Segment Routing IPv6 for Mobile User Plane

draft-ietf-dmm-srv6-mobile-uplane-21

Abstract

This document specifies the applicability of SRv6 (Segment Routing IPv6) to the user-plane of mobile networks. The network programming nature of SRv6 accomplishes mobile user-plane functions in a simple manner. The statelessness of SRv6 and its ability to control both service layer path and underlying transport can be beneficial to the mobile user-plane, providing flexibility, end-to-end network slicing, and SLA control for various applications.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 10 November 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.
This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction .............................................. 3
2. Conventions and Terminology ............................. 3
   2.1. Terminology ......................................... 3
   2.2. Conventions ......................................... 4
   2.3. Predefined SRv6 Endpoint Behaviors ................ 4
3. Motivation ................................................ 5
4. 3GPP Reference Architecture ............................ 5
5. User-plane modes ......................................... 6
   5.1. Traditional mode .................................... 7
   5.1.1. Packet flow - Uplink ............................ 8
   5.1.2. Packet flow - Downlink .......................... 9
   5.2. Enhanced mode ....................................... 9
   5.2.1. Packet flow - Uplink ............................ 10
   5.2.2. Packet flow - Downlink .......................... 11
   5.2.3. Scalability ...................................... 12
   5.3. Enhanced mode with unchanged gNB GTP behavior .... 12
   5.3.1. Interworking with IPv6 GTP ........................ 12
   5.3.2. Interworking with IPv4 GTP ........................ 15
   5.3.3. Extensions to the interworking mechanisms ....... 18
   5.4. SRv6 Drop-in Interworking ........................... 18
6. SRv6 Segment Endpoint Mobility Behaviors ............... 19
   6.1. Args.Mob.Session .................................... 20
   6.2. End.MAP ............................................. 20
   6.3. End.M.GTP6.D ...................................... 21
   6.7. H.M.GTP4.D ....................................... 25
   6.8. End.Limit: Rate Limiting behavior .................. 26
7. SRv6 supported 3GPP PDU session types .................. 27
8. Network Slicing Considerations .......................... 27
9. Control Plane Considerations ............................ 27
10. Security Considerations .................................. 28
11. IANA Considerations ...................................... 28
12. Acknowledgements ........................................ 29
13. Contributors ............................................. 29
14. References ................................................ 29
1. Introduction

In mobile networks, mobility systems provide connectivity over a wireless link to stationary and non-stationary nodes. The user-plane establishes a tunnel between the mobile node and its anchor node over IP-based backhaul and core networks.

This document specifies the applicability of SRv6 (Segment Routing IPv6) to mobile networks.

Segment Routing [RFC8402] is a source routing architecture: a node steers a packet through an ordered list of instructions called "segments". A segment can represent any instruction, topological or service based.

SRv6 applied to mobile networks enables a source-routing based mobile architecture, where operators can explicitly indicate a route for the packets to and from the mobile node. The SRv6 Endpoint nodes serve as mobile user-plane anchors.

2. Conventions and Terminology

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in [RFC2119].

2.1. Terminology

* CNF: Cloud-native Network Function
* NFV: Network Function Virtualization
* PDU: Packet Data Unit
* PDU Session: Context of a UE connects to a mobile network.
* UE: User Equipment
* UPF: User Plane Function
* VNF: Virtual Network Function (including CNFs)

The following terms used within this document are defined in [RFC8402]: Segment Routing, SR Domain, Segment ID (SID), SRv6, SRv6 SID, Active Segment, SR Policy, Prefix SID, Adjacency SID and Binding SID.
The following terms used within this document are defined in [RFC8754]: SRH, SR Source Node, Transit Node, SR Segment Endpoint Node and Reduced SRH.

The following terms used within this document are defined in [RFC8986]: NH, SL, FIB, SA, DA, SRv6 SID behavior, SRv6 Segment Endpoint Behavior.

2.2. Conventions

An SR Policy is resolved to a SID list. A SID list is represented as <S1, S2, S3> where S1 is the first SID to visit, S2 is the second SID to visit, and S3 is the last SID to visit along the SR path.

(SA, DA) (S3, S2, S1; SL) represents an IPv6 packet with:

* Source Address is SA, Destination Address is DA, and next-header is SRH
* SRH with SID list <S1, S2, S3> with Segments Left = SL
* Note the difference between the <> and () symbols: <S1, S2, S3> represents a SID list where S1 is the first SID and S3 is the last SID to traverse. (S3, S2, S1; SL) represents the same SID list but encoded in the SRH format where the rightmost SID in the SRH is the first SID and the leftmost SID in the SRH is the last SID. When referring to an SR policy in a high-level use-case, it is simpler to use the <S1, S2, S3> notation. When referring to an illustration of the detailed packet behavior, the (S3, S2, S1; SL) notation is more convenient.
* The payload of the packet is omitted.

SRH[n]: A shorter representation of Segment List[n], as defined in [RFC8754]. SRH[SL] can be different from the DA of the IPv6 header.

* gNB::1 is an IPv6 address (SID) assigned to the gNB.
* U1::1 is an IPv6 address (SID) assigned to UPF1.
* U2::1 is an IPv6 address (SID) assigned to UPF2.
* U2:: is the Locator of UPF2.

2.3. Predefined SRv6 Endpoint Behaviors

The following SRv6 Endpoint Behaviors are defined in [RFC8986].

* End.DT4: Decapsulation and Specific IPv4 Table Lookup
* End.DT6: Decapsulation and Specific IPv6 Table Lookup
* End.DT46: Decapsulation and Specific IP Table Lookup
* EndDX4: Decapsulation and IPv4 Cross-Connect
* EndDX6: Decapsulation and IPv6 Cross-Connect
* EndDX2: Decapsulation and L2 Cross-Connect
This document defines new SRv6 Segment Endpoint Behaviors in Section 6.

3. Motivation

Mobile networks are becoming more challenging to operate. On one hand, traffic is constantly growing, and latency requirements are tighter; on the other hand, there are new use-cases like distributed NFV that are also challenging network operations.

The current architecture of mobile networks does not take into account the underlying transport. The user-plane is rigidly fragmented into radio access, core and service networks, connected by tunneling according to user-plane roles such as access and anchor nodes. These factors have made it difficult for the operator to optimize and operate the data-path.

In the meantime, applications have shifted to use IPv6, and network operators have started adopting IPv6 as their IP transport. SRv6, the IPv6 dataplane instantiation of Segment Routing [RFC8402], integrates both the application data-path and the underlying transport layer into a single protocol, allowing operators to optimize the network in a simplified manner and removing forwarding state from the network. It is also suitable for virtualized environments, like VNF/CNF to VNF/CNF networking. SRv6 has been deployed in dozens of networks [I-D.matsushima-spring-srv6-deployment-status].

SRv6 defines the network-programming concept [RFC8986]. Applied to mobility, SRv6 can provide the user-plane behaviors needed for mobility management. SRv6 takes advantage of the underlying transport awareness and flexibility together with the ability to also include services to optimize the end-to-end mobile dataplane.

The use-cases for SRv6 mobility are discussed in [I-D.camarilloelmalky-springdmm-srv6-mob-usecases], and the architectural benefits are discussed in [I-D.kohno-dmm-srv6mob-arch].

4. 3GPP Reference Architecture

This section presents a reference architecture and possible deployment scenarios.

Figure 1 shows a reference diagram from the 5G packet core architecture [TS.23501].
The user plane described in this document does not depend on any specific architecture. The 5G packet core architecture as shown is based on the latest 3GPP standards at the time of writing this draft.

```
+-----+
| AMF |
+/-----+
[N11]
[N2] / +-----+
/ | SMF |
+++----/
/     \\
/       \\
/       \\
/N4
+-----+
+/-----+
[N3] ++++
[N9] ++++
[N6] ++++
DN
UE------gNB------UPF1------UPF2------DN
```

Figure 1: 3GPP 5G Reference Architecture

* UE: User Endpoint
* gNB: gNodeB with N3 interface towards packet core (and N2 for control plane)
* UPF1: UPF with Interfaces N3 and N9 (and N4 for control plane)
* UPF2: UPF with Interfaces N9 and N6 (and N4 for control plane)
* SMF: Session Management Function
* AMF: Access and Mobility Management Function
* DN: Data Network e.g. operator services, Internet access

This reference diagram does not depict a UPF that is only connected to N9 interfaces, although the mechanisms defined in this document also work in such case.

Each session from a UE gets assigned to a UPF. Sometimes multiple UPFs may be used, providing richer service functions. A UE gets its IP address from the DHCP block of its UPF. The UPF advertises that IP address block toward the Internet, ensuring that return traffic is routed to the right UPF.

5. User-plane modes

This section introduces an SRv6 based mobile user-plane.

In order to simplify the adoption of SRv6, we present two different "modes" that vary with respect to the use of SRv6. The first one is the "Traditional mode", which inherits the current 3GPP mobile architecture. In this mode GTP-U protocol [TS.29281] is replaced by...
SRv6, however the N3, N9 and N6 interfaces are still point-to-point interfaces with no intermediate waypoints as in the current mobile network architecture.

