - email from 7 May 2024 (RTC Reflector)

**On FS\_XRM\_Ph2 topics**

Huawei, HiSilicon – S4-240967, S4-240972

Meta – S4-241023

Lenovo – S4-241090

Qualcomm – S4-241092

Bo Burman email from 7 May 2024 (RTC Reflector) \*it is assumed to be Ericsson view but no input contribution from Ericsson was received.

The following table summarizes the company positions for the different topics. The last column represents the 5G RTP Ph2 Rapporteur’s view on a possible RTC recommendation for the reply LS to SA2 and further actions in RTC SWG.

The text marked with \* is a heavy simplification interpretation to facilitate initial discussion, for original proposed view of the companies refer to the indicated document.

|  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- |
|  | **Huawei**  **S4-240973 (AL-FEC)**  **S4-240972 (FS\_XRM\_ph2)** | **Meta**  S4-241023  **(AL-FEC, FS\_XRM\_ph2)** | **Qualcomm**  S4-241086 (AL-FEC)  **S4-241092 (FS\_XRM\_ph2)** | **Lenovo**  **241088 (AL-FEC)**  ***S4-241090 (FS-XRM\_ph2)*** | **Ericsson**  **Attachment A (AL-FEC, FS\_XRM\_ph2)** | **RTC Recommendation for LS replies** |
| **Used AL-FEC\*** | Decision is with SA2 |  | The schemes specified in IETF RFCs should be supported, including RaptorQ, Reed-Solomon, FlexFEC, and ULPFEC. | Active discard in RAN recommended (but not yet studied in SA4), with support of MDS (e.g., RS) or approximate MDS codes (Raptor, RaptorQ) | Feasibility should be studied before thinking what FEC codes could be recommended. | There are two classes of FEC codecs with related implications and challenges. Initial overview is available in TR 26.822 |
| **Static and dynamic FEC redundancy ratios\*** | Both options should be considered. | Dynamic FEC ratio should be supported. | Both options should be considered. | Dynamic FEC ratio should be supported. | Dynamic, but it requires further study | Dynamic FEC ratio may bring advantages. This may be used in combination with RTP retransmission if there is RTT is low enough. |
| **Need to distinguish dropped packets and reduce tx rate\*** | Yes. | Yes.  Dropping should be controlled by the sender. | No and SA4 should study more. | Yes | No. | The sender should possibly authorize the network to intentionally drop packets. Whenever this happen, the network should have a mechanism to inform the sender about what packets are dropped. |
| **Content FEC ratio solution is feasible\*** | No |  | No, there are better solutions. | Needs further study | It would need to be verified | SA4 should study more. |
| **Inter-PDU set correlation info** | * Release 18 TS 26.522 contains guidelines that assign PDU Set Importance implicitly taking dependency between PDU Sets into account. Complex and explicitly inter-PDU Set dependency information may not be practically feasible due to the complexity of codecs and the dependency relations. The late independent PDU Set with many dependencies can still be helpful for decoding of subsequent dependent PDU Sets. | 1. PDU set dependency discarding and b) independent PDU indicators should be supported. 2. To alleviate the potential energy impact due to PDU set discarding by RAN during congestion, it is proposed that PDU set dependency discarding feature should be an optional indicator that can be set by the server to control whether dependent PDU set can be discarded by RAN also if previous PDU set has be discarded due to congestion. In addition, as already documented in the SA2’s TR, an “independent indicator" can be provided to indicate independent PDU set (e.g. IDR frame) to allow RAN to attempt delivery of this PDU set even when PSDB can’t be met. Both should be supported. Overall, Server should have the capability to help RAN on how to discard the PDU set or deliver PDU set to the receiver. | To SA4’s understanding, in Sol#23 of TR 23.700-70 a mapping is as follows: 1) A PDU set is a video frame 2) a dependent PDU set is a video frame that is predicted from previous video frames forming a PDU set as well 3) dependent PDU Sets are dropped if the previous PDU Set is dropped. SA4 discourages such an operation and sees no benefit. The reason is that video decoders are built to compensate packet and video frame losses and with error concealment can provide a significantly better user experience that without any data. Secondly, it is unlikely that dependencies in itself can be expressed with a simple flag, but they are more complex taking into account advanced video codecs that for example use Gradual Decoder Refresh, RTCP Feedback in prediction loop and multi-layer coding. SA4 would not exclude that there may be applications that benefit from such an approach, but the feature should be turned on only if the application can benefit from the feature. | PDU Set interrelations and dependency trees may be quite dense depending on encoding parameters applied. For example, a typical I-P only video sequence may encode a P-frame based on a multiple set of previous frames (e.g., 3, 5, 6, 8, 12 etc.) and same applies for slices. This implies that in general capturing the inter-dependencies of PDU Sets may be convoluted and require many bits on the wire to identify and signal explicitly. Furthermore, capturing inter-dependencies require storing and processing state information of related PDU Sets adding considerable complexity at the RTC sender endpoint and any network node processing such information, e.g., NG-RAN.  