**SA WG4 Meeting #131 S4-250146**

**17 - 21 Februry 2025, Geneva,**

**Source: Huawei, HiSilicon**

**Title: Update of Improved QoS support for Media Streaming services**

**Document for: Approval**

**Agenda Item: 8.5**

**Work Item / Release: AMD-ARCH-MED / Rel-19**

# Introduction

During the telcos, the CR on Improved QoS support for Media Streaming services is endorsed in S4aI250042. However, some aspects are still missing, i.e. how the 5GMS Client gets aware of whether the Policy Template requires QoS monitoring before determining to initiate the QoS monitoring in 5G system.

# Proposal

## ===== CHANGE =====

### 4.0.5 Network assistance

The network assistance feature is applicable to both downlink media streaming and uplink media streaming. It enables the 5GMS Client in the UE to interrogate or manipulate the network Quality of Service for an ongoing media streaming session.

High-level procedures for this feature are defined in clause 5.9 (downlink media streaming) and in clauses 6.1, 6.5 and 6.7 (uplink media streaming). The network assistance feature is not explicitly provisioned by the 5GMS Application Provider. It is either available for a particular media streaming session or not, depending on system pre-configuration and/or policy.

Two mechanisms for obtaining network assistance are defined in the present document: one based on interactions with the PCF via network-based components of the 5GMS System (*AF-based network assistance*), the other based on ANBR signalling interactions between the UE modem and the RAN (*ANBR-based network assistance*).



Figure 4.0.5‑1: High-level arrangement for network assistance feature

The following network assistance sub-features are defined in this release for both the AF-based and ANBR-based mechanisms:

1. *Bit rate recommendation (or throughput estimation).* The 5GMS Client requests an estimate from a network-side component of the 5GMS System of the bit rate that can currently be offered by a media streaming session. The network-side component interrogates the PCF on behalf of the 5GMS Client to obtain this information about the PDU session corresponding to the media streaming session.

The 5GMS Client uses this information to adjust its own streaming bit rate to fit within the Quality of Service (QoS) envelope that the network is able to offer, for example by switching to a different representation listed in its Media Entry Point, or by adjusting the encoding bit rate for uplink streaming to fit within this bit rate budget. The media streaming Quality of Experience (QoE) is more stable and consistent as a consequence.

2. *Delivery boost.* The 5GMS Client speculatively requests a temporary boost to the bit rate of a media streaming session from a network-side component of the 5GMS System. The network-side component requests a modification to the PDU session corresponding to the media streaming session from the PCF on behalf of the 5GMS Client. If there is sufficient spare network capacity to accommodate the requested bit rate, it is granted by the 5GMS System on a temporary basis.

The 5GMS Client uses this temporary boost to speed up media streaming data transfer, for example to replenish a depleted downlink streaming buffer or to complete a download/upload faster than would otherwise be possible.

3. *QoS monitoring*. Based on the monitorable QoS parameters in the Service Access Information, the 5GMS Client requests a monitoring from a network-side component of the 5GMS System of the monitorable QoS parameters that is currently experienced by a media streaming session. The network-side component interrogates the PCF on behalf of the 5GMS Client to enable the QoS monitoring and obtain the corresponding QoS monitoring results about the QoS Flow(s) corresponding to the media streaming session.

The 5GMS Client uses this information to behave accordingly, e.g. using the monitored packet latency to determine when to request the next media segment, and/or to change the bit rate of the next media segment based on the monitored congestion status.

In addition, the use of network assistance by 5GMS Clients is logged by the 5GMS System and, if suitably provisioned, is exposed by it to subscribing 5GMS Application Providers in the form of events (see also clause 4.0.12).

## ===== CHANGE =====

### 4.0.6 Dynamic policies

The dynamic policies feature is applicable to both downlink media streaming and uplink media streaming. It enables the 5GMS Client in the UE to manipulate the network traffic handling policies for an ongoing media streaming session.



NOTE: The PCF is accessed via the NEF when the 5GMS network services are deployed outside the Trusted DN.

Figure 4.0.6‑1: High-level arrangement for dynamic policies



Figure 4.0.6‑2: Domain model for dynamic policies

With reference to figure 4.0.6‑2, dynamic policies work as follows:

1. A conceptual *Service Operation Point* is an abstract set of requirements that support a media streaming service (e.g., SD, HD, UHD). It is identified by an *External reference* that is used to tag *Policy Template* resources provisioned in the 5GMS System and *Service Descriptions* included in *Media Entry Point* documents.

2. The Service Operation Point is embodied in the 5G System by a *Policy Template* which is provisioned in the 5GMS network services by the 5GMS Application Provider within the scope of an umbrella *Provisioning Session*. A Policy Template may be defined as being applicable to a particular Data Network and/or Network Slice. The Policy Template carries the *External reference* and Network QoS parameters corresponding to a single Service Operation Point. (Any number of Policy Templates provisioned for different Data Networks and/or Network Slices may reference the same Service Operation Point.) The 5GMS network services may reject attempts to provision a Policy Template that specifies Network QoS parameters outside acceptable bounds imposed by local system configuration.

In addition, the Policy Template may include a reference to an existing Background Data Transfer policy. If no previously defined Background Data Transfer policy exists, the Policy Template may instead include the parameters that are used by the 5GMS network services to provision a Background Data Transfer policy for the current Provisioning Session. These parameters may include desired time windows when Background Data Transfer may be advertised to 5GMS Clients, a quota representing the maximum number of 5GMS Clients that are permitted to take advantage of Background Data Transfers in each such time window and a quota representing a ceiling for the aggregate volume of data that all 5GMS Clients are permitted to transfer in each Background Data Transfer window. Hence, an advertised time window is not a guarantee that a request for Background Data Transfer will actually be granted by the 5GMS System.