The second mode is the "Enhanced mode". This is an evolution from the "Traditional mode". In this mode the N3, N9 or N6 interfaces have intermediate waypoints -SIDs- that are used for Traffic Engineering or VNF purposes transparent to 3GPP functionalities. This results in optimal end-to-end policies across the mobile network with transport and services awareness.

In both, the Traditional and the Enhanced modes, we assume that the gNB as well as the UPFs are SR-aware (N3, N9 and -potentially- N6 interfaces are SRv6).

In addition to those two modes, we introduce two mechanisms for interworking with legacy access networks (those where the N3 interface is unmodified). In this document we introduce them as a variant to the Enhanced mode, however they are equally applicable to the Traditional mode.

One of these mechanisms is designed to interwork with legacy gNBs using GTP/IPv4. The second mechanism is designed to interwork with legacy gNBs using GTP/IPv6.

This document uses SRv6 Segment Endpoint Behaviors defined in [RFC8986] as well as new SRv6 Segment Endpoint Behaviors designed for the mobile user plane that are defined in this document in Section 6.

Note that the modes discussed throughout this section (with the exception of Section 5.4) only have informational purpose to implementors as well as operators deploying this technology. Indeed, it is expected that the operator defines his own operational model that best suits their needs.

5.1. Traditional mode

In the traditional mode, the existing mobile UPFs remain unchanged with the sole exception of the use of SRv6 as the data plane instead of GTP-U. There is no impact to the rest of the mobile system.

In existing 3GPP mobile networks, a PDU Session is mapped 1-for-1 with a specific GTP tunnel (TEID). This 1-for-1 mapping is mirrored here to replace GTP encapsulation with the SRv6 encapsulation, while not changing anything else. There will be a unique SRv6 SID associated with each PDU Session, and the SID list only contains a single SID.
The traditional mode minimizes the changes required to the mobile system; hence it is a good starting point for forming a common ground.

The gNB/UPF control-plane (N2/N4 interface) is unchanged, specifically a single IPv6 address is provided to the gNB. The same control plane signalling is used, and the gNB/UPF decides to use SRv6 based on signaled GTP-U parameters per local policy. The only information from the GTP-U parameters used for the SRv6 policy is the TEID, QFI, and the IPv6 Destination Address.

Our example topology is shown in Figure 2. The gNB and the UPFs are SR-aware. In the descriptions of the uplink and downlink packet flow, A is an IPv6 address of the UE, and Z is an IPv6 address reachable within the Data Network DN. A new SRv6 Endpoint Behavior, End.MAP, defined in Section 6.2, is used.

5.1.1. Packet flow – Uplink

The uplink packet flow is as follows:

\[
\begin{align*}
\text{UE}_\text{out} & : (A, Z) \\
\text{gNB}_\text{out} & : (\text{gNB}, U1::1) (A, Z) \quad \rightarrow \text{H.Encaps.Red} <U1::1> \\
\text{UPF1}_\text{out} & : (\text{gNB}, U2::1) (A, Z) \quad \rightarrow \text{End.MAP} \\
\text{UPF2}_\text{out} & : (A, Z) \quad \rightarrow \text{End.DT4 or End.DT6}
\end{align*}
\]

When the UE packet arrives at the gNB, the gNB performs a H.Encaps.Red operation. Since there is only one SID, there is no need to push an SRH. gNB only adds an outer IPv6 header with IPv6 DA U1::1. gNB obtains the SID U1::1 from the existing control plane (N2 interface). U1::1 represents an anchoring SID specific for that session at UPF1.

When the packet arrives at UPF1, the SID U1::1 is associated with the End.MAP SRv6 Endpoint Behavior. End.MAP replaces U1::1 by U2::1, that belongs to the next UPF (U2).
When the packet arrives at UPF2, the SID U2::1 corresponds to an End.DT4/End.DT6/End.DT46 SRv6 Endpoint Behavior. UPF2 decapsulates the packet, performs a lookup in a specific table associated with that mobile network and forwards the packet toward the data network (DN).

5.1.2. Packet flow - Downlink

The downlink packet flow is as follows:

- **UPF2_in**: (Z,A)
- **UPF2_out**: (U2::, U1::2) (Z,A) -> H.Encaps.Red <U1::2>
- **UPF1_out**: (U2::, gNB::1) (Z,A) -> End.MAP
- **gNB_out**: (Z,A) -> End.DX4, End.DX6, End.DX2

When the packet arrives at the UPF2, the UPF2 maps that flow into a PDU Session. This PDU Session is associated with the segment endpoint <U1::2>. UPF2 performs a H.Encaps.Red operation, encapsulating the packet into a new IPv6 header with no SRH since there is only one SID.

Upon packet arrival on UPF1, the SID U1::2 is a local SID associated with the End.MAP SRv6 Endpoint Behavior. It maps the SID to the next anchoring point and replaces U1::2 by gNB::1, that belongs to the next hop.

Upon packet arrival on gNB, the SID gNB::1 corresponds to an End.DX4, End.DX6 or End.DX2 behavior (depending on the PDU Session Type). The gNB decapsulates the packet, removing the IPv6 header and all its extensions headers, and forwards the traffic toward the UE.

5.2. Enhanced mode

Enhanced mode improves scalability, provides traffic engineering capabilities, and allows service programming [I-D.ietf-spring-sr-service-programming], thanks to the use of multiple SIDs in the SID list (instead of a direct connectivity in between UPFs with no intermediate waypoints as in Traditional Mode).

Thus, the main difference is that the SR policy MAY include SIDs for traffic engineering and service programming in addition to the anchoring SIDs at UPFs.

Additionally in this mode the operator may choose to aggregate several devices under the same SID list (e.g., stationary residential meters connected to the same cell) to improve scalability.
The gNB/UPF control-plane (N2/N4 interface) is unchanged, specifically a single IPv6 address is provided to the gNB. A local policy instructs the gNB to use SRv6.

The gNB MAY resolve the IP address received via the control plane into a SID list using a mechanism like PCEP, DNS-lookup, LISP control-plane or others. The resolution mechanism is out of the scope of this document.

Note that the SIDs MAY use the arguments Args.Mob.Session if required by the UPFs.

Figure 3 shows an Enhanced mode topology. The gNB and the UPF are SR-aware. The Figure shows two service segments, S1 and C1. S1 represents a VNF in the network, and C1 represents an intermediate router used for Traffic Engineering purposes to enforce a low-latency path in the network. Note that neither S1 nor C1 are required to have an N4 interface.

5.2.1. Packet flow - Uplink

The uplink packet flow is as follows:

UE_out : (A,Z)
gNB_out : (gNB, S1)(U1::1, C1; SL=2)(A,Z)->H.Encaps.Red<S1,C1,U1::1>
S1_out : (gNB, C1)(U1::1, C1; SL=1)(A,Z) ->End with PSP
C1_out : (gNB, U1::1)(A,Z) ->End.DT4,End.DT6,End.DT2U
UPF1_out: (A,Z) ->End.DT4,End.DT6,End.DT2U

UE sends its packet (A,Z) on a specific bearer to its gNB. gNB’s control plane associates that session from the UE(A) with the IPv6 address B. gNB’s control plane does a lookup on B to find the related SID list <S1, C1, U1::1>.
When gNB transmits the packet, it contains all the segments of the SR policy. The SR policy includes segments for traffic engineering (C1) and for service programming (S1).

Nodes S1 and C1 perform their related Endpoint functionality and forward the packet.

When the packet arrives at UPF1, the active segment (U1::1) is an End.DT4/End.DT6/End.DT2U which performs the decapsulation (removing the IPv6 header with all its extension headers) and forwards toward the data network.

5.2.2. Packet flow - Downlink

The downlink packet flow is as follows:

UPF1_in : (Z,A) ->UPF1 maps the flow w/ SID list <C1, S1, gNB>
UPF1_out: (U1::1, C1)(gNB::1, S1; SL=2)(Z,A)->H.Encaps.Red
C1_out  : (U1::1, S1)(gNB::1, S1; SL=1)(Z,A)
S1_out  : (U1::1, gNB::1)(Z,A) ->End with PSP
gNB_out : (Z,A) ->End.DX4/End.DX6/End.DX2

When the packet arrives at the UPF1, the UPF1 maps that particular flow into a UE PDU Session. This UE PDU Session is associated with the policy <C1, S1, gNB>. The UPF1 performs a H.Encaps.Red operation, encapsulating the packet into a new IPv6 header with its corresponding SRH.

The nodes C1 and S1 perform their related Endpoint processing.

Once the packet arrives at the gNB, the IPv6 DA corresponds to an End.DX4, End.DX6 or End.DX2 behavior at the gNB (depending on the underlying traffic). The gNB decapsulates the packet, removing the IPv6 header, and forwards the traffic towards the UE. The SID gNB::1 is one example of a SID associated to this service.

Note that there are several means to provide the UE session aggregation. The decision on which one to use is a local decision made by the operator. One option is to use the Args.Mob.Session (Section 6.1). Another option comprises the gNB performing an IP lookup on the inner packet by using the End.DT4, End.DT6, and End.DT2 behaviors.
5.2.3. Scalability

The Enhanced Mode improves since it allows the aggregation of several UEs under the same SID list. For example, in the case of stationary residential meters that are connected to the same cell, all such devices can share the same SID list. This improves scalability compared to Traditional Mode (unique SID per UE) and compared to GTP-U (dedicated TEID per UE).

