Distributed Mobility Management
Internet-Draft
Intended status: Informational
Expires: 27 August 2026

N. Tran
K. Nguyen-Trung
Y. Kim
Soongsil University
23 February 2026

Computing Aware Traffic Steering Consideration for Mobile User Plane Architecture

draft-dcn-dmm-cats-mup-07

Abstract

The document [I-D.draft-ietf-dmm-mup-architecture] describes the Mobile User Plane (MUP) architecture for Distributed Mobility Management. The proposed architecture converts the user mobility session information from the control plane entity to an IPv6 dataplane routing information. When there are multiple candidate instances located at different location to serve an user request, the MUP Provider Edge (PE) might prioritize the closest service location. However, the closest routing path might not be the optimal route.

This document discusses how the mentioned MUP architecture can be leveraged to set up dataplane routing paths to the optimal service instance location with the assistance of computing-aware traffic steering capabilities. For each session request, based on the up-to-date collected computing and network information, the MUP controller can convert the session information to the optimal route.

Status of This Memo

Table of Contents

1. Introduction

The document [I-D.draft-ietf-dmm-mup-architecture] describes the Mobile User Plane architecture for Distributed Mobility Management. This architecture is composed of a MUP controller and multiple MUP PEs. When applying the MUP architecture in 5G network, the MUP PEs accomodate the N3 RAN network as Interwork Segment or the N6 DN network as Direct Segment. The MUP PEs advertises the Interwork and Discovery Segment dataplane network reachability (e.g. Segment Routing IPv6 segment identifier (SRv6 SID)) to the MUP network via the interwork and direct segment Discovery routes. Meanwhile, the MUP controller transformed the received user mobility session information to the corresponding interwork and direct segment information. Then, it advertises the transformed information to MUP

PEs via Session Transformed routes. The MUP PEs use the matching Discovery routes to resolve the Session Transformed routes and forward the packet through the MUP SRv6 network.

This document discusses the optimal route configuration problem when applying the mentioned MUP architecture in a network scenario where an user request can be served by multiple computing instances of the same service located at different locations. The closest geographical service location to users might not be the optimal service instance’s location as pointed out in the problem statement document of IETF Computing-Aware Traffic Steering (CATS) working group [I-D.draft-ietf-cats-usecases-requirements]. In this scenario, an optimal service instance location can be decided at the mobile control plane or data plane.

In the control plane case, it is possible to use an Application Function (AF) to determine the optimal service instance and influence the 5G control plane to select the DN corresponding to the chosen instance. The MUP-C only needs to transform the optimal DN information in the session information into the corresponding route. Meanwhile, in the data plane approach, the MUP-C should decide the optimal service instance location by itself and transform the unoptimal session information into the optimal route based on its decision. The data plane approach can avoid additional signalling procedure at the control plane of the other approach. It also supports IP Routing paradigm benefit of SRv6 mobile user plane as mentioned in the edge computing use case of the document [I-D.draft-ietf-dmm-srv6mob-arch].

Therefore, a solution to integrate CATS capabilities into the mentioned MUP architecture is presented in this document. By considering service computing and network information of all candidate service instances, the MUP controller can convert the session information into the optimal dataplane route.

This document is proposed to discuss a possible extension consideration of the original MUP architecture document[I-D.draft-ietf-dmm-mup-architecture]. Regarding the Distributed Mobility Management requirements described in [RFC7333], the MUP architecture can partly address the "Non-optimal routes" problem and the "Multicast considerations" requirement by integrating CATS capabilties. As described in [RFC4786], anycast is the practice of making a particular service address available in multiple locations. Anycast support could be in the scope of multicast support for distributed mobility management.

2. Terminology used in this draft

CATS-MUP-C: Computing-aware traffic steering MUP-C which integrates CATS path selection and MUP-C features.

Besides, this document uses the following terminologies which has been defined in [I-D.draft-ietf-cats-framework]

CATS: Computing-Aware Traffic Steering takes into account the dynamic nature of computing resource metrics and network state metrics to steer service traffic to a service instance.

Service: An offering that is made available by a provider by orchestrating a set of resources (networking, compute, storage, etc.). The same service can be provided in many locations; each of them constitutes a service instance.

Service instance: An instance of running resources according to a given service logic.

