**3GPP TSG SA WG4#129-e S4-241521**

**Online, 19th – 23rd Aug 2024 (revision of S4-241229)**

|  |
| --- |
| *CR-Form-v12.0* |
| **CHANGE REQUEST** |
|  |
|  | **26**.**804** | **CR** | **0007** | **rev** | **4** | **Current version:** | **18.1.0** |  |
|  |
| *For* [***HE******LP***](http://www.3gpp.org/3G_Specs/CRs.htm#_blank)*on using this form: comprehensive instructions can be found at* [*http://www.3gpp.org/Change-Requests*](http://www.3gpp.org/Change-Requests)*.* |
|  |

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| ***Proposed change affects:*** | UICC apps |  | ME |  | Radio Access Network |  | Core Network | **X** |

|  |
| --- |
|  |
| ***Title:***  | **[FS\_AMD] Key Issue #X: Improved QoS support for Media Streaming services** |
|  |  |
| ***Source to WG:*** | Huawei, HiSilicon |
| ***Source to TSG:*** | S4 |
|  |  |
| ***Work item code:*** | FS\_AMD |  | ***Date:*** | 2024-05-14 |
|  |  |  |  |  |
| ***Category:*** | **B** |  | ***Release:*** | Rel-19  |
|  | *Use one of the following categories:****F*** *(correction)****A*** *(mirror corresponding to a change in an earlier release)****B*** *(addition of feature),* ***C*** *(functional modification of feature)****D*** *(editorial modification)*Detailed explanations of the above categories canbe found in 3GPP [TR 21.900](http://www.3gpp.org/ftp/Specs/html-info/21900.htm). | *Use one of the following releases:Rel-10 (Release 10)Rel-11 (Release 11)Rel-12 (Release 12)**Rel-13 (Release 13)Rel-14 (Release 14)Rel-15 (Release 15)Rel-16 (Release 16)* *Rel-17 (Release 17)* *Rel-18 (Release 18)* |
|  |  |
| ***Reason for change:*** | As agreed in SP-240514, how to improve the QoS support for Media Streaming services based on the QoS enhancements and the network information exposure is to be studied. Therefore, this paper proposes the Key Issue of "Improved QoS support for Media Streaming services". |
|  |  |
| ***Summary of change:*** | Proposal of KI#X: Improved QoS support for Media Streaming services. |
|  |  |
| ***Consequences if not approved:*** | SI cannot be completed. |
|  |  |
| ***Clauses affected:*** | 5.X |
|  |  |
|  | **Y** | **N** |  |  |
| ***Other specs*** |  |  |  Other core specifications  | TS/TR ... CR ...  |
| ***affected:*** |  |  |  Test specifications | TS/TR ... CR ...  |
| ***(show related CRs)*** |  |  |  O&M Specifications | TS/TR ... CR ...  |
|  |  |
| ***Other comments:*** |  |
|  |  |
| ***This CR's revision history:*** | SA4#128:1. Add clarifications on support of collaboration scenarios.2. Add high-level call flow and corresponding procedure descriptions.SA4#129-e:1. Merge CR0008 as baseline for solutions.
 |

\* \* \* \* First change (all new text)\* \* \* \*

### 5.23.2 Collaboration scenarios

#### 5.23.2.1 General

Collaboration scenarios 2–11 and 13–15 from TS 26.501 [15] are potential points of departure for improved QoS handling support with the following additions:

1. Similar to the Network Assistance feature in TS 26.501 [15], the network status of the 5G System may be exposed to media delivery sessions using the *QoS monitoring* feature and the *ECN marking for L4S* feature. The network status, including the data rate, latency, congestion, etc. may be used by the Media Delivery System for bit rate adaptation and/or congestion control.

 The PDU Set handling feature may be used to label PDUs belonging to a video frame or video slice as members of the same PDU Set.

NOTE: Whether the concept of PDU Set is feasible for video segment in a segment-based streaming service is not clear.