5.3. Enhanced mode with unchanged gNB GTP behavior

This section describes two mechanisms for interworking with legacy gNBs that still use GTP: one for IPv4, and another for IPv6.

In the interworking scenarios as illustrated in Figure 4, the gNB does not support SRv6. The gNB supports GTP encapsulation over IPv4 or IPv6. To achieve interworking, an SR Gateway (SRGW) entity is added. The SRGW maps the GTP traffic into SRv6.

The SRGW is not an anchor point and maintains very little state. For this reason, both IPv4 and IPv6 methods scale to millions of UEs.

![Figure 4: Example topology for interworking](image)

Both of the mechanisms described in this section are applicable to either the Traditional Mode or the Enhanced Mode.

5.3.1. Interworking with IPv6 GTP

In this interworking mode the gNB at the N3 interface uses GTP over IPv6.

Key points:

* The gNB is unchanged (control-plane or user-plane) and encapsulates into GTP (N3 interface is not modified).
* The 5G Control-Plane towards the gNB (N2 interface) is unmodified, though multiple UPF addresses need to be used - one IPv6 address (i.e. a BSID at the SRGW) is needed per <SLA, PDU session type>. The SRv6 SID is different depending on the required <SLA, PDU session type> combination.
* In the uplink, the SRGW removes GTP, finds the SID list related to the IPv6 DA, and adds SRH with the SID list.
* There is no state for the downlink at the SRGW.
* There is simple state in the uplink at the SRGW; using Enhanced mode results in fewer SR policies on this node. An SR policy is shared across UEs as long as they belong to the same context (i.e., tenant). A set of many different policies (i.e., different SLAs) increases the amount of state required.
* When a packet from the UE leaves the gNB, it is SR-routed. This simplifies network slicing [I-D.ietf-lsr-flex-algo].
* In the uplink, the SRv6 BSID steers traffic into an SR policy when it arrives at the SRGW.

An example topology is shown in Figure 5.

S1 and C1 are two service segments. S1 represents a VNF in the network, and C1 represents a router configured for Traffic Engineering.

![Figure 5: Enhanced mode with unchanged gNB IPv6/GTP behavior](image)

5.3.1.1. Packet flow - Uplink

The uplink packet flow is as follows:

UE_out : (A,Z)
gNB_out : (gNB, B)(GTP: TEID T)(A,Z) -> Interface N3 unmodified (IPv6/GTP)
SRGW_out: (SRGW, S1)(U2::T, C1; SL=2)(A,Z) -> B is an End.M.GTP6.D SID at the SRGW
S1_out : (SRGW, C1)(U2::T, C1; SL=1)(A,Z)
C1_out : (SRGW, U2::T)(A,Z) -> End with PSP
UPF2_out: (A,Z) -> End.DT4 or End.DT6

The UE sends a packet destined to Z toward the gNB on a specific bearer for that session. The gNB, which is unmodified, encapsulates the packet into IPv6, UDP, and GTP headers. The IPv6 DA B, and the GTP TEID T are the ones received in the N2 interface.
The IPv6 address that was signaled over the N2 interface for that UE PDU Session, B, is now the IPv6 DA. B is an SRv6 Binding SID at the SRGW. Hence the packet is routed to the SRGW.

When the packet arrives at the SRGW, the SRGW identifies B as an End.M.GTP6.D Binding SID (see Section 6.3). Hence, the SRGW removes the IPv6, UDP, and GTP headers, and pushes an IPv6 header with its own SRH containing the SIDs bound to the SR policy associated with this BindingSID. There at least one instance of the End.M.GTP6.D SID per PDU type.

S1 and C1 perform their related Endpoint functionality and forward the packet.

When the packet arrives at UPF2, the active segment is (U2::T) which is bound to End.DT4/6. UPF2 then decapsulates (removing the outer IPv6 header with all its extension headers) and forwards the packet toward the data network.

5.3.1.2. Packet flow - Downlink

The downlink packet flow is as follows:

UPF2_in : (Z,A) -> UPF2 maps the flow with <C1, S1, SRGW::TEID,gNB>

UPF2_out: (U2::1, C1)(gNB, SRGW::TEID, S1; SL=3)(Z,A) -> H.Encaps.Red

C1_out  : (U2::1, S1)(gNB, SRGW::TEID, S1; SL=2)(Z,A)

S1_out  : (U2::1, SRGW::TEID)(gNB, SRGW::TEID, S1, SL=1)(Z,A)

SRGW_out: (SRGW, gNB)(GTP: TEID=T)(Z,A) -> SRGW/96 is End.M.GTP6.E

gNB_out : (Z,A)

When a packet destined to A arrives at the UPF2, the UPF2 performs a lookup in the table associated to A and finds the SID list <C1, S1, SRGW::TEID, gNB>. The UPF2 performs an H.Encaps.Red operation, encapsulating the packet into a new IPv6 header with its corresponding SRH.

C1 and S1 perform their related Endpoint processing.

Once the packet arrives at the SRGW, the SRGW identifies the active SID as an End.M.GTP6.E function. The SRGW removes the IPv6 header and all its extensions headers. The SRGW generates new IPv6, UDP, and GTP headers. The new IPv6 DA is the gNB which is the last SID in the received SRH. The TEID in the generated GTP header is an argument of the received End.M.GTP6.E SID. The SRGW pushes the headers to the packet and forwards the packet toward the gNB. There is one instance of the End.M.GTP6.E SID per PDU type.
Once the packet arrives at the gNB, the packet is a regular IPv6/GTP packet. The gNB looks for the specific radio bearer for that TEID and forward it on the bearer. This gNB behavior is not modified from current and previous generations.

5.3.1.3. Scalability

For the downlink traffic, the SRGW is stateless. All the state is in the SRH pushed by the UPF2. The UPF2 must have the UE states since it is the UE’s session anchor point.

For the uplink traffic, the state at the SRGW does not necessarily need to be unique per PDU Session; the SR policy can be shared among UEs. This enables more scalable SRGW deployments compared to a solution holding millions of states, one or more per UE.

5.3.2. Interworking with IPv4 GTP

In this interworking mode the gNB uses GTP over IPv4 in the N3 interface

Key points:

* The gNB is unchanged and encapsulates packets into GTP (the N3 interface is not modified).
* N2 signaling is not changed, though multiple UPF addresses need to be provided - one for each PDU Session Type.
* In the uplink, traffic is classified by SRGW’s classification engine and steered into an SR policy. The SRGW may be implemented in a UPF or as a separate entity. How the classification engine rules are set up is outside the scope of this document, though one example is using BGP signaling from a Mobile User Plane Controller [I-D.mhkk-dmm-srv6mup-architecture].
* SRGW removes GTP, finds the SID list related to DA, and adds an SRH with the SID list.

An example topology is shown in Figure 6. In this mode the gNB is an unmodified gNB using IPv4/GTP. The UPFs are SR-aware. As before, the SRGW maps the IPv4/GTP traffic to SRv6.

S1 and C1 are two service segment endpoints. S1 represents a VNF in the network, and C1 represents a router configured for Traffic Engineering.
5.3.2.1. Packet flow - Uplink

The uplink packet flow is as follows:

\[
gNB_{\text{out}} : (\text{gNB}, B)(\text{GTP: TEID T})(A, Z) \quad \text{\rightarrow Interface N3}\ 
\]

\[
\text{SGGW}_{\text{out}} : (\text{SGGW}, S1)(U2::1, C1; SL=2)(A, Z) \quad \text{\rightarrow H.M.GTP4.D function}\ 
\]

\[
S1_{\text{out}} : (\text{SGGW}, C1)(U2::1, C1; SL=1)(A, Z) \quad \text{\rightarrow PSP}\ 
\]

\[
\text{UPF2}_{\text{out}} : (A, Z) \quad \text{\rightarrow End.DT4 or End.DT6}\ 
\]

The UE sends a packet destined to Z toward the gNB on a specific bearer for that session. The gNB, which is unmodified, encapsulates the packet into a new IPv4, UDP, and GTP headers. The IPv4 DA, B, and the GTP TEID are the ones received at the N2 interface.

When the packet arrives at the SRGW for UPF1, the SRGW has an classification engine rule for incoming traffic from the gNB, that steers the traffic into an SR policy by using the function H.M.GTP4.D. The SRGW removes the IPv4, UDP, and GTP headers and pushes an IPv6 header with its own SRH containing the SIDs related to the SR policy associated with this traffic. The SRGW forwards according to the new IPv6 DA.

S1 and C1 perform their related Endpoint functionality and forward the packet.

When the packet arrives at UPF2, the active segment is (U2::1) which is bound to End.DT4/6 which performs the decapsulation (removing the outer IPv6 header with all its extension headers) and forwards toward the data network.
Note that the interworking mechanisms for IPv4/GTP and IPv6/GTP differs. This is due to the fact that in IPv6/GTP we can leverage the remote steering capabilities provided by the Segment Routing BSID. In IPv4 this construct is not available, and building a similar mechanism would require a significant address consumption.