Service contact instance: A client-facing service function instance that is responsible for receiving requests in the context of a given service. A single service can be represented and accessed via several contact instances that run in different regions of a network.

CATS Path Selector (C-PS): A functional entity that computes and selects paths towards service locations and instances and which accommodates the requirements of service requests. Such a path computation engine takes into account the service and network status information.

CATS Service Metric Agent (C-SMA): A functional entity that is responsible for collecting service capabilities and status, and for reporting them to a C-PS.

CATS Network Metric Agent (C-NMA): functional entity that is responsible for collecting network capabilities and status, and for reporting them to a C-PS.

3. MUP enhancement requirements for supporting CATS

This section presents 3 enhancement points that need to be added in MUP for selecting an optimal service instance for serving an user request.

First, the MUP architecture should be capable of identifying the service and its candidate service instances. These service identifiers are well defined in CATS framework document [I-D.draft-ietf-cats-framework], CATS Service ID (CS-ID) is used to differentiate between different services. CATS Instance Selector ID (CIS-ID) is used to differentiate between different service instances of the same service.

Second, the MUP architecture should be capable of advertising service deployment information among its components. The egress MUP PE attaching to the MUP direct segment should gather the corresponding service and servince instance information (CS-ID and CIS-ID) and avertise to the MUP environment. Different methods can be considered for this requirement.

Third, the MUP architecture should be capable of advertising the computing and network metrics (CATS metrics) related to the each service instance. The egress MUP PE attaching to the MUP direct segment should gather the corresponding service CATS metrics and avertise to the MUP environment. Different methods can be considered for this requirement.

This document only discusses the requirements and architecture. Different methods for service information and CATS metrics distribution to the network can be applied. BGP extension is an example approach that can be referred from related IETF documents such as [I-D.draft-lin-idr-distribute-service-metric] or [I-D.draft-ietf-idr-5g-edge-service-metadata]. Solutions for mitigating routing instability issue that might be caused by varied and frequent changed CATS metrics can also be referred in related IETF documents. For example, updating frequency can be set to at least 30 seconds according to [I-D.draft-ietf-idr-5g-edge-service-metadata] or converting low-level fine-grained metrics into a normalized one as mentioned in [I-D.draft-ietf-cats-metric-definition]. Collecting metrics by a centralized entity as described below in this document is another candidate method for reducing the metric advertisement burden in the network.

4. MUP enhancement considerations for supporting CATS

4.1. CATS-MUP Centralized Deployment case

4.1.1. MUP Route enhancements

Compared with the original route definition introduced in [I-D.draft-ietf-dmm-mup-architecture], the Direct Segment Discovery Route (DSD) and the Type 2 Session Transformed Route (T2ST) need modifications to support the centralized CATS-MUP deployment case. Another CATS Metrics Update Route (CMU) is also introduced.

The Direct Segment Discovery route advertises the reachability information of the direct segment. This route is advertised from the PEs attaching to the direct segments to the PEs attaching to the mobile network access side. In CATS scenario, the direct segment is a specific instance of a service. The service identifier CS-ID and service instance identifier CIS-ID information are required in this route. The CS-ID can be used as the direct segment BGP extended community attribute. The list below shows the information that should be included in the BGP NLRI of the DSD route in CATS-MUP centralized deployment case:

* CS-ID

* CIS-ID

* Attached PE SID

The Type 2 Session Transformed Route convert the session information into dataplane routing information. This route is advertised from the CATS-MUP-C to the PEs attaching to the mobile network access side. In CATS scenario, the direct segment is a specific instance of a service. This route type includes the target service identifier CS-ID and the tunnel endpoint identifier on the mobile network core side information. The optimal service instance identifier CIS-ID determined by the CATS-MUP-C is also required in this route information. The list below shows the information that should be included in the BGP NLRI of the T2ST route in CATS-MUP centralized deployment case:

* CS-ID

* Optimal CIS-ID

* Tunnel Endpoint Identifier on the core side

The CATS Metric Update route convert the session information into dataplane routing information. This route is advertised from the PEs attaching to the direct segments to the CATS-MUP-C. This route type update the CATS metrics related to the attaching service instance of each PE to the CATS-MUP-C. The list below shows the information that should be included in the BGP NLRI of the CMU route in CATS-MUP centralized deployment case:

* CS-ID

* CIS-ID

* CATS metrics

4.1.2. Deployment architecture

```
                          +----------------+
                          |    Mobility    |
                          |   Management   |
                          |     System     |
                          +----------------+
                                   |
                     Session Information, CS-ID
                                   |
                  T2ST    +--------v-------+  CMU A
                  +-------|   CATS-MUP-C   |<-----+
               +--|-------|    +------+    |------|--+
               |  |       |    | C-PS |    |      |  |
               |  |       +----------------+      |  |   +----------+
   UE-         |  v                 CMU B ^       |  |   |   C-SMA  |
      \+---+   +------+ DSD A             |   +------+   |----------|-Service1
   UE--|RAN|---|  PE  |<------------------|---|  PE  |---|  Service | Instance
       +---+   +------+<---------------\  \   +------+   |   Site A | (CS-ID S1)
   UE-/        |        DSD B           \  \         |   +----------+ (CIS-ID S1A)
               |                         \  \        |
               |                          \  \       |
               |                           \  \      |
               |     MUP network            \ +------+   +----------+
               |      +-------+              \|  PE  |---|  Service |
               |      | C-NMA |               +------+   |   Site B |-Service1
               |      +-------+                      |   |----------| Instance
               +-------------------------------------+   |   C-SMA  | (CS-ID S1)
                                                         +----------+ (CIS-ID S1B)
```

Figure 1: CATS-MUP Centralized deployment option

Figure 1 describes the CATS-MUP Centralized deployment architecture. The controller MUP-C in previous mentioned document is enhanced with CATS path selection capability and renamed to CATS-MUP-C. The Centrailized deployment option has the following key features:

* The DSD routes advertise service and service instance identity information (CS-ID and CIS-ID) from the Service Site PEs to the RAN PEs. This happens only one time when a service instance is deployed or removed at a service site

* The CMU routes periodically advertise CATS metrics corresponding to each service instance from the Service Site PEs to the CATS-MUP-C

* The CATS-MUP-C decides the optimal service instance for each UE session based on the metrics from the CMU routes

* The T2ST routes convert the UE session information into the optimal underlay routing information and provide them to the RAN PEs.

4.1.3. New UE request underlay routing setup procedure

```
   +----+  +----+  +-----+  +-----+  +-----+
   | UE |  |RAN1|  |CATS |  |SiteA|  |SiteB|
   |    |  | PE |  |MUP-C|  | PE  |  | PE  |
   +----+  +----+  +-----+  +-----+  +-----+
     |       |       |        |        |
     |       0.DSD
     |       (CS-ID S1, CIS-ID S1A, SiteA PE SID)
     |       (CS-ID S1, CIS-ID S1B, SiteB PE SID)
     |       |<---------------|        |
     |       |<------------------------|
     |       |       |        |        |
     |       1.CMU - periodically advertised
     |       (CS-ID S1, CIS-ID S1A, CATS metrics)
     |       (CS-ID S1, CIS-ID S1B, CATS metrics)
     |       |       |<-------|        |
     |       |       |<----------------|
     2. UE request session establishment
     CATS-MUP-C recieves UE session information
     |       |       |        |        |
     |      3.CATS-MUP-C determines
     |      optimal service instance S1A
     |       |       |        |        |
     |      4.TS2T
     |      (CS-ID S1, CIS-ID S1A, TEID)
     |       |<------|        |        |
     |       |       |        |        |
     5. RAN1 PE use DSD route from SiteA (CIS-ID S1A)
     to route UE traffic to SiteA
     |------>|--------------->|        |
     |       |       |        |        |
     6. In case of optimal service instance change
     to SiteB based on current metrics
     |       |       |        |        |
     |      7.CATS-MUP-C determines
     |      optimal service instance S1B
     |       |       |        |        |
     |      8.TS2T
     |      (CS-ID S1, CIS-ID S1B, TEID)
     |       |<------|        |        |
     |       |       |        |        |
     9. RAN1 PE use DSD route from SiteB (CIS-ID S1B)
     to route UE traffic to SiteB
     |------>|----------------------->|
```

Figure 2: New UE request underlay routing setup procedure

Figure 3 describes the sequence of how the CATS-MUP Centralized deployment approach setup the underlay routing path for an UE request.

