Source: Samsung Electronics Co. Ltd

**Title: [FS\_MS\_NS\_Ph2] Key Issue #3: Moving media flows to other slices**

**Agenda Item: 2.6**

**Document for: Discussion and Agreement**

# **Introduction**

During the SA4#121 meeting, a contribution S4-221435 was discussed that covered the aspect of moving M4 media flows to other slices. Few comments were received for the proposal. This contribution revises the proposal text with the comments and feedback received during SA4#121 meeting.

Below is a summary of the changes and reason for those changes including a way forward:

1. Contribution S4-221435 submitted to SA4#121 meeting based on discussion of contributions during Post 120-e MBS SWG AH calls:
	1. Reviewed dynamic policy procedure in SA4 specifications (also included in clause 2b in this contribution).
	2. Described the use cases (including an SA2 study) and reason for moving M4 media flows to a different slice (also included in clause 3 of this contribution)
	3. Key issue description for moving media flows to other slices
	4. Couple of solution options and our opinion (also included in clause 4 of this contribution)
2. E-mail discussion during SA4#121 meeting:

One of the following can happen when MSH picks a policy for a given PDU Session in a slice that matches the policy’s *ApplicationSessionContext*, and requests application of that policy to that session traffic to the 5GMS AF:

* 1. AF can apply the policy - No problem. This case is already covered in TS 26.501 and TS 26.512. This option is shown in call flow in 6.3.2.1-1 in current contribution for information purposes.
	2. AF can decline to apply the policy because the slice cannot satisfy the requested QoS (for whatever reason, refer to SA2 study in this contribution). In this case, 5GMS AF informs the MSH that it is declining the request. This approach was suggested when this contribution was discussed during Post 120-e MBS SWG AH meeting, and included in S4-221435. When MSH gets the decline response, there are two choices:
		1. Flows continue in the same PDU Session with whatever QoS that is currently available, or
		2. MSH can check to see if there are any other application policy templates that also match the *ApplicationSessionContext* for that flow. If there is one, then try step b again with newly identified policy template.

Questions/suggestion during SA4#121: Add this to the solution call flow – to describe the retry strategy as a loop with a guard condition such as “[Until Success or applicable policy templates exhausted]”.

Our Explanation and changes in current contribution: Call flow 6.2.3.1-1 updated to show the above with a simple loop to try all applicable policy templates

* 1. Instead of declining, 5GMS AF can suggest alternative slice to realize the requested dynamic policy. When MSH gets this information, it can either use the PDU Session in alternate slice to have the dynamic policy realized, or have the flows continue to stay in current slice with current QoS. This solution option was discussed during Post 120-e MBS SWG AH calls, and submitted in S4-221435 to SA4#121 meeting

Question/suggestion during SA4#121: “why wouldn't the Service Access Information already have listed all possible policy templates in the original M5 Service Access Information response at the start of the media streaming session? Why can't the Media Session Handler just prune non-viable choices from the list based on its local knowledge of the DNNs/slices to which it has access?”

Our Explanation and changes in current contribution: There are couple of aspects to this question in our understanding:

* Does MSH have all available choices (of policy templates) and it picks the appropriate policy template(s) to request? - We believe the MSH has all the available policy templates that it can pick and choose from. This key issue also covers the possible case when none of the available policy templates can provide required QoS in current slice. For this scenario, we are discussing moving flows to alternate slices (also established by SA2 study discussed in this contribution)
* Why can’t the MSH have access to alternate slice information using the ServiceAccessInformation? – Yes, that is a possibility. Instead of 5GMS AF replying to MSH about alternate slice information during dynamic policy request, the 5GMS AF may provide this information to MSH using ServiceAccessInformation API. We updated the call flow in 6.3.2.1-2 to support both of these options.

Suggestion to remove details of how the 5GMS AF gets slice details, in both the call flows, and instead rely on SA2 and SA5 procedures – We updated both the call flows to add to note to reflect the above.

# **Aspects related to moving flows between different slices** (carried over from S4-221435)

1. **Support for moving flows to different slices in SA2 specifications**

An extract from clause 5.15.5.2.2 of TS 23.501 on determining whether ongoing traffic can be routed over other existing PDU Sessions in other slices is as follows:

*“The UE uses either the URSP rules (which includes the NSSP) or the UE Local Configuration as defined in clause 6.1.2.2.1 of TS 23.503 [45] to determine whether ongoing traffic can be routed over existing PDU Sessions belonging to other Network Slices or establish new PDU Session(s) associated with same/other Network Slice.”*

From the above extract, it is either the URSP rules delivered to the UE or the UE local configuration that defines how ongoing traffic can be routed over existing PDU Sessions belonging to other network slices.