5.3.2.2. Packet flow - Downlink

The downlink packet flow is as follows:

\[
\begin{align*}
\text{UPF2}_{\text{in}} & : (Z,A) \quad \rightarrow \text{UPF2 maps flow with SID} \<C1, S1, \text{GW}::\text{SA}:\text{DA}:\text{TEID}> \\
\text{UPF2}_{\text{out}} : (U2::1, C1)(\text{GW}::\text{SA}:\text{DA}:\text{TEID}, S1; \text{SL}=2)(Z,A) \rightarrow \text{H.Encaps.Red} \\
\text{Cl}_{\text{out}} & : (U2::1, S1)(\text{GW}::\text{SA}:\text{DA}:\text{TEID}, S1; \text{SL}=1)(Z,A) \\
\text{S1}_{\text{out}} & : (U2::1, \text{GW}::\text{SA}:\text{DA}:\text{TEID})(Z,A) \\
\text{SRGW}_{\text{out}} : (\text{GW}, \text{gNB})(\text{GTP: TEID=T})(Z,A) \quad \rightarrow \text{End.M.GTP4.E} \\
\text{gNB}_{\text{out}} & : (Z,A)
\end{align*}
\]

When a packet destined to A arrives at the UPF2, the UPF2 performs a lookup in the table associated to A and finds the SID list \(<C1, S1, \text{SRGW}::\text{SA}:\text{DA}:\text{TEID}>\). The UPF2 performs a \text{H.Encaps.Red} operation, encapsulating the packet into a new IPv6 header with its corresponding SRH.

The nodes C1 and S1 perform their related Endpoint processing.

Once the packet arrives at the SRGW, the SRGW identifies the active SID as an End.M.GTP4.E function. The SRGW removes the IPv6 header and all its extensions headers. The SRGW generates an IPv4, UDP, and GTP headers. The IPv4 SA and DA are received as SID arguments. The TEID in the generated GTP header is also the arguments of the received End.M.GTP4.E SID. The SRGW pushes the headers to the packet and forwards the packet toward the gNB.

When the packet arrives at the gNB, the packet is a regular IPv4/GTP packet. The gNB looks for the specific radio bearer for that TEID and forwards it on the bearer. This gNB behavior is not modified from current and previous generations.

5.3.2.3. Scalability

For the downlink traffic, the SRGW is stateless. All the state is in the SRH pushed by the UPF2. The UPF must have this UE-base state anyway (since it is its anchor point).
For the uplink traffic, the state at the SRGW is dedicated on a per UE/session basis according to a classification engine. There is state for steering the different sessions in the form of an SR Policy. However, SR policies are shared among several UE/sessions.

5.3.3. Extensions to the interworking mechanisms

In this section we presented two mechanisms for interworking with gNBs and UPFs that do not support SRv6. These mechanisms are used to support GTP over IPv4 and IPv6.

Even though we have presented these methods as an extension to the "Enhanced mode", it is straightforward in its applicability to the "Traditional mode".

5.4. SRv6 Drop-in Interworking

In this section we introduce another mode useful for legacy gNB and UPFs that still operate with GTP-U. This mode provides an SRv6-enabled user plane in between two GTP-U tunnel endpoints.

In this mode we employ two SRGWs that map GTP-U traffic to SRv6 and vice-versa.

Unlike other interworking modes, in this mode both of the mobility overlay endpoints use GTP-U. Two SRGWs are deployed in either N3 or N9 interface to realize an intermediate SR policy.

![Diagram](image)

Figure 7: Example topology for SRv6 Drop-in mode

The packet flow of Figure 7 is as follows:
When a packet destined to Z is sent to the gNB, which is unmodified (control-plane and user-plane remain GTP-U), gNB performs encapsulation into a new IP, UDP, and GTP headers. The IPv6 DA, U::1, and the GTP TEID are the ones received at the N2 interface.

The IPv6 address that was signaled over the N2 interface for that PDU Session, U::1, is now the IPv6 DA. U::1 is an SRv6 Binding SID at SRGW-A. Hence the packet is routed to the SRGW.

When the packet arrives at SRGW-A, the SRGW identifies U::1 as an End.M.GTP6.D.Di Binding SID (see Section 6.4). Hence, the SRGW removes the IPv6, UDP, and GTP headers, and pushes an IPv6 header with its own SRH containing the SIDs bound to the SR policy associated with this Binding SID. There is one instance of the End.M.GTP6.D.Di SID per PDU type.

S1 and C1 perform their related Endpoint functionality and forward the packet.

Once the packet arrives at SRGW-B, the SRGW identifies the active SID as an End.M.GTP6.E function. The SRGW removes the IPv6 header and all its extensions headers. The SRGW generates new IPv6, UDP, and GTP headers. The new IPv6 DA is U::1 which is the last SID in the received SRH. The TEID in the generated GTP header is an argument of the received End.M.GTP6.E SID. The SRGW pushes the headers to the packet and forwards the packet toward UPF. There is one instance of the End.M.GTP6.E SID per PDU type.

Once the packet arrives at UPF, the packet is a regular IPv6/GTP packet. The UPF looks for the specific rule for that TEID to forward the packet. This UPF behavior is not modified from current and previous generations.

