Workgroup: dmm
Internet-Draft:
draft-tjiang-dmm-5g-dupf-5mbs-00
Published: 7 March 2022
Intended Status: Informational
Expires: 8 September 2022
Authors: T. Jiang
China Mobile

5G Distributed UPFs for 5G Multicast and Broadcast Services (5MBS)

Abstract

The companion draft [I-D.zzhang-dmm-5g-distributed-upf] has described the 5G mobile user plane (MUP) via the refinement of distributed UPFs, along with various user plane implementations that some vendors and operators are exploring, with the requirement of not introducing changes to 3GPP architecture & signaling. The document 3GPP TS 23.247 [_3GPP-23.247] for 5G multicast and broadcast services, or 5MBS, specifies the 5GS architecture to support MBS communication. Thanks to the addition of new 5GS network functions (NFs) and MB-interfaces on 5G CP & UP, this might post additional provisioning & implementation challenges to the underlay transport infrastructure.

This document is not an attempt to do 3GPP SDO work in IETF. Instead, it discusses how to potentially integrate distributed UPFs with the delivery of 5MBS communication, as well as the benefits of using distributed UPFs to handle 5MBS traffic delivery.

Status of This Memo

Table of Contents

1. Distributed UPFs in 5G User Plane

Mobile User Plane (MUP) in 5G has two distinct parts: the Access Network part between UE and gNB, and the Core Network part between gNB and UPF. UPFs are traditionally deployed at central locations, with UEs' PDU sessions encapsulated and extended thru GTP-U tunnels via the N3 (and potentially N9) interfaces in 5GS. The interface N6 supports fundamentally a direct IP or Ethernet connection to the data network or DN.

Actually, UPFs could be distributed & deployed closer to gNBs. The draft [I-D.zzhang-dmm-5g-distributed-upf] has described the 5G mobile user plane (MUP) via the refinement of distributed UPFs or dUPFs. The following picture shows the dUPF architecture:

```
                          N3             N6
    UE1          gNB1      |     dUPF1    |
+---------+                |+------+-----+|
|   PDU   |                || PDU  |     ||      PE1
+---------+ +------+------+|+------+ IP/ ||    +-----+--+
|         | |      |GTP-U |||GTP-U |     ||----+ IP/ |  |
| 5G-AN   | |5G-AN +------+|+------+Ether||    |Ether|  |
| xHaul   | |xHaul |L3/2/1|||L3/2/1|     ||    +-----+--+
+---------+ +------|------+|+------------+|   (          )
                           |              |  ( Transport  )  PE3
                           |              |  (  Network   +--+-----+
    UE2          gNB2      |     dUPF2    |  (            |  | IP/ |
+---------+                |+------+-----+|  (   (DN)     |  |Ether|
|   PDU   |                || PDU  |     ||   (           +--+-----+
+---------+ +------+------+|+------+ IP/ ||    +-----+--+
|         | |      |GTP-U |||GTP-U |     ||    | IP/ |  |
| 5G-AN   | |5G-AN +------+|+------+Ether||    |Ether|  |
| xHaul   | |xHaul |L3/2/1|||L3/2/1|     ||    +-----+--+
+---------+ +-------------+|+------------+|      PE2
```

In distributed UPF architecture, the central (PSA) UPF is no longer needed. dUPF1 and UPF2 connect via PE1 and PE2, respectively, to the DN VPN (or network instance/NI) that UE1 and UE2 intend to access. There could exist other PEs, like PE3 in the picture, for other sites of the same network domain(VPN or NI) or for global Internet access.

There are some benefits of distributed UPFs:

* The N3 interface becomes very simple - over a direct or short transport connection between gNB and dUPF.

* The transport infrastructure off N3/N9 and N6 are straightforward, most likely over the same underlay VPN (MPLS, SR-MPLS or SRv6) supporting the traditional N3/N9 tunneling as in centralized PSA UPF case.