Clause 6.6.2 of TS 23.503 describes URSP (UE Route Selection Policy) information. As part of URSP, a set of traffic descriptors that help identifying application flows are specified. Also specified are a list of route selection descriptors that define how the identified flows to be routed through the 5G system. Below is an extract from Table 6.6.2.1-3 of TS 23.503 about slicing related entries in the Route Selection Descriptor of URSP.

*Table 6.6.2.1-3: Route Selection Descriptor*

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| *Information name* | *Description* | *Category* | *PCF permitted to modify in URSP* | *Scope* |
| ***Route selection components*** | *This part defines the route selection components* | *Mandatory(NOTE 2)* |  |  |
| *Network Slice Selection* | *Either a single value or a list of values of S-NSSAI(s).* | *Optional**(NOTE 3)* | *Yes* | *UE context* |
| *DNN Selection* | *Either a single value or a list of values of DNN(s).* | *Optional* | *Yes* | *UE context* |
| *..**NOTE 3: When the Subscription Information contains only one S-NSSAI in UDR, the PCF needs not provision the UE with S-NSSAI in the Network Slice Selection information. The "match all" URSP rule has one S-NSSAI at most.**..* |

From the table above, one of the route selection components is the network slice selection information which is either a single value or a list of values of S-NSSAIs, and this defines the NSSP (Network Slice Selection Policy). When this information is present in the URSP rules delivered to the UE by the PCF, the UE routes the identified application traffic through the slices identified by the NSSP. It is to be noted that the network slice selection information can be a list of S-NSSAIs, but the specification does not specify how the identified application traffic can be moved between the slices.

For UE Local Configuration, clause 6.1.2.2.1 of TS 23.503 describes UE policy control, and specifies that among the four policy objects – Access Network Discovery & Selection Policy (ANDSP), UE Route Selection Policy (URSP), V2X Policy (V2XP), and ProSe Policy (ProSeP), two policy objects – ANDSP and URSP may be pre-configured in the UE.

From above SA2 specifications, it is not clear how the application flows can be moved between different slices.

1. **Review of dynamic policy in SA4 specifications**

Clause 7.9 of TS 26.512 specifies Policy Templates Provisioning API using which an 5GMS Application Provider configures a set of Policy Templates within the scope of the Provisioning Session that can be subsequently applied to downlink or uplink media streaming sessions. A Policy Template is identified by its *policyTemplateId*  and represents a set of PCF/NEF parameters which define the service quality and associated charging for the corresponding downlink or uplink media streaming sessions. Also specified that when the Policy Template is used for QoS Flows the *M1QOSSpecification* shall be included. As specified in Table 6.4.3.2-1 of TS 26512, the *M1QoSSpecification* includes properties such as maximum bit rate for uplink and downlink, maximum authorized bit rate for uplink and downlink by 5GMS Application Provider, default packet loss rates for uplink and downlink.

In addition to the above, the data model for Policy Template resource as specified in clause 7.9.3.1 includes the *ApplicationSessionContext* information element which provides information about “*sliceInfo*” and “*dnn*” information sub-elements that represent the slice and DNN for which the policy template is applicable.

When UE needs the support for dynamic policy, it creates a dynamic policy request to the 5GMS AF. Clause 11.5 of TS 26.512 describes the M5 Dynamic Policies API that allows the Media Session Handler to request a specific policy and charging treatment to be applied to a particular application data flow of a downlink or uplink media session. As part of the request, as specified in clause 11.5.3 of TS 26512 describing data model of *DynamicPolicy* resource, among a set of parameters, the *policyTemplateId* and *M5QoSSpecification* information is sent by the MSH to the 5GMS AF. The *policyTemplateId* is the same Id that the 5GMS Application Provider configured using the M1 API above and is informed to the UE during the M5 Service Access Information API as specified in clause 11.2 of TS 26512. The *M5QoSSpecificaiton* represents the requested QoS, specified in clause 6.4.3.3 of TS 26512, and is very similar to the structure of *M1QOSSpecification*.

Note: It is not clear from TS 26512 what the behavior is if MSH includes both the *policyTemplateId* and M5QoSSpecification (requested QoS), and the requested M5QoSSpecification cannot be satisfied by the network slice for which policy template applies.

It is possible that the UE application flows be migrated to a different network slice to achieve the required QoS for the downlink or uplink session. However, there is no text in our SA4 specifications that deal with migration of UE media flows to different network slices due to dynamic policy procedure.