* First, when a service instance is available at a service site, a DSD route is advertised from the corresponding PE of that site. In the example in Figure 3, DSD routes are advertised from SiteA and SiteB PEs to all of the RAN PEs. These routes announce the service and service instance identities (CS-ID, CIS-ID) running at each service site and the corresponding PE SRv6 SID

* Then, the CATS metrics corresponding to each service instance is periodically updated from the Site PEs to the CATS-MUP-C via the CMU routes

* When a UE request session establishment to the mobile management system, the CATS-MUP-C recieves the session information which includes the requested service CS-ID and the TEID of the session. The CATS-MUP-C determines the optimal service instance for the session based on current CATS metrics.

* Then, the CATS-MUP-C maps the session information TEID with the optimal service instance CIS-ID and advertises this information to the RAN PE via the T2ST route.

* Upon recieving the T2ST route, the RAN PE selects the DSD route that has the same CIS-ID information with the recieved T2ST route (CIS-ID S1A in this example). The RAN PE uses this D2D route for steering the UE traffic to the optimal service site.

* In case of the optimal service instance change to another site (SiteB in this example) due to current CATS metrics update, the CATS-MUP-C advertises a new T2ST route to the RAN PE, which maps the session information TEID with the new optimal service instance CIS-ID (S1B in this example). The RAN PE resolves the new T2ST route with the corresponding DSD route of the new CIS-ID.

4.1.4. UE mobility handling procedure

```
+----+  +----+ +----+ +-----+  +-----+  +-----+
| UE |  |RAN1| |RAN2| |CATS |  |SiteA|  |SiteB|
|    |  | PE | | PE | |MUP-C|  | PE  |  | PE  |
+----+  +----+ +----+ +-----+  +-----+  +-----+
  |       |       |      |        |        |
  |       0.DSD
  |       (CS-ID S1, CIS-ID S1A, SiteA PE SID)
  |       (CS-ID S1, CIS-ID S1B, SiteB PE SID)
  |       |<----------------------|        |
```

```
     |       |<------------------------------|
     |       0.DSD
     |       (CS-ID S1, CIS-ID S1A, SiteA PE SID)
     |       (CS-ID S1, CIS-ID S1B, SiteB PE SID)
     |       |       |<--------------|       |
     |       |       |<----------------------|
     |       |       |      |        |       |
     |       1.CMU - periodically advertised
     |       (CS-ID S1, CIS-ID S1A, CATS metrics)
     |       (CS-ID S1, CIS-ID S1B, CATS metrics)
     |       |       |      |<-------|       |
     |       |       |      |<---------------|
        UE traffic is currently routed
             from RAN1 to SiteA
     |------>|---------------------->|       |
     |       |       |      |        |       |
     2.UE moves to RAN2
     CATS-MUP-C recieves modified UE session information
     |       |       |      |        |       |
     |       |       |      |        |       |
     |      3a.(If stateful service support required)
     |      Send TS2T map to previous service instance
     |      (CS-ID S1, CIS-ID S1A, TEID)
     |       |       |<-----|        |       |
     |      3a.After stateful migration/previous
     |      packet flow is completed
     |       |       |      |        |       |
     |      3.CATS-MUP-C determines
     |      optimal service instance S1B
     |       |       |      |        |       |
     |      4.TS2T
     |      (CS-ID S1, CIS-ID S1B, TEID)
     |       |       |      |        |       |
     |       |       |<-----|        |       |
     5. RAN2 PE use DSD route from SiteB (CIS-ID S1B)
     to route UE traffic to SiteB
     |-------------->|---------------------->|
```

Figure 3: UE mobility handling procedure

Figure 4 describes the sequence of how the CATS-MUP Centralized deployment approach setup the underlay routing path when UE moves to a new RAN.

* DSD routes are advertised from service site PEs to all RAN PEs. CMU routes are periodically advertised from the site PEs to the CATS-MUP-C

* When UE moves to a different RAN, the CATS-MUP-C recieves modified UE session information from the mobility management system

* The CATS-MUP-C select the new optimal service instance based on the current CATS metrics and send a corresponding T2ST route to the new RAN PE that maps the session information with the new optimal service instance CIS-ID (CIS-ID S1B in this example).

* In case of stateful/sticky service support are required, where connection to the previous service instance need to be maintained, the CATS-MUP-C should firstly send a T2ST route that maps the session information with the previous service instance CIS-ID. Then, depends on the session continuity policy, the CATS-MUP-C can update the PE with a new T2ST route later. For example, new TS2T can be sent to PE after service packet flow to the previous instance is completed or after service state migration is completed.