The Policy Template may include an *L4S enablement* flag to enable ECN marking for L4S in the 5G System (as described in clause 5.37.3 of TS 23.501 [2]). If set, this flag directs the 5GMS Client to select and activate ECN marking for L4S when it instantiates the Policy Template. The 5GMS network services accept the provisioning of such a Policy Template only if the underlying 5G System supports the detection of congestion and reaction to it.

NOTE 1: As described in RFC 9330 [X1], RFC 9331 [X2] and RFC 9332 [X3], the purpose of ECN marking for L4S (Low Latency, Low Loss and Scalable Throughput) is to inform a recipient host at the earliest opportunity that an IP packet has experienced network congestion at some point in its routing path. It exposes congestion information by marking ECN bits in the IP header of the user IP packets between the UE and the Application Server.

The Policy Template may include a *QoS monitoring configuration* to enable QoS monitoring in the 5G System (as described in clause 5.45 of TS 23.501 [2]) for measurement and reporting of QoS parameters when this Policy Template is instantiated. The QoS monitoring configuration indicates the trigger for reporting (event or periodic), the set of QoS parameters that may be monitored when this Policy Template is instantiated and, optionally, an indication that notifications are to be sent via the UPF. The resulting Service Access Information for the Policy Template indicates the (possibly empty) set of QoS parameters that may be monitored. Based on its own knowledge of the intended media delivery session, or based on input from an application, the 5GMS Client may select the entire available set or a subset of QoS parameters to be monitored when it instantiates the Policy Template. QoS monitoring is then activated by the 5GMS network services for these QoS parameters. The 5GMS network services notify the 5GMS Client about significant changes to these QoS parameters during the media delivery session.

3. The 5GMS Application Provider makes one or more *Media Entry Point* documents (e.g., DASH MPDs) available for use by the 5GMS Client. To take advantage of the dynamic policies feature, a Media Entry Point document includes one or more *Service Descriptions*, each identifying the streaming requirements of a presentation that correspond to a single Service Operation Point (e.g., SD, HD, UHD) and identified by means of an *External reference*. The same Service Description may be included in more than one Media Entry Point document in case a common Service Operation Point is applicable to multiple media presentations.

4. When a Media Entry Point is selected by the 5GMS Client at the start of a media streaming session, the 5GMS Client retrieves Service Access Information from a network-side component of the 5GMS System describing the set of available Policy Templates provisioned in step 2 and exposes this to a controlling application on the UE.

4a. If Background Data Transfer was provisioned as part of any Policy Templates in step 2 above, the Service Access Information includes details of the advertised time windows when Background Data Transfers are available and the data volume quota (if any). Maximum bit rates for the 5GMS Client in either or both the uplink and downlink direction may also be nominated by the 5G System and signalled to the 5GMS Client in the Service Access Information. Finally, an endpoint in the 5GMS network services may be provided allowing the 5GMS Client to subscribe to receive real-time notifications of Background Data Transfer warning notifications.

5. At the start of a media streaming session, the controlling application on the UE selects one of the Service Descriptions listed in the Media Entry Point document that realises its preferred Service Operation Point. Either the Media Player (when the Service Descriptions are within the Media Entry Point document) or the controlling application (when the Service Descriptions are not within the Media Entry Point document) informs the 5GMS Client of its choice by passing the corresponding External reference to it.

6. If there is a Policy Template available for the current media streaming session with the indicated External reference, the 5GMS Client instantiates this Policy Template by interacting with a network-side component of the 5GMS System in order to realise the Service Operation Point described by the Policy Template and the Service Description. The effect of this is that the corresponding network Quality of Service is applied to the media streaming session.

7. At any point during one of the advertised Background Data Transfer time windows the 5GMS Client may request a Background Data Transfer by instantiating a Policy Template with a Background Data Transfer specification in the 5GMS network services, including an estimate of the data volume it intends to transfer. The 5GMS network services may grant the request for the Background Data Transfer if the data volume estimate is acceptable and if the quota of requests for the time window in question has not already been exceeded. If the request is granted, the 5GMS network services apply the appropriate Background Data Transfer Quality of Service policy to the media streaming session from the Policy Template in question. The Background Data Transfer grant returned to the 5GMS Client includes an estimate of the time period for which Background Data Transfer is available for the 5GMS Client to use. After this period has expired, the 5GMS network services automatically revert the network Quality of Service back to its state before the grant.

8. The 5GMS media services also subscribe to receive Background Data Transfer warning notifications from the PCF related to the individual Background Data Transfer policy as defined in clause 4.16.7.3 of TS 23 502 [3]. The 5GMS media services shall notify the 5GMS Client when the network performance of that particular media streaming session degrades below the Background Data Transfer policy currently in force or when the aggregate data volume for all data transfers during the current Background Data Transfer time window has been reached.

In addition, the use of dynamic policies by 5GMS Clients is logged by the 5GMS System and, if suitably provisioned, is exposed by it to subscribing 5GMS Application Providers in the form of events (see also clause 4.0.12).

## ===== CHANGE =====

4.2.3 Service Access Information for downlink media streaming

The Service Access Information is the set of parameters and addresses which are needed by the 5GMSd Client to activate and control the reception of a downlink streaming session, and to report service/content consumption and/or QoE metrics.