# **Use Cases for moving M4 media flows to a different slice** (carried over from S4-221435)

**Use case 1: Network slice service continuity**

3GPP SA2 is conducting a Rel-18 study on “Enhancement of Network Slicing Phase 3” (UID 940063). As part of this, 3GPP SA2 is studying a key issue “Key Issue#1: Support of Network Slice Service Continuity” and the findings are documented in TR 23700-41. Clause 5.1 in this TR describes the use case of slice service continuity. An excerpt from this clause is below:

|  |
| --- |
| **“5.1 Key Issue #1: Support of Network Slice Service continuity**5.1.1 Description*This Key issues is aiming to address WT#1. The following scenarios can happen:****1) No mobility scenario:*** *Scenario 1a): network slice is overloaded in NG-RAN.* *Scenario 1b): network slice or network slice instance is overloaded or undergoing planned maintenance in CN (e.g. network slice termination).* *Scenario 1c): network performance of the network slice cannot meet the SLA.****2) Inter RA Mobility scenario:****..* *Scenario 2d): network slice or network slice instance is overloaded in the target CN.**This key issue is to study whether and how to provide service continuity for PDU sessions in network slices in the above scenarios 1b), 1c) and 2d).*” |

The key issue above documents the agreements that the network slices can be overloaded from time to time, they can undergo planned maintenance, and the performance of the network slice cannot meet the SLA (3GPP TS 29.520 describes stage-3 APIs using which NF consumers can get slice load level information and network slice instance load level information from NWDAF).

For this key issue, 11 solutions are documented in TR 23700-41:

* Solution #1: Additional S-NSSAI associated with the PDU session
* Solution #2: Slice Re-mapping Capabilities for Network Slice Service Continuity
* Solution #3: Support of Network Slice Service continuity using SSC mode 3
* Solution #4: PDU Session on compatible network slice
* Solution #5: PDU session handover to a target CN with an alternative S-NSSAI support
* Solution #15: Service continuity in case of Network Slice instance overload
* Solution #32: Solution for Network Control for UE Slice Use
* Solution #40: S-NSSAI change decided by PCF
* Solution #41: Network Slice change without service interruption
* Solution #42: Network controlled change to an alternative S-NSSAI
* Solution #43: Allowed NSSAI Determination in Initial Registration to Support Network Slice Service Continuity

Review of the solutions show the following:

1. All the solutions propose methods where an ‘alternative slice’ is identified to move the traffic from the current slice to that identified slice
2. Number of solutions based on which entity identifies the alternate slice: AMF (7), UE (2), SMF (1), PCF (1). It is not clear how each of these entities are configured with such alternate slice information.

As of the latest version of this TR 23700-41 (v1.1.0 of Rel-18):

* Clause 7.1 documents different evaluation points about each of the above 11 candidate solutions
* Clause 8.1 documents five interim conclusion points based on the evaluation of above solutions. There is still work to be done to have a clear conclusion on this topic

From the above, it is clear that network slices cannot be relied on for guaranteed performance or QoS, and thus needs control plane-based solutions to adapt if they start underperforming. This behavior can affect the SA4 specified dynamic policy procedures reviewed in clause 2 above i.e. it is entirely possible that when the MSH requests for a specific policy, the current network slice may not be able to provide the requested QoS, and therefore, the M4 media flows have to be migrated to another slice to receive the required QoS. The changes for clause 6.3 of TR 26941 proposed in this contribution uses a similar idea of an alternate slice as above, and proposes this alternate slice information configuration to happen over M1 messages so 5GMS AF can assist in migrating UE media flows to this alternate slice.

**Use case 2: Upsell vs dedicated premium gaming slice**

[1] discusses aspects related to commercializing 5G network slicing. The white paper discusses network slicing use cases and path to initial commercialization, and evolution of network slicing technology. The white paper discusses gaming slice scenarios, and talks of two use-cases:

* ***Premium service use-case****: where the user pays in a subscription model and will be able to use such premium treatment whenever it needs* [**provided using a separate dedicated premium network slice**]
* ***Upsell use-case****: where the user pays a one-time fee for a premium slice-enabled treatment, such as in the previously mentioned example of a temporary boost in performance for video or gaming*

The white paper further states the following:

*“One aspect we have hinted at but not discussed in detail is the on-demand enablement of slicing in the upsell scenario. While the Premium use case approach can be implemented by configuring the network with allowed NSSAI-s and modifying the user profile to use specific NSSAI when accessing the network, the Upsell approach requires communication between subscriber’s app, or the OS in case it intermediates the payment, and the Core for the purpose of dynamically managing the slicing activation/deactivation, as well as the monetization aspect.* ***An example of upsell in our gaming slice scenario involves asking the user to purchase the gaming slice treatment for a desired duration at a premium slice, at which point a network API would be accessed to enable the user to access the gaming slice. In this example, it would trigger the sending of a new URSP table that contains the rule and route for the gaming NSSAI****”*. While this use case requires a user subscription that triggers switching of slice that can be addressed in the UE using an URSP update, a similar use case can be imagined where the application provider can provision an alternate gaming slice that the UE can move to in case the current slice cannot provide required experience to the users.

