3GPP TSG SA WG4 128 S4-241086

Jeju, Korea, 20 - 24 May, 2024

Title: LS Reply to SA2 LS on Application-Layer FEC Awareness at RAN

Response to: LS (S2-2405064/[S4-240873](https://www.3gpp.org/ftp/TSG_SA/WG4_CODEC/TSGS4_128_Jeju/Docs/S4-240873.zip)) from SA2

Release: Release 19

Work Item: XRM, 5G\_RTP

Source: SA WG4

To: SA WG2,

Cc: RAN WG2, RAN WG3

**Contact Person:**

Name: Liangping Ma

Tel. Number:

E-mail Address: lpma DOT AT qti DOT qualcomm DOT com

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

Attachments: S2-2405064

**1. Overall Description:**

SA2 is studying enhancements to support for XR and media services. In this context, 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. The details of these proposals are documented as solutions #1, #2, #3, #4 and #21 in TR 23.700-70. In SA2, some companies are of the opinion that such solutions are useful to efficiently handle XR applications, e.g., XR split rendering and cloud gaming services that are using AL-FEC schemes regardless of the access technology that is used for the applications' traffic. Other companies' view is that XR applications should not use AL-FEC over NR in the first place as NR provides efficient means for reliable delivery.

Questions for 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?*

**SA4 Answer:** From a general reconstruction perspective there are two types of AL-FEC: MDS or near-MDS codes, and non-MDS codes. For MDS or near-MDS codes, e.g., RaptorQ and Reed-Solomon, it is sufficient to receive a sufficient number of coded packets, regardless of which packets to reconstruct the original, uncoded information. For non-MDS codes, e.g., FlexFEC, and ULPFEC, the exact dependency between the source packets and the parity packets needs to be known to be able to identify a sufficient set of packets for reconstruction of the original, uncoded information. Even though FlexFEC and ULPFEC are recommended for WebRTC (RFC 8854) and in use of mainstream WebRTC implementation (i.e., libwebrtc), they require more complex signalling overhead to aid RAN to always determine which combination of source and parity packets can recover the original, uncoded information. On the other hand, MDS codes/near-MDS codes only require a redundancy level indication and are not sensitive to the pattern of discarded packets as long as coded packets are received. Based on the above it would be advisable to support MDS or near-MDS codes, yet SA4 did not study in detail signalling aspects until now.

* 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?

**SA4 Answer:**

There are AL-FEC schemes in which the redundancy ratio is fixed based on code construction, and there are AL-FEC schemes in which the redundancy ratio can be dynamically changed, and hence in practice both scenarios exist. Yet, dynamic redundancy ratios may be more beneficial in practice since applications often appeal to dynamic AL-FEC control in combination with RTP retransmission and congestion control mechanisms to ensure robust and efficient adaptation to network conditions.

* *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:*

**SA4 Answer:** For the current congestion control mechanisms of which SA4 are aware, SA4 sees a need for the application layer to distinguish RAN’s intentionally dropped obsolete FEC packets from congestion related drops, and the need, depending on the operating point, for reducing the sending rate. However, if the RAN is enabled by a sender application to intentionally drop obsolete FEC packets, the RAN should indicate the action of intentionally dropping obsolete FEC packets to the application and the application may respond to intentional drops of obsolete FEC packets and congestion related drops differently, without reducing the sending rate.

* *Some companies in SA2 commented that transport protocols or applications need to reduce their sending rate in response to packet losses.*

**SA4 Answer:** Please see the previous answer.

* *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).*

**SA4 Answer:**

* When the network is in congestion, SA4 sees a need for reducing the sending rate in general. However, if discarding obsolete AL-FEC PDUs does not change the statistics of PDU Set losses, SA4 does not see a need for reducing the sending rate.
* SA4 will also study whether it is beneficial for senders for the NG-RAN to discard obsolete AL-FEC PDUs when the network is not in congestion.

Questions for RAN2:

* Can NG-RAN determine whether a PDU was successfully delivered over an unacknowledged mode data bearer? If so, does NG-RAN get this information sufficiently early to decide whether or not to drop subsequent AL-FEC packets?
* Provide feedback on the impact on NG-RAN to support dynamic redundancy ratios, i.e., a different ratio of PDUs that need to be successfully transferred to the UE for different PDU Sets within the same QoS flow?

*Questions for RAN2 and SA4:*

* *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?*

**SA4 Answer:** This solution would change the semantics of the PSI field defined in TS26.522 and this is not advisable upon a first review.., Yet, the general principle of indexing AL-FEC redundancy ratios for signalling optimization may be considered for further study.

**2. Actions:**

**To SA2/RAN2/RAN3:**

**ACTION:** SA4 kindly asks SA2/RAN2/RAN3 to consider the answers above.

**3. Date of Next TSG SA WG2 Meetings:**

<add telco with power to send/receive LS>

TSG-SA2 Meeting #164 19-23 August 2024 Maastricht, NL