2. In the case of network congestion, the NG-RAN may consider the PDU Set Importance for PDU Set level packet discarding. This is not expected for segment-based devliery where the TCP or QUIC transport connection used to carry the media streaming service requests reliable transmission.

#### Editor’s Note: Whether PDU Set feature is beneficial for Media Streaming services is for future study.5.23.2.1 Collaboration scenarios for L4S ECN marking

Collaboration scenarios for L4S ECN marking are depicted below. Both the Media AS and the 5GMSd Client make use of an L4S-enabled protocol stack. Figure 5.23.2.1-1 assumes that the Media AS resides within the external DN, while figure 5.23.2.1-2 assumes the Media AS within the Trusted DN.



Figure 5.23.2.1-1: Media AS in External DN



Figure 5.23.2.1-1: Media AS in Trusted DN

### 5.23.3 Architecture mapping

Not applicable.

### 5.23.4 High-level call flows

#### 5.23.4.1 Integrating QoS monitoring and/or ECN marking for L4S

The high-level call flow for integrating the QoS monitoring and/or ECN marking for L4S is shown below as well as the corresponding procedures.

It is assumed that the MNO and the 5GMS Application Provider have negotiated a Service Level Agreement that allows the 5GMS Application Provider to enable the ECN marking for L4S and QoS monitoring in the 5G System for media delivery.



Figure 5.23.4.1-1: Potential call flow for improved QoS handling support

Prerequisites:

- The 5GMS Application Provider has agreed an SLA with the Network Operator to allow the usage of network assistance for Media Streaming service.

Steps:

1. The 5GMS Application Provider provisions the 5G Media Streaming System configures content ingest. **A Network Assistance configuration is provided to allow the usage of ECN marking for L4S and/or QoS monitoring to notify the 5GMS Client of the latest network status.**

Editor’s Note: Whether to introduce the feature of "Improved QoS Support" or reuse the "Network Assistance" feature is for further study.

2. When the 5GMS-Aware Application starts, the Media Session Handler retrieves the Service Access Informaiton via M5 or M8. The 5GMS AF address that offers the network assistance is provided in the Service Access Information **and the options for QoS monitoring and/or ECN marking are also present**.

3. The Media Session Handler invokes the **Enhanced** Network Assistance API on the 5GMS AF **to find out about the latest network status. For instance, the 5GMS Media Session Handler may subscribe to the periodic congestion status report from 5GMS AF.**

4. The 5GMS AF interacts with the PCF or NEF to enable QoS monitoring and/or ECN marking for L4S in the 5G System via the Npcf\_PolicyAuthorization service at reference point N5 or the Nnef\_AFsessionWithQoS service at reference point N33.

5. In the case of QoS monitoring, the 5GMS AF can receive the notifications from PCF or NEF via the Npcf\_PolicyAuthorization\_Notify at reference point N5 or the Nnef\_AFsessionWithQoS\_Notify at reference point N33.

6. Alternatively, in the case of QoS monitoring, the 5GMS AF may receive the notifications directly from the UPF via the Nupf\_EventExposure\_Notify at reference point N5 or from NEF the Nnef\_EventExposure\_Notify service at reference point N33. This is beneficial when the 5GMS AF is deployed in the Edge DN and the SMF/PCF is generally deployed centrally.

7. The 5GMS AF further sends the notifications exposed by the network to the Media Session Handler using the MQTT notification channel for the Provisioning Session. The 5GMS Client may take this into account for rate adaptaion, congestion/flow control.

#### 5.23.4.2 QoS monitoring for media streaming



Figure 5.23.4.2-1: High-level call flow for QoS monitoring for Media Streaming

1. 5GMS Application Provider provisions the 5GMS AF with the **Network Assistance configuration** as described in step 1 of clause 5.23.4.1 The **Network Assistance configuration** contains the configuration of QoS monitoring, including the parameters to be monitored, reporting frequency (event triggered, periodic), optionally target entity of reporting and optionally the notification via UPF.