# **Solution options and opinion** (carried over from S4-221435)

Given the current standardization in SA2 and SA4 specifications, following options may be possible for addressing the issue of media flow migration between different network slices:

- *Option 1- Decline application of dynamic policy if current slice cannot provide requested QoS*: With this option, the 5GMS AF gets to know the current slice performance information as described in clause 3 above, but it declines the application of dynamic policy. In this case, the M4 media flows may not receive the required QoS support from the network.

- *Option 2: Use an alternate slice in case current slice cannot provide requested QoS:* With this option, the application provider configures an alternate slice information at the 5GMS AF for the UE media flows to be moved to in case the current slice cannot provide the requested QoS based dynamic policy. When the UE requests for such a dynamic policy application, and if the 5GMS AF infers that the current slice cannot provide requested QoS, it informs the MSH in the UE to move the M4 media flow to the alternate slice configured above.

# **References**

[1] “Commercializing 5G Network Slicing”, 5G Americas White Paper, https://www.5gamericas.org/wp-content/uploads/2022/07/Commercializing-5G-Network-Slicing-Jul-2022.pdf, July 2022

# **Proposal**

We propose following change be adopted into TR 26.941 for key issue #3.

**===== 1. CHANGE =====**

## 6.3 Key Issue #3: Moving media flows to other slices

### 6.3.1 Description

#### 6.3.1.1 Migration of UE flows between slices due to dynamic policy

Clause 7.9 of TS 26.512 [20] specifies Policy Templates Provisioning API using which an 5GMS Application Provider configures a set of Policy Templates within the scope of the Provisioning Session that can be subsequently applied to downlink or uplink media streaming sessions.

- The dynamic policy invocation configuration information is fetched by the Media Session Handler from the 5GMS AF using the M5 Service Access Information API specified in clause 11.2 of [20]. The response includes a list of 5GMS AF endpoint addresses that provide the dynamic policy API and set of valid policyTemplateIds.

- When the UE wants a different network QoS policy for 5G Media Streaming, the Media Session Handler in the 5GMS Client creates a dynamic policy request to the 5GMS AF. Clause 11.5 of [20] describes the M5 Dynamic Policies API that allows the Media Session Handler to request a specific policy and charging treatment to be applied to a particular application data flow of a downlink or uplink media streaming session.

- When the Media Session Handler intends to activate a QoS-related Dynamic Policy Template, it includes a M5QoSSpecification property, as specified in clause 6.4.3.3 of [20]. For transient operational reasons, it is possible that the requested M5QoSSpecification cannot be satisfied by the network slice for which policy template applies at the time of the dynamic policy invocation request. To satisfy the network QoS requirements of the target policy template, a potential solution is to extend the scope of the 5GMS architecture to allow the affected media flow to be migrated to a different PDU Session in a different network slice.

Clause 5.15.5.2.2 of TS 23.501 [7] specifies that the USRP rules delivered to the UE or the UE local configuration define how ongoing traffic can be routed over existing PDU Sessions belonging to other network slices. However, it is not clear how the application flows can be moved between different network slices using USRP information.

Open issues:

- How does the 5GMS System determine which slice to migrate the media flow to if the current slice cannot satisfy the requested QoS and how does the Media Session Handler migrate media flows between PDU Sessions on different network slices.

### 6.3.2 Candidate solutions

#### 6.3.2.1 Candidate solution #1: Dynamic policy in case current slice cannot provide requested QoS

Figure 6.3.2.1-1 illustrates the procedure for M5 dynamic policy invocation when the current network slice is unable to provide the requested QoS based dynamic policy.



Figure 6.3.2.1-1: Procedure for dynamic policy when the current network slice cannot provide requested QoS

The steps are as follows:

1. The 5GMS Application Provider configures a provisioning session at the 5GMS AF at reference point M1.

2. The 5GMS Application Provider provides service announcement information to the 5GMS-Aware Application in the UE as described in step 4 of clause 5.1 for downlink streaming and step 4 of clause 6.1 for uplink streaming in TS 26.501 [20].