6. SRv6 Segment Endpoint Mobility Behaviors
6.1. Args.Mob.Session

Args.Mob.Session provide per-session information for charging, buffering and lawful intercept (among others) required by some mobile nodes. The Args.Mob.Session argument format is used in combination with End.Map, End.DT4/End.DT6/End.DT46 and End.DX4/End.DX6/End.DX2 behaviors. Note that proposed format is applicable for 5G networks, while similar formats could be used for legacy networks.

```
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   QFI     |R|U|                PDU Session ID                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|PDU Sess(cont')|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 8: Args.Mob.Session format

* QFI: QoS Flow Identifier [TS.38415]
* R: Reflective QoS Indication [TS.23501]. This parameter indicates the activation of reflective QoS towards the UE for the transferred packet. Reflective QoS enables the UE to map UL User Plane traffic to QoS Flows without SMF provided QoS rules.
* U: Unused and for future use. MUST be 0 on transmission and ignored on receipt.
* PDU Session ID: Identifier of PDU Session. The GTP-U equivalent is TEID.

Arg.Mob.Session is required in case that one SID aggregates multiple PDU Sessions. Since the SRv6 SID is likely NOT to be instantiated per PDU session, Args.Mob.Session helps the UPF to perform the behaviors which require per QFI and/or per PDU Session granularity.

Note that the encoding of user-plane messages (e.g., Echo Request, Echo Reply, Error Indication and End Marker) is out of the scope of this draft. [I-D.murakami-dmm-user-plane-message-encoding] defines one possible encoding.

6.2. End.MAP

The "Endpoint behavior with SID mapping" behavior (End.MAP for short) is used in several scenarios. Particularly in mobility, End.MAP is used by the intermediate UPFs.

When node N receives a packet whose IPv6 DA is D and D is a local End.MAP SID, N does:
S01. If (IPv6 Hop Limit <= 1) {
S02.    Send an ICMP Time Exceeded message to the Source Address,
       Code 0 (Hop limit exceeded in transit),
       interrupt packet processing, and discard the packet.
S03. }
S04. Decrement IPv6 Hop Limit by 1
S05. Update the IPv6 DA with the new mapped SID
S06. Submit the packet to the egress IPv6 FIB lookup for
       transmission to the new destination

Notes: The SIDs in the SRH are not modified.

6.3.  End.M.GTP6.D

The "Endpoint behavior with IPv6/GTP decapsulation into SR policy" behavior (End.M.GTP6.D for short) is used in interworking scenario for the uplink towards SRGW from the legacy gNB using IPv6/GTP. Any SID instance of this behavior is associated with an SR Policy B and an IPv6 Source Address S.

When the SR Gateway node N receives a packet destined to D and D is a local End.M.GTP6.D SID, N does:

S01. When an SRH is processed {
S02.   If (Segments Left != 0) {
S03.      Send an ICMP Parameter Problem to the Source Address,
           Code 0 (Erroneous header field encountered),
           Pointer set to the Segments Left field,
           interrupt packet processing, and discard the packet.
S04.   }
S05.   Proceed to process the next header in the packet
S06. }

When processing the Upper-layer header of a packet matching a FIB entry locally instantiated as an End.M.GTP6.D SID, N does:
S01. If (Next Header (NH) == UDP & UDP_Dest_port == GTP) {
S02.    Copy the GTP TEID and QFI to buffer memory
S03.    Pop the IPv6, UDP, and GTP Headers
S04.    Push a new IPv6 header with its own SRH containing B
S05.    Set the outer IPv6 SA to S
S06.    Set the outer IPv6 DA to the first SID of B
S07.    Set the outer Payload Length, Traffic Class, Flow Label, Hop Limit, and Next-Header (NH) fields
S08.    Write in the SRH[0] the Args.Mob.Session based on the information of buffer memory
S09.    Submit the packet to the egress IPv6 FIB lookup and transmission to the new destination
S10. } Else {
S11.    Process as per [RFC8986] Section 4.1.1
S12. }

Notes: S07. The NH is set based on the SID parameter. There is one instantiation of the End.M.GTP6.D SID per PDU Session Type, hence the NH is already known in advance. For the IPv4v6 PDU Session Type, in addition we inspect the first nibble of the PDU to know the NH value.

The last segment (S3 in above example) SHOULD be followed by an Arg.Mob.Session argument space which is used to provide the session identifiers.


The "Endpoint behavior with IPv6/GTP decapsulation into SR policy for Drop-in Mode" behavior (End.M.GTP6.D.Di for short) is used in SRv6 drop-in interworking scenario described in Section 5.4. The difference between End.M.GTP6.D as another variant of IPv6/GTP decapsulation function is that the original IPv6 DA of GTP packet is preserved as the last SID in SRH.

Any SID instance of this behavior is associated with an SR Policy B and an IPv6 Source Address S.

When the SR Gateway node N receives a packet destined to D and D is a local End.M.GTP6.D.Di SID, N does:
S01. When an SRH is processed {
S02.   If (Segments Left != 0) {
S03.      Send an ICMP Parameter Problem to the Source Address,
           Code 0 (Erroneous header field encountered),
           Pointer set to the Segments Left field,
           interrupt packet processing, and discard the packet.
S04.   }
S05.   Proceed to process the next header in the packet
S06. }

When processing the Upper-layer header of a packet matching a FIB entry locally instantiated as an End.M.GTP6.Di SID, N does:

S01. If (Next Header = UDP & UDP_Dest_port = GTP) {
S02.   Copy D to buffer memory
S03.   Pop the IPv6, UDP, and GTP Headers
S04.   Push a new IPv6 header with its own SRH containing B
S05.   Set the outer IPv6 SA to S
S06.   Set the outer IPv6 DA to the first SID of B
S07.   Set the outer Payload Length, Traffic Class, Flow Label,
           Hop Limit, and Next-Header fields
S08.   Preprend D to the SRH (as SRH[0]) and set SL accordingly
S09.   Submit the packet to the egress IPv6 FIB lookup and
           transmission to the new destination
S10. } Else {
S11.   Process as per [RFC8986] Section 4.1.1
S12. }

Notes: S07. The NH is set based on the SID parameter. There is one instantiation of the End.M.GTP6.D SID per PDU Session Type, hence the NH is already known in advance. For the IPv4v6 PDU Session Type, in addition we inspect the first nibble of the PDU to know the NH value.

S SHOULD be an End.M.GTP6.E SID instantiated at the SR gateway.

6.5. End.M.GTP6.E

The "Endpoint behavior with encapsulation for IPv6/GTP tunnel" behavior (End.M.GTP6.E for short) is used among others in the interworking scenario for the downlink toward the legacy gNB using IPv6/GTP.

The prefix of End.M.GTP6.E SID MUST be followed by the
Arg.Mob.Session argument space which is used to provide the session identifiers.

When the SR Gateway node N receives a packet destined to D, and D is a local End.M.GTP6.E SID, N does the following:
S01. When an SRH is processed {
S02.   If (Segments Left != 1) {
S03.      Send an ICMP Parameter Problem to the Source Address,  
          Code 0 (Erroneous header field encountered),  
          Pointer set to the Segments Left field,  
          interrupt packet processing, and discard the packet.
S04.   }
S05.   Proceed to process the next header in the packet
S06. }

When processing the Upper-layer header of a packet matching a FIB entry locally instantiated as an End.M.GTP6.E SID, N does:

S01. Copy SRH[0] and D to buffer memory  
S02. Pop the IPv6 header and all its extension headers  
S03. Push a new IPv6 header with a UDP/GTP Header  
S04. Set the outer IPv6 SA to S  
S05. Set the outer IPv6 DA from buffer memory  
S06. Set the outer Payload Length, Traffic Class, Flow Label,  
      Hop Limit, and Next-Header fields  
S07. Set the GTP TEID (from buffer memory)  
S08. Submit the packet to the egress IPv6 FIB lookup and  
      transmission to the new destination

Notes: An End.M.GTP6.E SID MUST always be the penultimate SID. The TEID is extracted from the argument space of the current SID.

The source address S SHOULD be an End.M.GTP6.D SID instantiated at an SR gateway.


The "Endpoint behavior with encapsulation for IPv4/GTP tunnel" behavior (End.M.GTP4.E for short) is used in the downlink when doing interworking with legacy gNB using IPv4/GTP.

When the SR Gateway node N receives a packet destined to S and S is a local End.M.GTP4.E SID, N does:

S01. When an SRH is processed {
S02.   If (Segments Left != 0) {
S03.      Send an ICMP Parameter Problem to the Source Address,  
          Code 0 (Erroneous header field encountered),  
          Pointer set to the Segments Left field,  
          interrupt packet processing, and discard the packet.
S04.   }
S05.   Proceed to process the next header in the packet
S06. }

Matsushima, et al. Expires 10 November 2022
When processing the Upper-layer header of a packet matching a FIB entry locally instantiated as an End.M.GTP4.E SID, N does:

1. Store the IPv6 DA and SA in buffer memory
2. Pop the IPv6 header and all its extension headers
3. Push a new IPv4 header with a UDP/GTP Header
4. Set the outer IPv4 SA and DA (from buffer memory)
5. Set the outer Total Length, DSCP, Time To Live, and Next-Header fields
6. Set the GTP TEID (from buffer memory)
7. Submit the packet to the egress IPv6 FIB lookup and transmission to the new destination

Notes: The End.M.GTP4.E SID in S has the following format:

```
0 127
+-----------------------+-------+----------------+---------+
|  SRGW-IPv6-LOC-FUNC   |IPv4DA |Args.Mob.Session|0 Padded |
+-----------------------+-------+----------------+---------+
128-a-b-c            a            b           c
```

Figure 9: End.M.GTP4.E SID Encoding

The IPv6 Source Address has the following format:

```
0 127
+----------------+-------------------------------+
|  Source UPF Prefix |IPv4 SA | any bit pattern(ignored) |
+----------------+-------------------------------+
128-a-b            a                  b
```

Figure 10: IPv6 SA Encoding for End.M.GTP4.E

6.7. H.M.GTP4.D

The "SR Policy Headend with tunnel decapsulation and map to an SRv6 policy" behavior (H.M.GTP4.D for short) is used in the direction from legacy IPv4 user-plane to SRv6 user-plane network.

When the SR Gateway node N receives a packet destined to a IW-IPv4-Prefix, N does:
S01. IF Payload == UDP/GTP THEN
S02. Pop the outer IPv4 header and UDP/GTP headers
S03. Copy IPv4 DA, TEID to form SID B
S04. Copy IPv4 SA to form IPv6 SA B'
S05. Encapsulate the packet into a new IPv6 header ;;Ref1
S06. Set the IPv6 DA = B
S07. Forward along the shortest path to B
S08. ELSE
S09. Drop the packet

Ref1: The NH value is identified by inspecting the first nibble of the inner payload.

The SID B has the following format:

```
  0
+-----------------------------------------------+-------+----------------+---------+
|Destination UPF Prefix | IPv4DA | Args.Mob.Session| 0 Padded |
+-----------------------------------------------+-------+----------------+---------+
  128-a-b-c               a            b           c
```

Figure 11: H.M.GTP4.D SID Encoding

The SID B MAY be an SRv6 Binding SID instantiated at the first UPF (U1) to bind an SR policy [I-D.ietf-spring-segment-routing-policy].

6.8. End.Limit: Rate Limiting behavior

The mobile user-plane requires a rate-limit feature. For this purpose, we define a new behavior "End.Limit". The "End.Limit" behavior encodes in its arguments the rate limiting parameter that should be applied to this packet. Multiple flows of packets should have the same group identifier in the SID when those flows are in the same AMBR (Aggregate Maximum Bit Rate) group. The encoding format of the rate limit segment SID is as follows:

```
+-----------------------------------------------+----------+-----------+
| LOC+FUNC rate-limit | group-id | limit-rate|
+-----------------------------------------------+----------+-----------+
  128-i-j                i          j
```

Figure 12: End.Limit: Rate limiting behavior argument format

If the limit-rate bits are set to zero, the node should not do rate limiting unless static configuration or control-plane sets the limit rate associated to the SID.

7. SRv6 supported 3GPP PDU session types

The 3GPP [TS.23501] defines the following PDU session types:

* IPv4
* IPv6
* IPv4v6
* Ethernet
* Unstructured

SRv6 supports the 3GPP PDU session types without any protocol overhead by using the corresponding SRv6 behaviors (End.DX4, End.DT4 for IPv4 PDU sessions; End.DX6, End.DT6, End.T for IPv6 PDU sessions; End.DT46 for IPv4v6 PDU sessions; End.DX2 for L2 and Unstructured PDU sessions).

8. Network Slicing Considerations

A mobile network may be required to implement "network slices", which logically separate network resources. User-plane behaviors represented as SRv6 segments would be part of a slice.

[I-D.ietf-spring-segment-routing-policy] describes a solution to build basic network slices with SR. Depending on the requirements, these slices can be further refined by adopting the mechanisms from:

* IGP Flex-Algo [I-D.ietf-lsr-flex-algo]
* Inter-Domain policies
  [I-D.ietf-spring-segment-routing-central-epe]

Furthermore, these can be combined with ODN/AS (On Demand Nexthop/Automated Steering) [I-D.ietf-spring-segment-routing-policy] for automated slice provisioning and traffic steering.

Further details on how these tools can be used to create end to end network slices are documented in [I-D.ali-spring-network-slicing-building-blocks].

9. Control Plane Considerations

This document focuses on user-plane behavior and its independence from the control plane. While the SRv6 mobile user-plane behaviors may be utilized in emerging architectures, such as [I-D.gundavelli-dmm-mfa], [I-D.mhkk-dmm-srv6mup-architecture] for example, require control plane support for the user-plane, this document does not impose any change to the existent mobility control plane.
Section 11 allocates SRv6 Segment Endpoint Behavior codepoints for the new behaviors defined in this document.

10. Security Considerations

The security considerations for Segment Routing are discussed in [RFC8402]. More specifically for SRv6 the security considerations and the mechanisms for securing an SR domain are discussed in [RFC8754]. Together, they describe the required security mechanisms that allow establishment of an SR domain of trust to operate SRv6-based services for internal traffic while preventing any external traffic from accessing or exploiting the SRv6-based services.

The technology described in this document is applied to a mobile network that is within the SR Domain.

This document introduces new SRv6 Endpoint Behaviors. Those behaviors do not need any special security consideration given that it is deployed within that SR Domain.

11. IANA Considerations

The following values have been allocated within the "SRv6 Endpoint Behaviors" [RFC8986] sub-registry belonging to the top-level "Segment Routing Parameters" registry:

<table>
<thead>
<tr>
<th>Value</th>
<th>Hex</th>
<th>Endpoint behavior</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>40</td>
<td>0x0028</td>
<td>End.MAP</td>
<td>[This.ID]</td>
</tr>
<tr>
<td>41</td>
<td>0x0029</td>
<td>End.Limit</td>
<td>[This.ID]</td>
</tr>
<tr>
<td>69</td>
<td>0x0045</td>
<td>End.M.GTP6.D</td>
<td>[This.ID]</td>
</tr>
<tr>
<td>70</td>
<td>0x0046</td>
<td>End.M.GTP6.Di</td>
<td>[This.ID]</td>
</tr>
<tr>
<td>71</td>
<td>0x0047</td>
<td>End.M.GTP6.E</td>
<td>[This.ID]</td>
</tr>
<tr>
<td>72</td>
<td>0x0048</td>
<td>End.M.GTP4.E</td>
<td>[This.ID]</td>
</tr>
</tbody>
</table>

Table 1: SRv6 Mobile User-plane Endpoint Behavior Types
12. Acknowledgements

The authors would like to thank Daisuke Yokota, Bart Peirens, Ryokichi Onishi, Kentaro Ebisawa, Peter Bosch, Darren Dukes, Francois Clad, Sri Gundavelli, Sridhar Bhaskaran, Arashmid Akhavain, Ravi Shekhar, Aeneas Dodd-Noble, Carlos Jesus Bernardos, Dirk v. Hugo and Jeffrey Zhang for their useful comments of this work.

13. Contributors

Kentaro Ebisawa
Toyota Motor Corporation
Japan
Email: ebisawa@toyota-tokyo.tech

Tetsuya Murakami
Arrcus, Inc.
United States of America
Email: tetsuya.ietf@gmail.com

14. References

14.1. Normative References

[I-D.ietf-spring-segment-routing-policy]


14.2. Informative References


[I-D.ietf-spring-sr-service-programming]
Clad, F., Xu, X., Filsfils, C., Bernier, D., Li, C.,
Decraene, B., Ma, S., Yadlapalli, C., Henderickx, W., and
S. Salsano, "Service Programming with Segment Routing",
Work in Progress, Internet-Draft, draft-ietf-spring-sr-
service-programming-05, 10 September 2021,
<https://datatracker.ietf.org/doc/html/draft-ietf-spring-sr-
service-programming-05>.

[I-D.kohno-dmm-srv6mob-arch]
Kohno, M., Clad, F., Camarillo, P., and Z. Ali,
"Architecture Discussion on SRv6 Mobile User plane",
Work in Progress, Internet-Draft, draft-kohno-dmm-srv6mob-arch-
05, 8 November 2021,
<https://datatracker.ietf.org/doc/html/draft-kohno-dmm-
srv6mob-arch-05>.

[I-D.matsushima-spring-srv6-deployment-status]
Matsushima, S., Filsfils, C., Ali, Z., Li, Z., Rajaraman,
K., and A. Dhamija, "SRv6 Implementation and Deployment
Status",
Work in Progress, Internet-Draft, draft-
matsushima-spring-srv6-deployment-status-15, 5 April 2022,
<https://datatracker.ietf.org/doc/html/draft-matsushima-
spring-srv6-deployment-status-15>.

[I-D.mhkk-dmm-srv6mup-architecture]
Matsushima, S., Horiba, K., Khan, A., Kawakami, Y.,
Murakami, T., Patel, K., Kohno, M., Kamata, T., Garvia, P.
C., Voyer, D., Zadok, S., Meilik, I., Agrawal, A.,
Perumal, K., and J. Horn, "Segment Routing IPv6 Mobile
User Plane Architecture for Distributed Mobility
Management",
Work in Progress, Internet-Draft, draft-
mhkk-dmm-srv6mup-architecture-03, 20 March 2022,
<https://datatracker.ietf.org/doc/html/draft-mhkk-dmm-
srv6mup-architecture-03>.

[I-D.murakami-dmm-user-plane-message-encoding]
Murakami, T., Matsushima, S., Ebisawa, K., Camarillo, P.,
and R. Shekhar, "User Plane Message Encoding",
Work in Progress, Internet-Draft, draft-murakami-dmm-user-plane-
message-encoding-05, 5 March 2022,
<https://datatracker.ietf.org/doc/html/draft-murakami-dmm-
user-plane-message-encoding-05>.

[TS.29281] 3GPP, "General Packet Radio System (GPRS) Tunnelling
Protocol User Plane (GTPv1-U)", 3GPP TS 29.281 15.1.0,
December 2017.
Appendix A. Implementations

This document introduces new SRv6 Endpoint Behaviors. These behaviors have an open-source P4 implementation available in https://github.com/ebiken/p4srv6.

Additionally, a full implementation of this document is available in Linux Foundation FD.io VPP project since release 20.05. More information available here: https://docs.fd.io/vpp/20.05/d7/d3c/srv6_mobile_plugin_doc.html.

There are also experimental implementations in M-CORD NGIC and Open Air Interface (OAI).

Authors’ Addresses

Satoru Matsushima (editor)
SoftBank
Japan
Email: satoru.matsushima@g.softbank.co.jp

Clarence Filsfils
Cisco Systems, Inc.
Belgium
Email: cf@cisco.com

Miya Kohno
Cisco Systems, Inc.
Japan
Email: mkohno@cisco.com

Pablo Camarillo Garvia (editor)
Cisco Systems, Inc.
Spain
Email: pcamaril@cisco.com

Daniel Voyer
Bell Canada
Canada
Email: daniel.voyer@bell.ca
Mobility aware Transport Network Slicing for 5G
draft-ietf-dmm-tn-aware-mobility-03

Abstract

This document specifies a framework and mapping of slices in 5G mobile systems to transport network slices in IP, Layer 2 or Layer 1 transport networks. Slices in 5G systems are characterized by latency bounds, reservation guarantees, jitter, data rates, availability, mobility speed, usage density, criticality and priority. These characteristics are mapped to transport network slice include bandwidth, latency and criteria such as isolation, directionality and disjoint routes. Mobile slice criteria are mapped to the appropriate transport slice and capabilities offered in backhaul, midhaul and fronthaul connectivity segments between radio side network functions and user plane function(gateway).

This document describes how a mobile network slice is mapped to a slice in IP or Layer 2 transport network between 3GPP provisioning end points. The same mapping mechanisms apply during initial UE session setup and following UE mobility. Applicability of this framework and underlying transport networks, which can enable different slice properties are also discussed. This is based on mapping between mobile and transport underlays (L2, Segment Routing, IPv6, MPLS and IPv4).

Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC2119 [RFC2119].
Internet-Draft Mobility aware Transport Network Slicing for 5G  March 2022

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 6, 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction ............................................. 3
   1.1. IETF Network Slicing Terminology .................. 4
   1.2. Problem Statement ................................. 5
   1.3. Solution Approach ................................ 5
   1.4. Acronyms .......................................... 5
2. Transport and Slice aware Mobility in 5G Networks .......... 7
   2.1. Backhaul and Mid-Haul Transport Network .......... 8
      2.1.1. IETF Network Slicing Applicability .......... 10
      2.1.2. Fronthaul Transport Network ................. 10
   2.2. Mobile Transport Network Context (MTNC) and Scalability 10
   2.3. Transport Network Function (TNF) .................. 11
   2.4. Transport Provisioning ............................ 12
   2.5. MTNC-ID in the Data Packet ........................ 13
   2.6. Functionality for E2E Management ................. 15
1. Introduction