* MEC becomes much simpler since no need to deploy centralized PSA UPF plus ULCL UPFs; UE-UE traffic can be optimized for LAN-type services (via host-route).

In short, the distributed UPFs model achieves "N3/N9/N6 shortcut and central UPF bypass", which is desired by many operators.

2. 5G Multicast and Broadcast Services (5MBS)

The 3GPP document TS 23.247 [3GPP-23.247] for 5G multicast and broadcast services, or 5MBS, specifies the 5GS architecture to support MBS communication. The following picture shows the brief system architecture of 5MBS:

```
              ----+----------(SBA for 5GC) ---------+-----
                  |          |                      |
               +--+--+   +---+---+              +---+----+
               | AMF |   |  SMF  |              | MB-SMF |
               +--+--+   +-+-+-+-+              +---+----+
                 /           |                      |
             N2 /         N4 |                  N4mb|
               /             |                      |
              /    N3    +-+-+---+     N19mb    +---+----+ N6mb +-----+
         +-----+---------+  UPF  +--------------| MB-UPF |------| DN |
+----+   |     |         +-------+ (Individual) +---+----+      +-----+
| UE +---+ gNB |                                    |
+----+   +-----+                                    |
               |_________N3mb (shared delivery)_____|
```

TS 23.247 [_3GPP-23.247] adds new 5GS network functions (NFs) on both 5G control-plane (CP) and user-plane (UP). For example, the CP NF MB-SMF is, in collaboration with the regular SMF, to provision and signal to the UP NF MB-UPF (via the interface N4mb) for setting up MBS delivery path.

5MBS has specified two data delivery modes, individual delivery vs. shared delivery:

* Individual delivery: When the (downlink or DL) MBS packets are received by the MB-UPF from the interface N6mb, MB-UPF replicates & forwards those packets towards (multiple) UPFs, via the interface N19mb, through either unicast (requiring multiple GTP tunnels if unicast underlay transport is applied) or multicast (if multicast underlay transport over N19mb is applied) transmission.

* Shared delivery: When the (DL) MBS packets are received by the MB-UPF from N6mb, MB-UPF replicates & forwards those packets towards (multiple) gNBs, via the interface N3mb (the lower-path in the picture), through either (multiple) separate GTP tunnels if unicast underlay transport over N3mb is applied, or a single GTP tunnel if multicast underlay over N3mb is supported.

3. 5G Distributed UPF for 5G MBS Communication

3.1. 5MBS Transport Challenges

The 5MBS architecture in TS 23.247 [_3GPP-23.247] introduces some network challenges:

* Because of the addition of new CP and UP NFs, this will post additional provisioning & implementation challenges to the underlay transport infrastructure. For example, in the individual delivery mode, both SMF and MB-SMF have to synchronize with each

other to help set up the relay/stitching path between UPF, MB-UPF and DN.

* The picture in previous section shows three new interface types corresponding to three different segments: N3mb, N6mb and N19mb. Based on the traffic delivery mode, once MB-UPF receives DL traffic from N6mb, it will have to do either individual or shared delivery.

* In accordance with TS 23.247 [_3GPP-23.247], the underlay transport infrastructure of all three segments can use either unicast or multicast transmission, based on the capabilities of underlay networks. For example, for the DL *shared* delivery from MB-UPF to gNB via the interface N3mb, 5G MBS packets can be transmitted to multiple gNBs via multicast transmission if the underlay network supports. Otherwise, MB-UPF will have to use unicast to transmit separately to (multiple) gNBs. Considering that this unicast/multicast flexibility is applicable to all the three above-mentioned segments, the implementation will have to face more challenges.

3.2. 5G Distributed UPF for 5MBS Implementation

The REQ8 of [RFC7333] talks about the multicast efficiency between non-optimal and optimal routes, where it states that, in term of multicast considerations, DMM SHOULD enable multicast solutions to be developed to avoid network inefficiency in multicast traffic delivery.