The Service Access Information may be provided together with other service announcement information using M8d. Alternatively, the 5GMSd Client fetches the Service Access Information from the 5GMSd AF. The Service Access Information may be provided as, or may be accessed via, a 3GPP-defined Service URL that provides a unique resolvable identifier to the 5GMSd Provisioning Session and that may also include a reference to the Media Player Entry. Regardless of how it is provided, the Service Access Information contains different information, depending on the collaboration model between the 5GMS System and the 5GMSd Application Provider, and also depending on offered features. Baseline parameters are listed in Table 4.2.3‑1 below:

**Table 4.2.3-1: Parameters of baseline Service Access Information**

|  |  |
| --- | --- |
| **Parameters** | **Description** |
| Provisioning Session identifier | Unique identification of the M1d Provisioning Session. |

When the content hosting feature is activated for a downlink streaming session, the parameters from Table 4.2.3-1a below can additionally be present.

**Table 4.2.3-1a: Streaming Access parameters**

|  |  |
| --- | --- |
| **Parameters** | **Description** |
| Media Player Entries | A set of pointers to documents that each define an equivalent media presentation (see NOTE), e.g. MPD for DASH content or URL to a video clip file.  Each member of the set may specify additional details to aid selection by the 5GMS Client, including content type, profile indicators and precedence.  A Media Player Entry document may additionally include Service Descriptions, each one identified by an *External reference* that enables it to be matched with a Policy Template, and each describing the set of media streaming parameters (e.g., bit rate, target latency) that realise a Service Operation Point.  A Media Player Entry URL may be embedded in a 3GPP Service URL. |
| NOTE: An equivalent media presentation is one which has the same content but may result in a different Quality of Experience. | |

When the consumption reporting feature is activated for a downlink streaming session, the parameters from Table 4.2.3‑2 below are additionally present.

**Table 4.2.3-2: Parameters for consumption reporting configuration**

|  |  |
| --- | --- |
| **Parameters** | **Description** |
| Reporting interval | Identifies the interval between consumption reports being sent by the Media Session Handler. |
| Server address | A list of 5GMSd AF addresses where the consumption reports are sent by the Media Session Handler. |
| Sample percentage | The proportion of clients that shall report media consumption.  If not specified, all clients shall send reports. |
| Location reporting | Identify whether the Media Session Handler provides location data to the 5GMSd AF (in case of MNO or trusted third parties) |

When the dynamic policy invocation feature is activated for a downlink streaming session the parameters from Table 4.2.3‑3 below are additionally present.

**Table 4.2.3-3: Parameters for dynamic policy invocation configuration**

|  |  |
| --- | --- |
| **Parameters** | **Description** |
| Server address | A list of 5GMSd AF addresses (in the form of opaque URLs) which offer the APIs for dynamic policy invocation sent by the 5GMS Media Session Handler. |
| Valid Policy Template Ids | A list of Policy Template identifiers which the 5GMSd Client is authorized to use. |
| Service Data Flow Methods | A list of recommended Service Data Flow description methods (descriptors), e.g. 5-Tuple, ToS, 2-Tuple, etc, which should be used by the Media Session Handler to describe the Service Data Flows for the traffic to be policed. |
| External reference | Additional identifier for this Policy Template that can be cross-referenced with external metadata about the streaming session.  The same external reference may appear on more than one dynamic policy invocation configuration within the scope of the same Provisioning Session provided the parameters below differ in the underlying Policy Template. |
| L4S enablement | A flag indicating that this Policy Template requires ECN marking for L4S.  The 5GMSd Client should not instantiate this Policy Template unless it supports L4S. |
| Monitorable QoS parameters | The set of QoS parameters that may be monitored when this Policy Template is instantiated. |

When the metrics collection and reporting feature is activated for a downlink streaming session, one or more parameter sets for metrics configuration, according to Table 4.2.3‑4, are additionally present. Each metrics configuration set contains specific settings valid for that configuration, which is typically metric scheme dependent, and collection and reporting shall be done separately for each set.

**Table 4.2.3-4: Parameters for each metrics configuration set**

|  |  |
| --- | --- |
| **Parameters** | **Description** |
| Scheme | The scheme associated with this metrics configuration set. A scheme may be associated with 3GPP or with a non-3GPP entity. If not specified, a default 3GPP metrics scheme shall apply.  Metrics schemes shall be uniquely identified by URIs. |
| Server address | A list of 5GMSd AF addresses to which metric reports shall be sent for this metrics configuration set. |
| DNN | The Data Network Name (DNN) which shall be used when sending metrics report for this metrics configuration set.  If not specified, the default DNN shall be used. |
| Slice scope | A list of network slice(s) for which metrics collection and reporting shall be executed for this metrics configuration set.  If not specified, the metrics collection and reporting shall be done for all network slices. |
| Reporting interval | The sending interval between metrics reports for this metrics configuration set.  If not specified, a single final report shall be sent after the streaming session has ended. |
| Sample percentage | The proportion of streaming sessions that shall report metrics for this metrics configuration set.  If not specified, reports shall be sent for all sessions. |
| Streaming source filter | A list of content URL patterns for which metrics reporting shall be done for this metrics configuration set.  If not specified, reporting shall be done for all URLs. |
| Communication Service type | The type of Communication Service (Unicast and/or MBS broadcast and/or MBS multicast) for which metrics collection and reporting is requested.  If not specified, metrics collection and reporting shall be performed for all communication service types. |
| Metrics | A list of metrics which shall be collected and reported for this metrics configuration set.  For progressive download and DASH streaming services, the listed metrics are associated with the 3GPP metrics scheme and shall correspond to one or more of the metrics as specified in clauses 10.3 and 10.4, respectively, of TS 26.247 [7].  In addition, for the 3GPP metrics scheme as applied to DASH streaming, the quality reporting scheme and quality reporting protocol as defined in clauses 10.5 and 10.6, respectively, of [7] shall be used.  If not specified, a complete (or default if applicable) set of metrics will be collected and reported. |

When 5GMSd AF-based Network Assistance is activated for a downlink streaming session the parameters from Table 4.2.3‑5 below shall be additionally present.