The 3GPP architecture for 5GS defined in [TS.23.501-3GPP], [TS.23.502-3GPP] and [TS.23.503-3GPP] for 5GC (5G Core) and the NG-RAN architecture and procedures defined in [TS.38.300-3GPP] and [TS.38.401-3GPP] include procedures for setting up network slices in the 3GPP network.

3GPP has defined slice types to cover enhanced mobile broadband (eMBB) communications, ultra-reliable low latency communications (URLLC) and massive internet of things (mIoT) and may extend to include new slice types as needed. ATIS [ATIS075] has defined an additional slice type for V2X services. 3GPP slicing and RAN aspects are further described Appendix A.1.

3GPP slice types and multiple instances of a slice type satisfy various characteristics for 5G network resources. A slice in 3GPP is a logical chunk of 3GPP network resources that is dynamically created and may include core network control and user plane functions as well as access network functions. A slice instance that spans user plane network functions including the UPF (User Plane Function), gNB-CU (generalized Node-B Centralized Unit) and gNB-DU (generalized Node-B Distributed Unit) and its interfaces N3, N9, F1-U) are clearly defined, however:

- 3GPP standards do not specify the underlying IP transport network capabilities or slices thereof.
- Though 3GPP standards define how interfaces N3, N9, F1-U are reselected following mobility but do not specify the underlying transport network reselection aspects following mobility.
Slice details in 3GPP, ATIS or NGMN do not specify how slice characteristics for QoS, hard/soft isolation, protection and other aspects should be satisfied in IP transport networks.

IP transport is used to interconnect the data forwarding entities, UPFs, gNB-CU and gNB-DU in the 5GC and NG-RAN architecture but 3GPP specifications only define the interfaces (N3, N9, F1U etc.) and the 3GPP transport end points [TS.28.541-3GPP]. The architecture allows the flexibility to dynamically place Branching Point (BP) and Uplink Classifier (ULCL) UPFs closer to the access network (5G-AN). The 5G-AN can be a radio access network (NG-RAN) or any non-3GPP access network, for example, WLAN. The resources of gNB-CU and gNB-DU corresponding to a slice in a 5G radio access network (NG-RAN) may be interconnected using a Layer 2 or IP network transport.

A transport underlay across 3GPP interfaces N3, N9 and F1U may use multiple technologies or network providers on path and the slice in 3GPP domain should have a corresponding mapping in the transport domain. This document proposes to map a slice in the 3GPP domain to a transport domain slice. Key considerations including simplicity (e.g., use of L2 VLAN), routed networks on path (i.e., IP based mapping), efficiency of inspecting the slice mapping parameter and others are described in subsequent sections.

1.1. IETF Network Slicing Terminology

[I-D.ietf-teas-ietf-network-slices] draft defines the ‘IETF Network slice’, its scope and characteristics. It lists use cases where IETF technologies can be used for slicing solutions, for various connectivity segments. Transport slice terminology as used in this document refers to the connectivity segment between various 5G systems (i.e. 5G-AN which includes NG-RAN, ULCL UPF, BP UPF and PSA UPF) and some of these segments are referred to as IETF Network slices.

[I-D.ietf-teas-ietf-network-slices] also defines a generic framework and how abstract requests to set up slices can be mapped to more specific technologies (e.g., VPN and traffic-engineering technologies). This document is aimed to be specific to 3GPP use case where many such connectivity segments are used in E2E slicing solutions. Some of the terminologies defined in these referred drafts and applicability to this document are further described in Section 2.1.1.
1.2. Problem Statement

- The 5G System (5GS) as defined in 3GPP specifications, does not consider the resources and functionalities needed from the transport network for the path between UPF, gNB corresponding to the N3, N9 and F1-U interfaces. The lack of underlying Transport Network (TN) awareness in 3GPP may lead to selection of sub-optimal UPF(s) and/or 5G-AN during various procedures in 5GS (e.g., session establishment and various mobility scenarios).

- Meeting the specific slice characteristics on the F1-U, N3 and N9 interfaces depends on the IP transport underlay providing these resources and capabilities. There should also be a means by which 3GPP slices are mapped to corresponding transport network slices and the means to carry this mapping in N3, N9, F1-U packets over the transport network. This is needed to meet SLAs for real-time, mission-critical or latency sensitive services.

- 3GPP defines configuration for its transport end-points in [TS.28.541-3GPP]. These end-points may be for Layer 2 alternatives such as VLAN or L3/routed networks on the F1-U, N3 or N9 path based on desired capabilities. When L3/routed networks and IP transport are used, IP header fields like DSCP are not sufficient as they convey QoS but not the other aspects like isolation or protection. Other fields and extensions to carry a slice mapping, simplicity of lookup, implementation and scale are discussed in Section 2.

1.3. Solution Approach

This document specifies an approach to fulfill the needs of 5GS to transport user plane traffic from 5G-AN (including NG-RAN) to UPF in an optimized fashion. This is done by keeping establishment and mobility procedures aware of the underlying transport network along with slicing requirements.

Section 2 describes in detail on how TN aware mobility can be built irrespective of underlying TN technology used. How other IETF TE technologies applicable for this draft is specified in Section 3.2.

1.4. Acronyms

- 5QI - 5G QoS Indicator
- 5G-AN - 5G Access Network
- AMF - Access and Mobility Management Function (5G)
BP       -  Branch Point (5G)
CSR      -  Cell Site Router
CP       -  Control Plane (5G)
CU       -  Centralized Unit (5G, gNB)
DN       -  Data Network (5G)
DU       -  Distributed Unit (5G, gNB)
eMBB     -  enhanced Mobile Broadband (5G)
FRR      -  Fast ReRoute
gNB      -  5G NodeB
GBR      -  Guaranteed Bit Rate (5G)
GTP-U    -  GPRS Tunneling Protocol - User plane (3GPP)
IGP      -  Interior Gateway Protocols (e.g. IS-IS, OSPFv2, OSPFv3)
LFA      -  Loop Free Alternatives (IP FRR)
mIOT     -  Massive IOT (5G)
MPLS     -  Multi Protocol Label Switching
NG-RAN   -  Next Generation Radio Access Network (i.e., gNB, NG-eNB - RAN functions which connect to 5GC)
NSSMF    -  Network Slice Selection Management Function
QFI      -  QoS Flow ID (5G)
PPR      -  Preferred Path Routing
PDU      -  Protocol Data Unit (5G)
PW       -  Pseudo Wire
RAN      -  Radio Access Network (i.e. 3GPP radio access network used synonymously with NG-RAN in this document)
RAN      -  Radio Access Network
2. Transport and Slice aware Mobility in 5G Networks

3GPP architecture [TS.23.501-3GPP], [TS.23.502-3GPP] describe slicing in 5GS and is provided here for information. The application of 5GS slices in transport network for backhaul, mid-haul and front haul are not explicitly covered in 3GPP and is the topic here. To support specific characteristics in backhaul (N3, N9), mid-haul (F1) and front haul, it is necessary to map and provision corresponding resources in the transport domain. This section describes how to provision the mapping information in the transport network and apply it so that user plane packets can be provided the transport resources (QoS, isolation, protection, etc.) expected by the 5GS slices.

The figure shows the entities on path for 3GPP Network Functions (gNB-DU, gNB-CU, UPF) to obtain slice aware classification from an IP/L2 transport network.
Figure 1: Backhaul and Mid-haul Transport Network for 5G

2.1. Backhaul and Mid-Haul Transport Network

Figure 1 depicts IP Xhaul network with SDN-C and PE (Provider Edge) routers providing IP transport service to 5GS user plane entities 5G-AN (e.g. gNB) and UPF. 5GS architecture with high level management, control and user plane entities and its interaction with the IP transport plane is shown here. The slice capability required in IP transport networks are estimated and provisioned by the functionality as specified in Section 2.3 (TNF) with support from various other control plane functions such as the Network Data Analytics Function (NWDAF), Network Function Repository Function (NRF) and Policy Control Function (PCF). The TNF is only a logical function that may be realized in a 3GPP management function such as Network Slice Selection Management Function (NSSMF) defined in [TS.28.533-3GPP].
The TNF requests the SDN-C to provision the IP XHaul network using ACTN [RFC8453].

The 5G management plane in Figure 1 interacts with the 5G control plane – the 5GC (5G Core), gNB-CU (5G NodeB Centralized Unit) and gNB-DU (5G Node B Distributed Unit). Non-access stratum (NAS) signaling from the UE for session management, mobility is handled by the 5GC. When a UE initiates session establishment, it indicates the desired slice type in the S-NSSAI (Specific Network Slice Selection Assistance Information) field. The AMF uses the S-NSSAI, other subscription information and configuration in the NSSF to select the appropriate SMF and the SMF in turn selects UPFs (User Plane Functions) that are able to provide the specified slice resources and capabilities.

The AMF, SMF, NSSF, PCF, NRF, NWDAF and other control functions in 5GC are described in [TS.23.501-3GPP]. Some of the slice capabilities along the user plane path between the (R)AN and UPFs (F1-U, N3, N9 segments) such as a low latency path, jitter, protection and priority needs these to be provided by the IP transport network.

The 5G user plane from UE to DN (Data Network) includes a mid-haul segment (F1-U between gNB DU(UP), gNB CU(UP)) and backhaul (N3 between gNB - UPF; N9 between UPFs). If the RAN uses lower layer split architecture as specified by O-RAN alliance, then the user plane path from UE to DN also includes the fronthaul interface. The fronthaul interface carries the radio frames in the form of In-phase (I) and Quadrature (Q) samples using eCPRI encapsulation over Ethernet or UDP over IP.