It is important to note that SA4 already provided guidelines of PSI marking based on PDU Set dependencies as per TS 26.522 [1], Clause 4.2.6.2.5. This is an implicit inter-dependency marking and may provide some assistance, as required, in RAN PDU Set discarding decisions as per Rel-18. No benefit was so far identified in providing explicit inter-dependencies information of PDU Sets beyond this. | On Question 1, I don’t see the need to add correlation information on top of the PSI field introduced in Rel-18. There is no evidence that adding such correlation information is useful. It is assumed that 'correlation' would be used to decide which and to what extent PDU Sets are used as reference by other PDU Sets and thereby determining how ‘essential’ a PDU Set is to other PDU Sets, which effectively becomes a measure of PDU Set importance (PSI).  Adding additional detailed correlation information, such as for example exactly to which other PDU Sets a PDU Set is correlated, would likely be both complex and voluminous. Making use of such information in RAN also seems to require detailed tracking of the success status of a dynamically updated, long history of previously sent PDU Sets in the same QoS flow, with a size of that history corresponding to the longest PDU Set prediction chain that can easily be tens if not hundreds of PDU Sets. Limiting that complexity to, e.g., the number of dependent PDU Sets for a PDU Set (without detailing which), seems very similar to the information carried by the existing PSI parameter. | There seems to be no benefits, or the topic should be studied more. |
| **A stream can include packets not related to a PDU set** | Yes, it is possible. | - | The answer is yes. This relates to the question whether there are packets that are not a PDU Set ("lonely PDUs"). This was already discussed in LSs (S4-240168) between SA2 and SA4 in Rel-18. | Yes. | Yes. | Yes |
| **Burst size in the 1st packet of the burst** | Yes. SA4 is studying it. | - Yes. It is possible since there may be cases where the application server can add this information without introducing additional delay. SA4 is also studying the enhancement of data burst related traffic characteristics in 5G RTP phase 2, which will be documented in TR 26.822. | There may be cases where the application server can add this information without introducing any sending delay. However, in general it is not always possible to add any such information without adding any delay. In one example, a burst may consist of multiple PDU Sets. If a PDU Set is a video frame, when the video encoder generates the first video frame, it does not know the size of the second video frame. | Any indication of upcoming/future traffic characteristics (e.g., burst size, large PDU Set) involves in general buffering at the RTC sender endpoint as a media encoder typically cannot determine the size of the encoded payload before the encoding finishes. The larger the desired indication scope (i.e., burst, or PDU set) the larger the buffering and delay incurred on the RTC sender endpoint. Some buffering may be possible, but it is specific and up to application end-to-end requirements. In some cases, it may be possible to signal large incoming PDU Sets in the previous PDU Set at the expense of incurring one FPS delay cycle. If RAN finds such approach useful, this could be studied further by SA4 as part of the KI#11 in FS\_5G\_RTP\_Ph2. | On Question 4, I don’t think it would be desirable to always provide a burst size in the first packet of a burst. Doing so would require buffering the entire burst at the sender side before calculating the size of it, causing a potentially significant delay in sending. A media encoder, especially video encoder, cannot know the exact resulting size of an encoded frame before encoding it, as the number of encoded bits in a frame depends on the video content of that frame and cannot be known until the actual encoding is performed. | SA4 should study more. |
| **Solutions only for QUIC-encrypted traffic** | QUIC based protocols are emerging quickly in real implementations and expected to be popularly used in the future. Therefore, it would be fine to focus on the solutions that support QUIC as transport protocol. | - | While QUIC gets more popular, it cannot be assumed to be the only supported protocol for encrypted traffic. In general, SA4 just has started to study applicability of QUIC and more elaborate feedback can only be provided at a later stage. | QUIC has been so far limited in Rel-19 to segmented media delivery only as per approved SIDs of FS\_AMD and FS\_5G\_RTP\_Ph2. Hence commenting on QUIC transport for RTC or generally for packet-based media delivery requires proper study which is currently not in scope of Rel-19. It is however good to understand the layer at which “end-to-end” encrypted traffic is applicable. If UDP SDU is implied there are mechanisms in place, e.g., RFC 6904, which would allow confidentiality and integrity protection of RTP SDU, RTP header extensions, and integrity protection of the RTP header extension for PDU Set marking, as well. | I believe commenting on QUIC-based solutions in SA4 is premature, since QUIC discussions have only just begun in SA4 Rel-19. | SA4 should study more. |