**Table 4.2.3-5: Parameters for 5GMSd AF-based Network Assistance configuration**

|  |  |
| --- | --- |
| **Parameters** | **Description** |
| Server address | 5GMSd AF address that offers the APIs for 5GMSd AF-based Network Assistance, accessed by the 5GMSd Media Session Handler. The server address shall be an opaque URL, following the 5GMS URL format. |
| Monitorable QoS parameters | The set of QoS parameters that may be monitored as part of a Network Assistance session. |

## ===== CHANGE =====

4.3.3 Service Access Information for uplink media streaming

The Service Access Information is the set of parameters and addresses which are needed by the 5GMSu Client to activate and control the uplink streaming session.

The Service Access Information may be provided by the 5GMSu Application Provider to the 5GMSu-Aware Application together with other service announcement information using M8u. Alternatively, the 5GMSu Client fetches the Service Access Information from the 5GMSu AF at reference point M5u. Regardless of how it is provided, the Service Access Information contains different information, depending on the collaboration model between the 5GMS System and the 5GMSu Application Provider (which are assumed to be independent entities), and also depending on offered features. The Service Access Information may be provided as, or may be accessed via, a 3GPP-defined Service URL that provides a unique resolvable identifier to the 5GMSu media session and that may also include a reference to the Media Entry Point.

Baseline parameters are listed in table 4.3.3‑1 below:

**Table 4.3.3-1: Parameters of baseline Service Access Information**

|  |  |
| --- | --- |
| **Parameters** | **Description** |
| Provisioning Session identifier | Unique identification of the M1u Provisioning Session. |

The parameters from table 4.3.3-2 below shall also be present.

**Table 4.3.3-2: Streaming Access parameters**

|  |  |
| --- | --- |
| **Parameters** | **Description** |
| Media Streamer Entries | A set of entry points. Each entry point consists of one of the following:  a. A URL endpoint on the 5GMSu AS to which media can be streamed directly at M4u and its associated data, or  b. The URL of a document that can be downloaded from the 5GMSu AS which contains the parameters for uplink media streaming at M4u.  A Media Streamer Entry document may additionally include Service Descriptions, each one identified by an *External reference* that enables it to be matched with a Policy Template, and each describing the set of media streaming parameters (e.g., bit rate, target latency) that realise a Service Operation Point.  A Media Streamer Entry URL may be embedded in a 3GPP Service URL. |

Each entry point is defined by its parameters and identifiers. The set shall have at least one member.

When the dynamic policy invocation feature is activated for an uplink streaming session the parameters from table 4.3.3‑3 below are additionally present.

**Table 4.3.3-3: Parameters for dynamic policy invocation configuration**

|  |  |
| --- | --- |
| **Parameters** | **Description** |
| Server address | A list of 5GMSu AF addresses (in the form of opaque URLs) which offer the APIs for dynamic policy invocation sent by the 5GMS Media Session Handler. |
| Valid Policy Template Ids | A list of Policy Template identifiers which the 5GMSu Client is authorized to use. |
| Service Data Flow Methods | A list of recommended Service Data Flow description methods (descriptors), e.g. 5-Tuple, ToS, 2-Tuple, etc, which should be used by the Media Session Handler to describe the Service Data Flows for the traffic to be policed. |
| External reference | Additional identifier for this Policy Template that can be cross-referenced with external metadata about the streaming session.  The same external reference may appear on more than one dynamic policy invocation configuration within the scope of the same Provisioning Session provided the parameters below differ in the underlying Policy Template. |
| L4S enablement | A flag indicating that this Policy Template requires ECN marking for L4S.  The 5GMSu Client should not instantiate this Policy Template unless it supports L4S. |
| Monitorable QoS parameters | The set of QoS parameters that may be monitored when this Policy Template is instantiated. |

When 5GMSu AF-based Network Assistance is activated for an uplink streaming session the parameters from table 4.3.3‑4 below shall be additionally present.

**Table 4.3.3-4: Parameters for 5GMSu AF-based Network Assistance configuration**

|  |  |
| --- | --- |
| **Parameters** | **Description** |
| Server address | 5GMSu AF address that offers the APIs for 5GMSu AF-based Network Assistance, accessed by the 5GMSu Media Session Handler. The server address shall be an opaque URL, following the 5GMS URL format. |
| Monitorable QoS parameters | The set of QoS parameters that may be monitored as part of a Network Assistance session. |

## ===== CHANGE =====

### 5.3.2 Baseline provisioning procedure

The present clause describes the baseline procedure to provision the features using the 5GMS System.

NOTE 1: SLA negotiations between the 5GMSd Application Provider and the 5GMS System provider are outside the scope of the present specification and are included in the figure below for illustrative purposes only.



Figure 5.3.2-1: High-level procedure for provisioning the 5GMS System  
for downlink media streaming sessions

Steps:

1. The 5GMSd Application Provider discovers the address (URL) of the 5GMSd AF (M1d) for Session Provisioning.

2. Void.

3. The 5GMSd Application Provider creates a Provisioning Session, providing its 5GMSd Application Provider identifier as input. 5GMSd Application Provider queries the capabilities and authorized features.

4. The 5GMSd Application Provider specifies one or more 5GMSd features in the Provisioning Session. A set of authorized features is activated, such as content consumption measurement, logging, collection and reporting; QoE metrics measurement, logging, collection and reporting; dynamic policy; network assistance; and content hosting (including ingest).

One or more *External service identifiers* are supplied by the 5GMSd Application Provider to support the later retrieval of Service Access Information from the 5GMSd AF by the Media Session Handler.