NOTE: In case the 5GMS AS is deployed as an EAS instance in the Edge DN, a local UPF can also be inserted for local access to the 5GMS EAS. In order to reduce the latency used for exposure of the QoS monitoring results, the local UPF is expected to provide the notifications of network status directly to the 5GMS AF and 5GMS AS, or via a locally deployed NEF as defined in clause 5.8.2.17 of TS 23.501 [23].

1. The Media Session Handler retrieves Service Access Information with the configuration of QoS monitoring provided inside the client Network Assistance configuration.
2. If the Media Session Handler is interested in understanding the network status (e.g., congestion status, packet latency) it creates an enhanced Network Assistance Session **that includes the requested QoS montoring configuration** on the 5GMS AF at reference point M5.
3. Based on the QoS monitoring configuration received in the previous step, the 5GMS AF interacts with the PCF (or NEF) to enable QoS monitoring via the Npcf\_PolicyAuthorization service at reference point N5 or the Nnef\_AFsessionWithQoS service at reference point N33.

Besides, based on the provisioning from the 5GMS Application Provider, the 5GMS AF understands that QoS monitoring is required for 5GMS AS traffic control, e.g. congestion control, bit rate adaptation for progressive download, the 5GMS AF may also request the PCF or NEF to enable the QoS monitoring.

In the case where the 5GMS AS is deployed in the Edge DN, the 5GMS AF may additionally enable the exposure of QoS montoring results via the local UPF or local NEFby configuring the PCF (or NEF).

1. The 5GMS AF invokes the Npcf\_PolicyAuthorization service or the Nnef\_AFsessionWithQoS service **with the requested QoS monitoring configurations**.
2. The PCF accepts the request and enables QoS monitoring within the 5G System, i.e., by configuring the RAN and/or the UPF for monitoring and reporting of target QoS parameters.
3. Following the QoS monitoring request(s), the PCF exposes the QoS monitoring results to the 5GMS AF periocially or by event triggers.
4. **Alternatively, the QoS monitoring results can be exposed to the 5GMS AF by the UPF directly via Nupf\_EventExposure\_Notify service or via a locally deployed NEF via Nnef\_EventExposure\_Notifyservice at reference point N33.**
5. If QoS monitoring was requested by the Media Session Handler, **the 5GMS AF sends the notifications of the QoS monitoring results to the Media Session Handler** via the MQTT notification channel at reference point M5 associated with the Network Assistance Session.
6. **The Media Session Handler further provides the QoS monitoring results to the Media Stream Handler at reference point M11.**
7. **The Media Stream Handler may use the notified QoS monitoring results to modify its behaviour.**

For example, in the case of downlink media streaming, the Media Player may use the monitored packet latency to determine when to request the next media segment, and/or to change the bit rate of the next media segemtn based on the monitored congestion status.

1. If QoS monitoring for the 5GMS AS was provisioned by the 5GMS Application Provider in step 1, **the 5GMS AF provides QoS monitoring notifications to the 5GMS AS via reference point M3**.

NOTE: How the 5GMS AS receives notifications via reference point M3 is for further study.

1. The 5GMS AS may use the notified QoS monitoring results to modify its behaviour.

For example, the 5GMS AS may use the monitored packet latency, congestion status to adjust the congestion window.

#### NOTE: Whether notification of network status to the 5GMS AS is practical, useful and desirable is for further study. 5.23.4.3 L4S-on-request for downlink media streaming

An Application Function may request L4S support from the 5G Network for a certain QoS Flow, e.g. by invoking the Nnef\_AfsessionWithQoS service. The concept of this solution is that an application only requests L4S support from the network when the application layer provides support. The activation leverages the existing 5GMS Dynamic Policy invocation API, allowing the 5GMS-Aware Application to request L4S support as and when it is needed.

A high-level call flow for downlink media streaming is sketched in figure 5.23.4.3-1 below. The following is assumed:

- The service here is a unicast downlink streaming service with dynamic policy support, as described in clause 5.7 of TS 26.501 [15].

- The Layer 4 protocol used for application flows is TCP and the TCP stack used supports L4S.

- The network supports L4S packet marking.

- The application has specifically requested ECN marking for its media delivery session.