* Upon recieving the T2ST route, the RAN PE selects the DSD route that has the same CIS-ID information with the recieved T2ST route (CIS-ID S1B in this example). The RAN PE uses this D2D route for steering the UE traffic to the optimal service site.

4.2. CATS-MUP Distributed Deployment case

4.2.1. MUP Route enhancements

Compared with the original route definition introduced in [I-D.draft-ietf-dmm-mup-architecture], the Direct Segment Discovery Route (DSD) and the Type 2 Session Transformed Route (T2ST) need modifications to support the distributed CATS-MUP deployment case.

The Direct Segment Discovery route advertises the reachability information of the direct segment. This route is advertised from the PEs attaching to the direct segments to the PEs attaching to the mobile network access side. For the distributed CATS-MUP deployment case, in addition to the CS-ID and the CIS-ID, the CATS metrics of the corresponding service instance of the PE is also included. The CS-ID can be used as the direct segment extended community ID. The list below shows the information that should be included in the BGP NLRI of the DSD route in CATS-MUP centralized deployment case:

* CS-ID

* CIS-ID

* CATS metrics

* Attached PE SID

The Type 2 Session Transformed Route convert the session information into dataplane routing information. This route is advertised from the CATS-MUP-C to the PEs attaching to the mobile network access side. For the distributed CATS-MUP deployment case, this route type only includes the target service identifier CS-ID and the tunnel endpoint identifier on the mobile network core side information. The list below shows the information that should be included in the BGP NLRI of the T2ST route in CATS-MUP centralized deployment case:

* CS-ID

* Tunnel Endpoint Identifier on the core side

4.2.2. Deployment architecture

```
                             +----------------+
                             |    Mobility    |
                             |   Management   |
                             |     System     |
                             +----------------+
                                      |
                        Session Information, CS-ID
                                      |
                     T2ST    +--------v-------+
                     +-------|                |
                  +--|-------|      MUP-C     |---------+
                  |  |       |                |         |
                  |  |       +----------------+         |   +-----------+
      UE-         |  v                                  |   |   C-SMA   |
         \+---+   +------+ DSD A                 +------+   |-----------|-Service1
      UE--|RAN|---|  PE  |<----------------------|  PE  |---|  Service  | Instance
          +---+   +------+<---------------\      +------+   |   Site A | (CS-ID S1)
      UE-/        | C-PS | DSD B           \            |   +-----------+ (CIS-ID S1A)
                  +------+                  \           |
                  |                          \          |
                  |                           \         |
                  |     MUP network            \ +------+   +-----------+
                  |      +-------+              \|  PE  |---|  Service  |
                  |      | C-NMA |               +------+   |   Site B |-Service1
                  |      +-------+                      |   |-----------| Instance
                  +-------------------------------------+   |   C-SMA   | (CS-ID S1)
                                                            +-----------+ (CIS-ID S1B)
```

Figure 4: CATS-MUP Distributed deployment option

Figure 5 describes the CATS-MUP Distributed deployment architecture. This option has the following key features:

* The DSD routes periodically advertise service, service instance identity information (CS-ID and CIS-ID) and CATS metrics from the Service Site PEs to the RAN PEs.

* The T2ST routes provides the UE session information and requested service CS-ID to the RAN-PE.

* The C-PS at the RAN-PE selects the optimal service instance based on current CATS metrics provided from the DSD routes.

4.2.3. New UE request underlay routing setup procedure

```
   +----+  +----+  +-----+  +-----+  +-----+
   | UE |  |RAN1|  |     |  |SiteA|  |SiteB|
   |    |  | PE |  |MUP-C|  | PE  |  | PE  |
   +----+  +----+  +-----+  +-----+  +-----+
     |       |       |        |        |
     |       0.DSD (periodically advertised)
     |       (CS-ID S1, CIS-ID S1A, SiteA PE SID, CATS metrics)
     |       (CS-ID S1, CIS-ID S1B, SiteB PE SID, CATS metrics)
     |       |<---------------|        |
     |       |<------------------------|
     |       |       |        |        |
     1. UE request session establishment
     MUP-C recieves UE session information
     |       |       |        |        |
     |      2.TS2T
     |      (CS-ID S1, TEID)
     |       |<------|        |        |
     |       |       |        |        |
     3. C-PS at RAN1 PE determines service instance CIS-ID S1A
     as the optimal one, use the DSD route that has CIS-ID S1A
     to route UE request
     |------>|--------------->|        |
     |       |       |        |        |
     4. In case of optimal service instance change
     to SiteB based on current metrics
     |       |       |        |        |
     5. C-PS at RAN1 PE determines service instance CIS-ID S1B
     as the optimal one, use the DSD route that has CIS-ID S1B
     to route UE request
     |------>|------------------------>|
```