When the content hosting feature is offered and selected, the 5GMS Application Provider configures the content hosting behaviour of the 5GMSd AS. This Content Hosting Configuration is specified in clause 5.4 and includes selecting the ingest protocol and format, caching and proxying of media objects, content preparation, access protection (e.g. URL signing) and indicating a target distribution area (e.g. through geofencing).

When the dynamic policy feature is offered and selected, the 5GMSd Application Provider specifies a set of policies which can be invoked for the unicast downlink streaming session. This may include an *L4S enablement* flag indicating whether ECN marking is a requirement of the Policy Template. It may also include a *QoS monitoring configuration* indicating that QoS monitoringis required when this Policy Template is instantiated, and which QoS parameters may be selected by the 5GMSd Client for monitoring. The UE becomes aware of the selected policies in the form of a list of valid Policy Template Ids.

When the content consumption measurement, logging, collection and reporting feature is offered and selected, the 5GMSd Application Provider indicates the desired reporting interval. When the 5GMSd Application Provider has delegated Service Access Information handling to the 5GMS System, then location reporting is also selected or de-selected.

When the QoE metrics measurement, logging, collection and reporting feature is offered and selected, the 5GMSd Application Provider provides configuration input on the QoE post processing. When the 5GMSd Application Provider has delegated Service Access Information handling to the 5GMS System, then more detailed metrics reporting is configured.

When the edge computing feature is offered and selected, the 5GMSd Application Provider provides one or more Edge Resources Configurations that can be used to support either client-driven management or Application Provider-driven management of edge resources associated with the Provisioning Session.

When the event data processing feature is offered and selected, the 5GMSd Application Provider provides one or more Event Data Processing Configurations that determine how, in the scope of the Provisioning Session, content consumption and QoE metrics collected from the UE and application logs collected from the 5GMSd AS are processed into events and exposed to subscribers.

5. When content hosting is desired, the 5GMSd AF interacts with the 5GMSd AS at reference point M3d to allocate M2d resources and to configure the ingest format by means of a Content Hosting Configuration (defined in clause 5.4) which may reference Server Certificates and Content Preparation Templates, as required. The 5GMSd AS responds with the M2d content ingest address.

6. The 5GMSd AF compiles the Service Access Information. The Service Access Information contains access details and options such as the Provisioning Session identifier, M5d (Media Session Handling) addresses for content consumption reporting, QoE metrics reporting, dynamic policy, network assistance, etc. When content hosting is offered and has been selected in step 4, then also M4d (Media Streaming) information such as the DASH MPD is included.

7. The 5GMSd AF provides the results to the 5GMSd Application Provider.

a. When the 5GMSd Application Provider has selected full Service Access Information, then the results are provided in the form of addresses and configurations for M2d (Ingest), M5d (Media Session Handling) and M4d (Media Streaming).

b. When the 5GMSd Application Provider delegated the Service Access Information handling to the 5GMS System, then a reference to the Service Access Information (e.g., a URL) is provided. The Media Session Handler fetches the full Service Access Information later from the 5GMSd AF.

8. When content hosting is offered and has been selected in step 4, the 5GMSd Application Provider can start supplying content at the M2d ingest interface. In the case of progressive download or on-demand DASH sessions, the 5GMSd Application Provider makes the content assets available. In the case of Live DASH streaming sessions, the 5GMSd Application Provider starts supplying the live content.

9. The 5GMSd Application Provider executes Service Announcement and updates the UEs (during the lifetime of the Provisioning Session).

Optional:

10. The 5GMSd Application Provider may update the Provisioning Session.

Depending on the parameters of the Provisioning Session:

11. The 5GMSd AF may send event-related or periodic notifications to the 5GMSd Application Provider.

According to schedule, or upon request:

12. The 5GMSd Application Provider may manually terminate the Provisioning Session (at any time). All associated resources are released. Content may be removed from the 5GMSd AS. The 5GMSd Application Provider may configure a schedule for Provisioning Session termination.

13. The 5GMSd AF sends a notification upon Provisioning Session termination.

The 5GMSd AF may request the creation or reuse of one or more network slices for distributing the content of the provisioned session. If more than one network slice is provisioned for the distribution of the content of a session, the list of allowed S‑NSSAIs shall be conveyed to the target UEs (e.g. through URSP or through M5d or M8d).

NOTE 2: The 5GMSd AS(s) serving the content are only accessible through the DNN(s) used by the network slice(s) provisioned for the distribution of that content.

## ===== CHANGE =====

### 5.7.10 QoS monitoring of downlink media streaming based on Dynamic Policy

Figure 5.7.10-1 below shows a high-level call flow for the configuration and usage of QoS monitoring with downlink media streaming.



Figure 5.7.10-1: High-level call flow for QoS monitoring for downlink Media Streaming

The steps are as follows:

1. *Policy Template Provisioning.* A Policy Template is provisioned and shall **include the QoS monitoring configuration**. The QoS monitoring configuration includes the parameters to be monitored, the reporting frequency (event triggered, periodic), and optionally the notification via local UPF.

NOTE: In case the 5GMSd AS is deployed as an EAS instance in the Edge DN, a local UPF can also be inserted for local access to the 5GMSd 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 5GMSd AF and 5GMSd AS, or via a locally deployed NEF as defined in clause 5.8.2.17 of TS 23.501 [2].

2. *Service Access Information retrieval*. The Media Session Handler retrieves Service Access Information from 5GMSd AF via reference point M5d.

2a. *Service Access Information retrieval*. The 5GMSd AS retrieves Service Access Information from 5GMSd AF via reference point M3d.