- NG-RAN manipulaties the ECN bits (per clause 5.37.3.2 of TS 23.501 [23]). It is equally possible that the PSA-UPF manipulates the ECN bits (per clause 5.37.3.3 of [23]).



Figure 5.23.4.3-1: Call flow for L4S on request

The steps are as follows:

0: *Policy Template Provisioning.* A Policy Template is provisioned with L4S capability.

1: *Dynamic Policy activation.* The Media Session Handler within the 5GMSd Client obtains Service Access Information and triggers a dynamic policy activation. The selected Policy Template is one configured with L4S capability.

2: QoS request. The 5GMSd AF requests QoS handling adding the L4S indication using e.g. the Nnef\_AfSessionWithQoS service or the Npcf\_PolicyAuthorization service. This indicates that the new QoS flow should be L4S-enabled. The new QoS flow with the L4S indication setting propagates through the 5G System.

3: The Media Player within the 5GMSd Client ensures that ECN is enabled within the used transport protocol.

NOTE: This step may happen implicitly by selecting an L4S supporting transport protocol stack.

4: The Media Player within the 5GMSd Client triggeres the establishment of a TCP connection. The ECT(1) codepoint is set in the IP header, indicating an L4S-Capable Transport, and the SDAP entity ensures that the packet is forwarded via the matching QoS flow.

5: The 5GMSd AS responds to the TCP connection establishment request. The 5GMSd AS sets ECT(1) in the IP headers, indicating an L4S-Capable Transport.

6: The UPF finds the matching QoS flow identifier for the downlink packet and sends the packet via the according QoS flow to the UE. TCP Connection setup continues, with one ECT bit set in all packets.

7. When the RAN detects an upcoming congestion (based on continuous congestion monitoring), the 5G System sets the CE (Congestion Experienced) codepoint in the IP header of the downlink packet.

8. The TCP protocol stack used by the Media Player in the 5GMSd Client reflects the Early Congestion Notification to the TCP sender by setting the ECN-Echo (ECE) flag in the TCP header of an uplink PDU of the same TCP connnection. The TCP sender reacts to the ECN-Echo accordingly (i.e., by reducing its sending congestion window).

NOTE 1: The ECN-Echo flag is also acknowledged by the TCP sender setting the Congestion Window Reduced (CWR) flag in an outgong TCP frame, but this acknowledgement is not illustrated in this call flow.

NOTE 2: Classic ECN [X6] requires an ECN signal to be treated as equivalent to a packet drop. L4S [X1] specifies a more fine-grained response and an early congestion signal triggers a less severe reaction. How a TCP sender behaves “accordingly” is not in scope of the specification.

9. Based on the CE indication received in step 7, or by detecting a reduced bit rate in the downlink application flow, the Media Player in the 5GMSd Client reacts by, for example, changing the requested representation.

### 5.23.5 Gap analysis and requirements

Editor’s Note: Other issues that need to be solved are FFS.

#### 5.23.3.1 Integrating QoS monitoring and/or ECN marking for L4S

#### 5.23.5.2 QoS monitoring for media streaming

#### 5.23.5.3 L4S-on-request for downlink media streaming

Based on the call flow in clause 5.23.4.3, the following observations are made:

- L4S/ECN does not require modifications to the Media Player.

- The 5GMSd AF needs to explicitly request L4S handling of packets by the 5G System by interacting with the PCF at reference point N5 (or else via the NEF at reference poiont N33).

- The Policy Template structure needs to be extended to include an L4S enablementflag.

- An L4S-capable transport protocol stack is required in both the 5GMSd Client and at the 5GMSd AS.

NOTE: When the transport protocol stack used on the UE or the Application Server does not support ECN marking, the ECT flags are set accordingly to explicitly indicate lack of support.

- Whether L4S requires explicit activation at session start is for further study.

### 5.23.6 Candidate solutions

Editor’s Note: Candidate solutions including call flows, protocols and APIs for identified issues are FFS.

### 5.23.7 Summary and conclusions

\* \* \* \* End of changes \* \* \* \*