The current 5MBS architecture requires all DL multicast traffic go through the (centralized) MB-UPF, regardless of using the individual or shared delivery. In many operators' networks, 5GS might be deployed in a location that is relatively distant from customer (edge) sites. In this scenario, the efficiency of multicast transmission will be compromised. On the other aspect, 5G dUPF, deployed closer to gNB, will make the implementation more efficient:

* For shared delivery, the MB-UPF can be distributed closer to gNB. The N6mb is a normal IP interface which is connected to DN over underlay network. This transport connection will most likely use the VPN infrastructure that has been provisioned by operators for 5GS. As a dUPF, the N3mb tunnel off MB-UPF could be made much simpler. In some field edge sites, a dUPF could co-locate on-prem with gNB, which can even remove the usage of complex (inter-site) VPN to favor native IP transport.

* For individual delivery, it involves two UPFs, one regular UPF and one MB-UPF. To follow the current 3GPP specification, we can distribute and deploy both UPFs closer to gNB. While the DL

traffic off the N6mb interface may achieve the same gain as in the shared-delivery mode, the transport for N19mb tunnel and (regular) N3 tunnel can be significantly simplified. Remember we have mentioned that either unicast or multicast (underlay) transmission can be used for N19mb (and actually also for N6mb and N3mb). Therefore, applying dUPF will help simplify the N19mb VPN transmission.

* For individual delivery, if we expand the scope beyond the current 3GPP spec, we could integrate the regular UPF and MB-UPF together as a distributed UPF, and then deploy the dUPF closer to gNB. In this scenario, both the N19mb and N3 tunnels can be simplified significantly. TS 23.247 [_3GPP-23.247] specifies the behaviors of MB-UPF, as a standalone NF. Indeed, all the features and behaviors that would be implemented by a MB-UPF can be collaboratively integrated into a regular UPF. This type of 'merging' will lead to more network efficiency and better multicast traffic forwarding, conforming the [RFC7333] REQ8.

The draft [I-D.zzhang-dmm-5g-distributed-upf] discussed and compared briefly different tunneling mechanisms to implement 3GPP GTP, i.e., SRv6, MPLS as the underlay, or in [I-D.mhkk-dmm-srv6mup-architecture] specifying a new SRv6 based MUP architecture to replace GTP. While these proposals may experience different issues upon 5MBS transport implementation, dUPF will make it more feasible.

4. Security Considerations

TBD.

5. IANA Considerations

This document requests no IANA actions.

6. References

6.1. Normative References

[RFC7333] Chan, H., Ed., Liu, D., Seite, P., Yokota, H., and J. Korhonen, "Requirements for Distributed Mobility Management", RFC 7333, DOI 10.17487/RFC7333, August 2014, <https://www.rfc-editor.org/info/rfc7333>.

6.2. Informative References

[I-D.mhkk-dmm-srv6mup-architecture]
Matsushima, S., Horiba, K., Khan, A., Kawakami, Y., Murakami, T., Patel, K., Kohno, M., Kamata, T., Camarillo, P., Voyer, D., Zadok, S., Meilik, I., Agrawal, A., and K. Perumal, "Segment Routing IPv6 Mobile User

Plane Architecture for Distributed Mobility Management", Work in Progress, Internet-Draft, draft-mhkk-dmm-srv6mup-architecture-01, 10 November 2021, <https://www.ietf.org/archive/id/draft-mhkk-dmm-srv6mup-architecture-01.txt>.

[I-D.zzhang-dmm-5g-distributed-upf] Zhang, Z., Patel, K., and T. Jiang, "5G Distributed UPFs", Work in Progress, Internet-Draft, draft-zzhang-dmm-5g-distributed-upf-00, 6 March 2022, <https://www.ietf.org/archive/id/draft-zzhang-dmm-5g-distributed-upf-00.txt>.

[_3GPP-23.247] "Architectural enhancements for 5G multicast-broadcast services; V17.1.0", December 2021.

Author's Address

Tianji Jiang
China Mobile

Email: tianjijiang@chinamobile.com