3. *Dynamic Policy activation.* The Media Session Handler within the 5GMSd Client obtains Service Access Information and triggers a dynamic policy activation. A Policy Template Binding is present within the Service Access Information for each provisioned Policy Template. **Policy Template Bindings suitable for QoS monitoring are indicated by a set of QoS parameters that may be monitored. As described in clause 4.0.6, the Media Session Handler determines the subset of QoS parameters to be monitored based on its own knowledge or based on input from the 5GMSd-Aware Application, and this subset is included in the Dynamic Policy activation****.**

3a. *QoS monitoring notification subscription.* The Media Session Handler subscribes to receive QoS monitoring notifications from the 5GMSd AF.

3a. *QoS monitoring notification subscription.* The 5GMSd AS subscribes to receive QoS monitoring notifications from the 5GMSd AF.

4. *QoS monitoring request.* The 5GMSd AF invokes the Npcf\_PolicyAuthorization service or the Nnef\_AFsessionWithQoS service **with the requested QoS monitoring configurations**. In the case where the 5GMSd AS is deployed in the Edge DN, the 5GMSd AF may additionally enable the exposure of QoS montoring results via the local UPF or local NEF in this step.

5. The PCF accepts the request and enables QoS monitoring within the 5G System, i.e., by configuring the RAN and/or the (local) UPF for monitoring and reporting of target QoS parameters for the downlink media streaming.

Following the QoS monitoring request(s):

6. The PCF may expose QoS monitoring results to the 5GMSd AF periodically or by event triggers using the Npcf\_PolicyAuthorization\_Notify service operation directly at reference point N5, or else using the Nnef\_EventExposure\_Notifyservice operation via the NEF at reference point N33.

7. Alternatively, the QoS monitoring results may be exposed to the 5GMSd AF by the UPF directly using the Nupf\_EventExposure\_Notify service or via a locally deployed NEF using the Nnef\_EventExposure\_Notifyservice at reference point N33.

8. If QoS monitoring was requested by the Media Session Handler, **the 5GMSd AF sends the notifications of the QoS monitoring results to the Media Session Handler** via reference point M5d.

**9. The Media Session Handler provides the QoS monitoring results to the Media Stream Handler at reference point M11d.**

**10. 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 segment based on the monitored congestion status.**

**11. The 5GMSd AF may provide the QoS monitoring results to the 5GMSd AS at reference point M3d.**

**12. The 5GMSd AS may use the notified QoS monitoring results to modify its behaviour.**

## ===== CHANGE =====

## 5.9 Downlink Network Assistance

### 5.9.1 Introduction

The Network Assistance (NA) feature enables a UE that is receiving a downlink media stream to improve the QoE of the media streaming session, by being able to make use of three distinct facilities.

The first facility is **bit rate recommendation** (**throughput estimation**). This enables the UE to start a downlink streaming session at the most appropriate bit rate for the network conditions at hand, or to obtain a recommendation from the network which will remain valid until further notice during a media streaming session. The recommended bit rate is based on network estimations or predictions of available link bandwidth. This function is provided as an additional tool to support the UE, in addition to the common approach of the UE performing its own estimation based on measurement of the downlink traffic in the past.

The second facility is the **delivery boost**. The 5GMSd Client uses this function to indicate to the network that a temporary boost, i.e., a temporary increase of network throughput for this client is needed in order to avoid the risk of media playback stalling due to buffer under-run, which could otherwise occur during the next media segment or soon after. Throughput boosting may also be used at the start of a playback session to shorten the time to start media playback, giving a better experience for the user.

The third facility is the **QoS monitoring**. The 5GMSd Client use this function to get aware of the network status, e.g. UL/DL/Round-trip latency, UL/DL congestion, UL/DL data rate. Based on the monitored QoS parameters, the 5GMSd Client behaves accordingly. This function is provided as an additional tool to support the UE, in addition to the common approach of the UE performing its own estimation based on measurement of the downlink traffic in the past.

Network Assistance for downlink media streaming may be offered to the UE in one of two ways:

- Based on interaction between the UE and the 5GMSd AF, with a subsequent interaction between the 5GMSd AF and the PCF (or the NEF);

- Based on interaction between the UE and the RAN, re-using the ANBR-based RAN signalling.

The UE shall not use both approaches on the same Network Assistance session.

Figure 5.9.1-1 depicts the Network Assistance feature in the context of the 5GMS architecture, showing the scope of both approaches.



Figure 5.9.1-1: Downlink Network Assistance alternative approaches

### 5.9.2 5GMSd AF-based downlink Network Assistance

The Network Assistance (NA) feature enables a UE to receive a bit rate recommendation from the 5GMSd AF that provides the NA server function. The 5GMSd AF provides the response with an estimation of throughput, or the recommend­ation of a bit rate which will remain valid until further notice, or the QoS monitoring results. The UE uses this estimation to derive the most suitable bit rate for its downlink content session from the versions that might be available. The network is expected to commit resources to be able to fulfil the recommendation, although no guarantee for the bandwidth estimation can be assumed. The 5GMSd Client may provide additional information, such as available media versions, in terms of the required bit rates, to the 5GMSd AF.

The second facility is the delivery boost. The 5GMSd Client uses this function to indicate to the network that a temporary boost, i.e. a temporary increase of network throughput for this client, is needed.

Each interaction for the 5GMSd AF-based downlink Network Assistance procedures consists of two steps in sequence:

1. Between the UE (Media Session Handler) and the 5GMSd AF using a 5GMS API at interface M5d;

2. Between the 5GMSd AF and the PCF (or the NEF):

- using the Npcf\_PolicyAuthorization procedure or Nnef\_AFSessionWithQoS procedure to request modification of the PDU session with the requested QoS information.