Figure 5: New UE request underlay routing setup procedure

Figure 6 describes the sequence of how the CATS-MUP Distributed deployment approach setup the underlay routing path for an UE request.

* DSD routes are advertised from the PEs of the service sites that hosts the service instances to all of the RAN PEs. Because CATS metrics are periodically updated, these routes are also periodically advertises to the RAN PEs with the latest CATS metrics alongside the service, service instance identities (CS-ID, CIS-ID) and the PE SRv6 SID.

* When a UE request session establishment to the mobile management system, the MUP-C recieves the session information which includes the requested service CS-ID and the TEID of the session. The MUP-C provides this information to the RAN PEs via the T2ST routes.

* Upon recieving the T2ST route, the C-PS at the RAN PE selects optimal service instance for the requested service (identified by the CS-ID) based on the current CATS metrics of the service advertised from the DSD routes. The RAN PE uses the D2D route that has the optimal service instance CIS-ID for steering the UE traffic to the optimal service site.

* In case of the optimal service instance change to another site (SiteB in this example) due to current CATS metrics update, the C-PS at the RAN-PE selects a new optimal service instance based on current CATS metrics. Then, the RAN PE uses the corresponding DSD route of the new optimal service instance to route the UE request.

4.2.4. UE mobility handling procedure

```
   +----+  +----+ +----+ +-----+  +-----+  +-----+
   | UE |  |RAN1| |RAN2| |CATS |  |SiteA|  |SiteB|
   |    |  | PE | | PE | |MUP-C|  | PE  |  | PE  |
   +----+  +----+ +----+ +-----+  +-----+  +-----+
     |       |       |      |        |        |
     |       0.DSD (periodically advertised)
     |       (CS-ID S1, CIS-ID S1A, SiteA PE SID, CATS metrics)
     |       (CS-ID S1, CIS-ID S1B, SiteB PE SID, CATS metrics)
     |       |<----------------------|        |
     |       |<-------------------------------|
     |       0.DSD (periodically advertised)
     |       (CS-ID S1, CIS-ID S1A, SiteA PE SID, CATS metrics)
     |       (CS-ID S1, CIS-ID S1B, SiteB PE SID, CATS metrics)
     |       |       |<--------------|        |
     |       |       |<-----------------------|
     |       |       |      |        |        |
        UE traffic is currently routed
             from RAN1 to SiteA
     |------>|---------------------->|        |
     |       |       |      |        |        |
     2.UE moves to RAN2
     MUP-C recieves modified UE session information
     |       |       |      |        |        |
     |      3a.(If stateful service support required)
     |      Send TS2T map to previous service instance
     |      (CS-ID S1, CIS-ID S1A, TEID)
     |       |       |<-----|        |        |
     |      3a.After stateful migration/previous
     |      packet flow is completed
     |       |       |      |        |        |
     |      3.TS2T
     |      (CS-ID S1, TEID)
     |       |       |      |        |        |
     |       |       |<-----|        |        |
     4. C-PS at RAN2 PE determines service instance CIS-ID S1B
     as the optimal one, use the DSD route that has CIS-ID S1B
     to route UE request
     |-------------->|----------------------->|
```

Figure 6: UE mobility handling procedure

Figure 7 describes the sequence of how the CATS-MUP Distributed deployment approach setup the underlay routing path when UE moves to a new RAN.

* DSD routes are periodically advertised from service site PEs to all RAN PEs with the corresponding CATS metrics

* When UE moves to a different RAN, the MUP-C recieves modified UE session information from the mobility management system, and provided the session information TEID and the requested service CS-ID to the new RAN PE via the T2ST route.

* The C-PS at the new RAN PE selects new optimal service instance based on current CATS metrics provided by the DSD routes. Then it uses the DSD route that maps to the optimal service instance CIS-ID to route the UE request.