3. (Optional) In case the 5GMSd Client in UE received only a reference to the Service Access Information, then it acquires the Service Access Information from the 5GMS AF as described in step 6 of clause 5.1 of [20].

4. The 5GMS AF gets current slice information.

NOTE: How the 5GMS AF obtains slice status information is based on SA2 and SA5 procedures.

For dynamic policy, given a set of applicable policy templates whose ApplicationSessionContext information matches that of the current application data flow (i.e. matching slice and dnn information), steps 5-8 are repeated until either the 5GMS AF accepts application of a policy template or until all the applicable policy templates are exhausted. The details of the M5 dynamic policy procedure is specified in step 7 of clause 5.1 for downlink media streaming and step 7 of clause 6.1 for uplink media streaming in [20].

5. The Media Session Handler picks an applicable policy template for dynamic policy operation.

6. The Media Session Handler requests the 5GMS AF to apply the network QoS described by the policy template to the application data flow in the current slice by sending the policyTemplateId and M5QoSSpecification as described in clause 11.5 of [20].

NOTE: It is FFS as to how to 5GMS AF determines what the current slice and DNN is for the incoming request.

7. In the first alternative, the 5GMS AF may conclude that the network QoS described by the requested policy template can be applied in the current slice. The 5GMS AF applies the requested dynamic policy and returns back to the Media Session Handler with a success response.

8. In another alternative, the 5GMS AF may conclude that the network QoS described by the requested policy template cannot be satisfied in the current slice. The 5GMS AF denies application of requested dynamic policy and returns back to the Media Session Handler with a denied response. When the Media Session Handler receives this response, steps 5-8 are repeated with the next applicable policy template

9. The M4 Media Streaming procedures are then carried out as specified in step 8 of clause 5.1 for downlink media streaming and step 8 of clause 6.1 for uplink media streaming in [20] in the current slice with possible performance degradation if all applicable policy templates were exhausted without success.

Figure 6.3.2.1-2 illustrates the procedure for M5 dynamic policy when the current network slice is unable to provide the requested QoS-based dynamic policy, and the 5GMS AF assists in moving the UE media flow to an alternate network slice to receive the requested QoS.



Figure 6.3.2.1-2: Procedure for dynamic policy when the current network slice cannot provide requested policy which triggers migration of media flows to alternate slice

The steps are as follows:

1. The 5GMS Application Provider configures a provisioning session at the 5GMS AF at reference point M1 with information about alternate S-NSSAI to migrate the UE session to in case current slice cannot provide required QoS.

2. The 5GMS Application Provider provides service announcement information to the 5GMS-Aware Application in the UE as described in step 4 of clause 5.1 for downlink streaming and step 4 of clause 6.1 for uplink streaming in TS 26.501 [20]. The service announcement information may have the alternate S-NSSAI configured by the 5GMS Application Provider in step-1.

3. (Optional) In case the 5GMSd Client in UE received only a reference to the Service Access Information, then it acquires the Service Access Information from the 5GMS AF as described in step 6 of clause 5.1 of [20].

4. The 5GMS AF gets current slice status information.

NOTE: How the 5GMS AF obtains slice information is based on SA2 and SA5 procedures.

5. The M5 Media Session Handling procedure is then performed as specified in step 7 of clause 5.1 for downlink media streaming and step 7 of clause 6.1 for uplink media streaming in [20].

a. The Media Session Handler in UE requests application of dynamic policy for the application data flow in the current slice by sending a policyTemplateId and M5QoSSpecification as described in clause 11.5 of [20].

b. Using the information received in step 4 above, the 5GMS AF may conclude that the requested dynamic policy cannot be satisfied in the current slice. The 5GMS AF suggests the Media Session Handler to move the M4 flow to an alternate slice, and may include the alternate S-NSSAI information in response to the Dynamic Policies API request.

6. The Media Session Handler in the UE interacts with the AMF to request the establishment of a new PDU Session or modification of an existing PDU Session to use the alternate network slice as indicated by the alternate S-NSSAI received in step-2 or step-5. The PDU Session establishment and update procedures are defined in clause 4.3 of TS 23.502 [15].

7. The M4 Media Streaming procedures are then carried out as specified in step 8 of clause 5.1 for downlink media streaming and step 8 of clause 6.1 for uplink media streaming in [20] in the alternate slice with the network QoS in the requested dynamic policy.

**===== END CHANGES =====**