- using the Npcf\_PolicyAuthorization notification service or Nnef Monitoring Event procedure to receive QoS changes. The 5GMSd AF receives these policy change notifications asynchronously.

Network Assistance is performed within dedicated NA sessions that are a part of the Media Session Handler procedures. Only 5GMSd Clients that have been granted an NA session by the 5GMSd AF may execute the NA bit rate recommendation and boost request procedures. This enables the 5GMSd AF to enforce policies that could include the access to NA by certain 5GMSd Clients only, such that authorisation to use NA facilities can be verified once at the granting of an NA session, and does not need to be verified for each NA request from the 5GMSd Client.

The procedures for 5GMSd AF-based downlink Network Assistance are:

1. NA session initiation. The 5GMSd Client requests to initiate an NA session with the 5GMSd AF. If the request fulfils any pre-requisites for access to NA functionality, for example policy and charging, then the 5GMSd AF responds with a confirmation that the NA session has been established.

2. NA throughput estimation or QoS monitoring by 5G System. The 5GMSd Client requests a throughput estimation for a downlink media session from the 5GMSd AF. A unique identifier for the downlink media session is provided by the 5GMSd Client.

If the set of available bit rates in the downlink media session is provided with the request then the 5GMSd AF responds with the recommended bit rate based on its throughput estimation. If the throughput estimation is lower than the lowest value from the set of available bit rates, then the actual throughput estimation is provided.

If no set of available bit rates is provided with the request then the 5GMSd AF responds with the throughput estimation.

If the set or subset of monitorable QoS parameters in the downlink media session are provided with the request, then the 5GMSd AF activates the QoS monitoring for the set or subset of monitorable QoS parameters in 5G System and responds with the corresponding QoS monitoring results. If no monitorable QoS parameter is provided with the request, the 5GMSd AF activates the QoS monitoring for the set of monitorable QoS parameters in 5G System and responds with the corresponding QoS monitoring results.

3. NA delivery boost. The 5GMSd Client requests from the AF a downlink delivery boost. A unique identifier for the downlink media session is provided by the 5GMSd Client.

4. NA session termination. The 5GMSd Client requests to terminate an NA session with the 5GMSd AF. The unique identifier of the NA session to be terminated is provided by the 5GMSd Client.

The AF responds positively if the indicated session could be terminated, and negatively otherwise.

### 5.9.3 ANBR-based downlink Network Assistance

This approach for the downlink Network Assistance feature is based upon the exchange of the bit rate recommendation message and bit rate recommendation query message between the RAN and UE as described in TS 38.321 [9]. In this case the Network Assistance function in the 5GMSd Client utilises a UE-internal facility to access the RAN modem driver to access ANBR signalling operations and to receive corresponding responses from the network. Hence, this control interface is outside the scope of the 5GMS architecture.

The RAN in this architecture contains control plane and user plane entities that interact with peer control plane and user plane entities in the UE.

- The control plane interactions involve the sending by the RAN of downlink bit rate recommendations to the UE, which may occur either independently, or in response to a boost request from the UE, via the exchange of the bit rate recommendation message and bit rate recommendation query message between the RAN and UE as described in TS 38.321 [9]. This control interface is outside the scope of the 5GMS architecture.

- The user plane interaction pertains to downlink streaming media transport over interface M4d.

## ===== CHANGE =====

#### 6.2.2.2 Baseline provisioning procedure

This clause describes the baseline procedure to provision the features using the 5GMS System.

NOTE 1: SLA negotiations between the 5GMSu Application Provider and the 5GMS System provider are outside the scope of the present specification and are included in the figure below for illustrative purposes only.



Figure 6.2.2.2-1: High-level procedure for provisioning the 5GMS System  
for uplink media streaming sessions

Steps:

1. The 5GMSu Application Provider authenticates itself with the system. This procedure reuses existing authentication/authorization procedures, e.g., as defined for CAPIF [13].

2. The 5GMSu Application Provider creates a Provisioning Session, providing its 5GMSu Application Provider identifier as input. 5GMSu Application Provider queries the capabilities and authorized features.

3. The 5GMSuApplication Provider specifies one or more 5GMSu features in the Provisioning Session. A set of authorized features is activated, such as content dynamic policy; network assistance; and content publishing (including egest).

When the content publishing feature is offered and selected, the 5GMS Application Provider provides a Content Publishing Configuration to configure the content publishing behaviour of the 5GMSu AS (see next step), including selecting the uplink ingest protocol and format, content preparation and egest protocol and format.

When the dynamic policy feature is offered and selected, the 5GMSu Application Provider specifies a set of policies which can be invoked for the uplink streaming session. This may include an *L4S enablement* flag indicating whether ECN marking is a requirement of the Policy Template. It may also include a *QoS monitoring configuration* indicating that QoS monitoring is required when this Policy Template is instantiated, and which QoS parameters may be selected by the 5GMSd Client for monitoring. The UE becomes aware of the selected policies in the form of a list of valid Policy Template Ids.

When the edge computing feature is offered and selected, the 5GMSu Application Provider provides one or more Edge Resources Configurations that can be used to support either client-driven management or Application Provider-driven management of edge resources associated with the Provisioning Session.

4. When content publication is desired, the 5GMSu AF interacts with the 5GMSu AS at reference point M3u to configure any necessary Server Certificates and/or Content Publishing Templates and to allocate resources for M2u egest protocol and format by means of a Content Publishing Configuration. The 5GMSu AS responds to the 5GMSu AF with the M2u content egest address.

5. The 5GMSu AF compiles the Service Access Information. The Service Access Information contains access details and options such as the Provisioning Session identifier, M5u (Media Session Handling) addresses for uplink entry point, dynamic policy, network assistance, etc.