# Attachment A email Bo burman to RTC reflector on May 7 2024

Dear all,

Please find below my comments to the two liaisons that the telco requested should be further discussed on this list before SA4#128.

|  |  |  |
| --- | --- | --- |
| **[S4aR240033](https://www.3gpp.org/ftp/TSG_SA/WG4_CODEC/3GPP_SA4_AHOC_MTGs/SA4_RTC/Docs/S4aR240033.zip)** | LS on Application-Layer FEC Awareness at RAN | ETSI |

I believe this LS is incorrectly assuming the usefulness of AL-FEC, just asking which AL-FEC to use or recommend. I find no proof or even motivation why AL-FEC would be beneficial in cellular 3GPP networks. It even seems possible that, if adding AL-FEC increases the overall flow bitrate on a QoS flow that is already suffering from difficult transport conditions, the resulting source PDU loss rate would increase compared to before adding AL-FEC information to the flow. Shouldn’t AL-FEC overall feasibility be verified before asking how to add it?

If AL-FEC is found to be feasible and under the assumption that adding “too much” AL-FEC could be counter-productive, it seems a dynamic ratio is advisable. How to control that ratio would need further study.

It seems inadvisable for RAN to intentionally drop “obsolete” FEC packets, as any packet losses are to be interpreted as signs of network congestion, which the sender must react to by reducing the sending rate. Thus, if FEC packets continues to be intentionally dropped, the sending rate will also continue to be reduced until the effective throughput is zero or at an absolute minimum. For the sending application reaction to packet drops, it is assumed that applications will generally not be 3GPP-aware and would not know the difference between intentionally dropped packets and congestion-based losses. For such intentional drop solution to work, this assumes a misbehaving application that does not comply with the internet paradigm to reduce the rate as result of packet losses. Hence, I do not think that 3GPP should attempt to introduce any solution to address that scenario.

For the potential use of PSI to indirectly signal the source-to-FEC ratio, I believe the feasibility of this would also have to be verified. For example, if “more important” PDU Sets would get assigned a higher amount of AL-FEC and thus the total size of the PDU Set and its AL-FEC would be larger. This size alone would increase the loss probability for the PDU Set + AL-PDU as the number of PDUs to carry that information increases. That does not seem to necessarily lead to a benefit in the end. Similarly, if less important PDU Sets gets a low amount or no AL-FEC, they might always be dropped, which does not necessarily lead to a better user experience than if no AL-FEC is used at all and traffic is just prioritized in accordance with the PSI field.

|  |  |  |
| --- | --- | --- |
| **[S4aR240034](https://www.3gpp.org/ftp/TSG_SA/WG4_CODEC/3GPP_SA4_AHOC_MTGs/SA4_RTC/Docs/S4aR240034.zip)** | LS on FS\_XRM Ph2 | SA2 |

On Question 1, I don’t see the need to add correlation information on top of the PSI field introduced in Rel-18. There is no evidence that adding such correlation information is useful. It is assumed that 'correlation' would be used to decide which and to what extent PDU Sets are used as reference by other PDU Sets and thereby determining how ‘essential’ a PDU Set is to other PDU Sets, which effectively becomes a measure of PDU Set importance (PSI).

Adding additional detailed correlation information, such as for example exactly to which other PDU Sets a PDU Set is correlated, would likely be both complex and voluminous. Making use of such information in RAN also seems to require detailed tracking of the success status of a dynamically updated, long history of previously sent PDU Sets in the same QoS flow, with a size of that history corresponding to the longest PDU Set prediction chain that can easily be tens if not hundreds of PDU Sets. Limiting that complexity to, e.g., the number of dependent PDU Sets for a PDU Set (without detailing which), seems very similar to the information carried by the existing PSI parameter.

On Question 2, yes, there can be PDUs in a media (e.g., RTP) stream that are not related to any PDU Set. For example, RTP sessions can decide to multiplex RTP and RTCP onto a single QoS flow, in which case the RTCP packets do not belong to a PDU Set. In other cases, even PDUs with different media types can be multiplexed onto a single QoS flow, for example as described by RFC 9143, and not all of those PDUs need be related to any PDU Set.

On Question 4, I don’t think it would be desirable to always provide a burst size in the first packet of a burst. Doing so would require buffering the entire burst at the sender side before calculating the size of it, causing a potentially significant delay in sending. A media encoder, especially video encoder, cannot know the exact resulting size of an encoded frame before encoding it, as the number of encoded bits in a frame depends on the video content of that frame and cannot be known until the actual encoding is performed.

On Question 5, I believe commenting on QUIC-based solutions in SA4 is premature, since QUIC discussions have only just begun in SA4 Rel-19.

Best Regards,

Bo