* In case of stateful/sticky service support are required, where connection to the previous service instance need to be maintained, the MUP-C should firstly send a special T2ST route that additionally contains the previous service instance CIS-ID information. The RAN-PE then uses the DSD route that maps to the previous instance CIS-ID. Then, depends on the session continuity policy, the CATS-MUP-C can update the PE with a normal T2ST route later. For example, new TS2T can be sent to PE after service packet flow to the previous instance is completed or after service state migration is completed. Once the RAN-PE receives the normal T2ST route, it uses the DSD route that maps to the optimal service instance CIS-ID to route the UE request.

5. References

5.1. Informative References

[I-D.draft-ietf-cats-framework]
Li, C., Du, Z., Boucadair, M., Contreras, L. M., and J. Drake, "A Framework for Computing-Aware Traffic Steering (CATS)", Work in Progress, Internet-Draft, draft-ietf-cats-framework, 17 February 2026, <https://datatracker.ietf.org/doc/draft-ietf-cats-framework/>.

[I-D.draft-ietf-cats-metric-definition]
Yao, K., Li, C., Contreras, L.M., Ros-Giralt, J., and G. Zeng, "CATS Metrics Definition", 2 February 2026, <https://datatracker.ietf.org/doc/draft-ietf-cats-metric-definition/>.

[I-D.draft-ietf-cats-usecases-requirements]
Yao, K., Contreras, L.M., Shi, H., Zhang, S., and Q. An, "Mobile User Plane Architecture using Segment Routing for Distributed Mobility Management", 4 February 2026, <https://datatracker.ietf.org/doc/draft-ietf-cats-usecases-requirements/>.

[I-D.draft-ietf-dmm-mup-architecture]
Matsushima, S., Horiba, K., Kawakami, Y., Murakami, T., Patel, K., and J. Horn, "Mobile User Plane Architecture using Segment Routing for Distributed Mobility Management", Work in Progress, Internet-Draft, ietf-dmm-mup-architecture, 20 October 2025, <https://datatracker.ietf.org/doc/draft-ietf-dmm-mup-architecture/>.

[I-D.draft-ietf-dmm-srv6mob-arch]
Kamata, T., Horn, J., Jalil, L., Cheng, W., and M. Kohno, "Architecture Discussion on SRv6 Mobile User plane", 2 September 2025, <https://datatracker.ietf.org/doc/draft-ietf-dmm-srv6mob-arch/>.

[I-D.draft-ietf-idr-5g-edge-service-metadata]
Dunbar, L., Majumdar, K., Li, C., Mishra, G., and Z. Du, "BGP Extension for 5G Edge Service Metadata", 18 September 2025, <https://datatracker.ietf.org/doc/draft-ietf-idr-5g-edge-service-metadata/>.

[I-D.draft-lin-idr-distribute-service-metric]
Lin, C., Yao, H., and Q. Xiong, "Distribute Service Metric By BGP", 21 December 2025, <https://datatracker.ietf.org/doc/draft-lin-idr-distribute-service-metric/>.

[ieee-access-cats-mup]
Tran, M-N., Duong, V-B., and Y. Kim, "Design of Computing-Aware Traffic Steering Architecture for 5G Mobile User Plane", 24 June 2024, <https://doi.org/10.1109/ACCESS.2024.3418960>.

[RFC4786] Abley, J. and K. Lindqvist, "Operation of Anycast Services", December 2006, <https://datatracker.ietf.org/doc/rfc4786/>.

[RFC7333] Chan, A., Liu, D., Seite, P., Yokota, H., and J. Korhonen, "Requirements for Distributed Mobility Management", August 2014, <https://datatracker.ietf.org/doc/rfc7333/>.

Authors' Addresses

Minh-Ngoc Tran
Soongsil University
369, Sangdo-ro, Dongjak-gu
Seoul
06978
Republic of Korea
Email: mipearlska1307@dcn.ssu.ac.kr

Kiem Nguyen Trung
Soongsil University
369, Sangdo-ro, Dongjak-gu
Seoul
06978
Republic of Korea
Email: kiemnt@dcn.ssu.ac.kr

Younghan Kim
Soongsil University
369, Sangdo-ro, Dongjak-gu
Seoul
06978
Republic of Korea
Phone: +82 10 2691 0904
Email: younghak@ssu.ac.kr