6. The 5GMSu AF provides the results to the 5GMSu Application Provider.

The following steps:

7. When the 5GMSu Application Provider has selected full Service Access Information, the results are provided in the form of addresses and configurations for M2u (content egest), M5u (Media Session Handling) and M4u (Media Uplink Streaming). The 5GMSu Application Provider provides a subset of this information to the 5GMSu-Aware Application through M8u.

8. When the 5GMSu-Aware Application decides to activate the streaming service transmission, the Service Access Information is provided to the 5GMSu Client.

9. The 5GMSu Client requests the 5GMSu AF to initialise uplink media streaming (M5u), including reservation of any resources required for content preparation.

Or, alternatively:

10. The 5GMS-Aware Application requests the 5GMSu Client to start an uplink streaming session (M6u/M7u).

11. When the 5GMSu Application Provider has delegated Service Access Information handling to the 5GMS System, a reference to the Service Access Information (e.g. an URL) is provided. The Media Session Handler fetches the full Service Access Information later from the 5GMSu AF.

Then:

12. The 5GMSu Client streams the content to the 5GMSu AS.

13. When content publishing is offered and has been selected in step 4, the 5GMSu Application Provider can start retrieving the content from the M2u egest interface.

Optionally:

14. The 5GMSu Application Provider may update the Provisioning Session.

According to schedule, or upon request by the 5GMSu-Aware Application:

15. The 5GMSu Application Provider may manually terminate the Provisioning Session (at any time). All associated resources are released. Content may be removed from the 5GMSu AS. The 5GMSd Application Provider may configure a schedule for Provisioning Session termination.

16. The 5GMSu AF sends a notification to the 5GMSu Client upon Provisioning Session termination.

The 5GMSu AF may request the creation or reuse of one or more network slices for ingesting the content of the provisioned session. If more than one network slice is provisioned for the ingest of the content of a session, the list of allowed S‑NSSAIs shall be conveyed to the target UE (e.g. through URSP or through M8u, step 7, or M5u, step 10).

NOTE 2: The 5GMSu AS receiving the content is only accessible through the DNN(s) used by the network slice(s) provisioned for the distribution of that content.

## ===== CHANGE =====

### 6.9.9 QoS monitoring of uplink media streaming based on Dynamic Policy

Figure 6.9.9-1 below shows a high-level call flow for the configuration and usage of QoS monitoring with uplink media streaming.



Figure 6.9.9-1: High-level call flow for QoS monitoring for uplink Media Streaming

The steps are as follows:

1. *Policy Template Provisioning.* A Policy Template is provisioned and shall include **the QoS monitoring configuration**. The QoS monitoring configuration includes the parameters to be monitored, the reporting frequency (event triggered, periodic), and optionally the notification via UPF.

NOTE 1: In case the 5GMSu AS is deployed as an EAS instance in the Edge DN, a local UPF can also be inserted for local access to the 5GMSu Edge AS. 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 5GMSu AF and 5GMSu AS, or via a locally deployed NEF as defined in clause 5.8.2.17 of TS 23.501 [2].

2. *Service Access Information retrieval*. The Media Session Handler retrieves Service Access Information from 5GMSu AF via reference point M5u.

2a. *Service Access Information retrieval*. The 5GMSu AS retrieves Service Access Information from 5GMSu AF via reference point M3u.

3. *Dynamic Policy activation.* The Media Session Handler within the 5GMSu Client obtains Service Access Information and triggers a dynamic policy activation. A Policy Template Binding is present within the Service Access Information for each provisioned Policy Template. **Policy Template Bindings suitable for QoS monitoring are indicated by a set of QoS parameters that may be monitored. As described in clause 4.0.6, the Media Session Handler determines the subset of QoS parameters to be monitored based on its own knowledge or based on input from the 5GMSd-Aware Application, and this subset is included in the Dynamic Policy activation.**

4. *QoS monitoring request.* The 5GMSu AF invokes the Npcf\_PolicyAuthorization service or the Nnef\_AFsessionWithQoS service **with the requested QoS monitoring configurations**. In the case where the 5GMSu AS is deployed in the Edge DN, the 5GMSu AF may additionally enable the exposure of QoS montoring results via the local UPF or local NEF in this step.

5. The PCF accepts the request and enables QoS monitoring within the 5G System, i.e., by configuring the RAN and/or the (local) UPF for monitoring and reporting of target QoS parameters for the uplink media streaming.

Following the QoS monitoring request(s):

6. The PCF may expose the QoS monitoring results to the 5GMSu AF periocially or by event triggers using the Npcf\_PolicyAuthorization\_Notify service operation directly at reference point N5, or else using the Nnef\_EventExposure\_Notifyservice operation via the NEF at reference point N33.

7. Alternatively, the QoS monitoring results may be exposed to the 5GMSu AF by the UPF directly using the Nupf\_EventExposure\_Notify service or via a locally deployed NEF using the Nnef\_EventExposure\_Notifyservice at reference point N33.

8. If QoS monitoring was requested by the Media Session Handler, **the 5GMSu AF sends the notifications of the QoS monitoring results to the Media Session Handler** via reference point M5u.

**9. The Media Session Handler provides QoS monitoring results to the Media Stream Handler at reference point M11u.**

**10. The Media Stream Handler may use the notified QoS monitoring results to modify its behaviour.** For example, in the case of uplink media streaming, the Media Player may use the monitored packet latency, congestion status, etc. to determine the bit rate of the uplink streaming.

**11. The 5GMSu AF may provide the QoS monitoring results to the 5GMSu AS at reference point M3u.**

**12. The 5GMSu AS may use the notified QoS monitoring results to modify its behaviour.**