The N3, N9 and F1-U user planes use GTP-U [TS.29.281-3GPP] to transport UE PDUs (IPv4, IPv6, IPv4v6, Ethernet or Unstructured). For the front-haul described further in Section 2.1.2, an Ethernet transport with VLANs can be expected to be the case in many deployments.

Figure 1 also depicts the PE router, where transport paths are initiated/terminated can be deployed separately with UPF or both functionalities can be in the same node. The TNF provisions this in the SDN-C of the IP XHaul network using ACTN [RFC8453]. When a GTP-U encapsulated user packet from the (R)AN (gNB) or UPF with the slice information traverses the F1-U/N3/N9 segment, the PE router of the IP transport underlay can inspect the slice information and provide the provisioned capabilities. This is elaborated further in Section 2.4.
2.1.1. IETF Network Slicing Applicability

Some of the functional elements depicted in the Figure 1 can be mapped to the terminology set forth in the [I-D.ietf-teas-ietf-network-slices]. From 3GPP perspective, UE and UPF are the network slice endpoints and routers, gNB-DU, gNB-CU, switches, PE nodes are the slice realization endpoints. The TNF represented in the Figure 1 can be seen as IETF Network Slice Controller (NSC) functionality and SDN-C maps to Network Controller (NC). NSC-NBI interface is the interface from 3GPP Management plane (e.g., NSSMF) and NSC-SBI interface is the interface between TNF and SDN-C. Various possibilities for implementation of these interfaces and the relation to ACTN are also further described in the [I-D.ietf-teas-ietf-network-slices].

2.1.2. Fronthaul Transport Network

The O-RAN Alliance has specified the fronthaul interface between the O-RU and the O-DU in [ORAN-WG4.CUS-O-RAN]. The radio layer information, in the form of In-phase (I) and Quadrature (Q) samples are transported using Enhanced Common Public Radio Interface (eCPRI) framing over Ethernet or UDP. On the Ethernet based fronthaul interface, the slice information can be carried in the Ethernet header through the VLAN tags. The Ethernet switches in the fronthaul transport network can inspect the slice information (VLAN tag) in the Ethernet header and provide the provisioned capabilities. The mapping of I and Q samples of different radio resources (radio resource blocks or carriers etc.,.) to different slices and to their respective VLAN tags on the fronthaul interface is controlled by the O-RAN fronthaul C-Plane and M-Plane interfaces. On a UDP based fronthaul interface, the slice information can be carried in the IP or UDP header. The PE routers of the fronthaul transport network can inspect the slice information in the IP or UDP header and provide the provisioned capabilities. The fronthaul transport network is latency and jitter sensitive. The provisioned slice capabilities in the fronthaul transport network MUST take care of the latency and jitter budgets of the specific slice for the fronthaul interface. The provisioning of the fronthaul transport network is handled by the SDN-C pertaining to the fronthaul transport.

2.2. Mobile Transport Network Context (MTNC) and Scalability

The MTNC represents a slice of a transport path for a tenant between two 3GPP user plane functions. The Mobile-Transport Network Context Identifier (MTNC-ID) is generated by the TNF to be unique for each instance (for a tenant) and per traffic class (including QoS and slice aspects). Thus, there may be more than one MTNC-ID for the same QoS and instance if there is a need to provide isolation (slice)
of the traffic. It should be noted that MTNC are per class/instance and not per user session. The MTNC-IDs are configured by the TNF to be unique within a provisioning domain.

Since the MTNC-IDs are generated per instance / tenant, there is no need for unique MTNC-IDs per flow/session. In addition, since the traffic estimation is performed prior to UE’s session establishment, there is no provisioning delay experienced by the UE during its session setup. For an instance/tenant, the MTNC-ID space scales roughly as a square of the number sites between which 3GPP user plane functions have paths. If there are T traffic classes and C Tenants, the number of MTNC-IDs in a fully meshed network is T * C. An MTNC-ID space of 16 bits (65K identifiers) can be expected to be sufficient.

2.3. Transport Network Function (TNF)

Figure 1 shows a view of the functions and interfaces for provisioning the MTNC-IDs. The focus is on provisioning between the 3GPP management plane (NSSMF), transport network (SDN-C) and carrying the MTNC-IDs in PDU packets for the transport network to grant the provisioned resources.

In Figure 1, the TNF (logical functionality within the NSSMF) requests the SDN-C in the transport domain to program the TE path using ACTN [RFC8453]. The SDN-C programs the Provider Edge (PE) routers and internal routers according to the underlay transport technology (e.g., MPLS, SR, PPR). The PE router inspects incoming PDU data packets for the UDP SRC port which mirrors the MTNC-ID, classifies and provides the VN service provisioned across the transport network.

The detailed mechanisms by which the NSSMF provides the MTNC-IDs to the control plane and user plane functions are for 3GPP to specify. Two possible options are outlined below for completeness. The NSSMF may provide the MTNC-IDs to the 3GPP control plane by either providing it to the Session Management Function (SMF), and the SMF in turn provisions the user plane functions (UP-NF1, UP-NF2) during PDU session setup. Alternatively, the user plane functions may request the MTNC-IDs directly from the TNF/NSSMF. Figure 1 shows the case where user plane entities request the TNF/NSSMF to translate the Request and get the MTNC-ID. Another alternative is for the TNF to provide a mapping of the 3GPP Network Instance Identifier, described in Section 2.6 and the MTNC-ID to the user plane entities via configuration.

The TNF should be seen as a logical entity that can be part of NSSMF in the 3GPP management plane [TS.28.533-3GPP]. The NSSMF may use
network configuration, policies, history, heuristics or some combination of these to derive traffic estimates that the TNF would use. How these estimates are derived are not in the scope of this document. The focus here is only in terms of how the TNF and SDN-C are programmed given that slice and QoS characteristics across a transport path can be represented by an MTNC-ID. The TNF requests the SDN-C in the transport network to provision paths in the transport domain based on the MTNC-ID. The TNF is capable of providing the MTNC-ID provisioned to control and user plane functions in the 3GPP domain. Detailed mechanisms for programming the MTNC-ID should be part of the 3GPP specifications.

2.4. Transport Provisioning

Functionality of transport provisioning for an engineered IP transport that supports 3GPP slicing and QoS requirements in [TS.23.501-3GPP] is described in this section.

During a PDU session setup, the AMF using input from the NSSF selects a network slice and SMF. The SMF with user policy from Policy Control Function (PCF) sets 5QI (QoS parameters) and the UPF on the path of the PDU session. While QoS and slice selection for the PDU session can be applied across the 3GPP control and user plane functions as outlined in Section 2, the IP transport underlay across F1-U, N3 and N9 segments do not have enough information to apply the resource constraints represented by the slicing and QoS classification. Current guidelines for interconnection with transport networks [IR.34-GSMA] provide an application mapping into DSCP. However, these recommendations do not take into consideration other aspects in slicing like isolation, protection and replication.

IP transport networks have their own slice and QoS configuration based on domain policies and the underlying network capability. Transport networks can enter into an agreement for virtual network services (VNS) with client domains using the ACTN [RFC8453] framework. An IP transport network may provide such slice instances to mobile network operators, CDN providers or enterprises for example. The 3GPP mobile network, on the other hand, defines a slice instance for UEs as are the mobile operator’s ‘clients’. The Network Slice Selection Management Function (NSSMF) [TS 28.533] that interacts with a TN controller like an SDN-C (that is out of scope of 3GPP).

The ACTN VN service can be used across the IP transport networks to provision and map the slice instance and QoS of the 3GPP domain to the IP transport domain. An abstraction that represents QoS and slice instances in the mobile domain and mapped to ACTN VN service in the transport domain is represented here as MTNC-IDs. Details of how
the MTNC-IDs are derived are up to functions that can estimate the level of traffic demand.

The 3GPP network/5GS provides slices instances to its clients (UE) that include resources for radio and mobile core segments. The UE’s PDU session spans the access network (radio) and F1-U/N3/N9 transport segments which have an IP transport underlay. The 5G operator needs to obtain slice capability from the IP transport provider since these resources are not seen by the 5GS. Several UE sessions that match a slice may be mapped to an IP transport segment. Thus, there needs to be a mapping between the slice capability offered to the UE (NSSAI) and what is provided by the IP transport.

When the 3GPP user plane function (5G-AN, UPF) does not terminate the transport underlay protocol (e.g., MPLS), it needs to be carried in the IP protocol header from end-to-end of the mobile transport connection (N3, N9). [I-D.ietf-dmm-5g-uplane-analysis] discusses these scenarios in detail.

### 2.5. MTNC-ID in the Data Packet

When the 3GPP user plane function (5G-AN, UPF) and transport provider edge is on different nodes, the PE router needs to have the means by which to classify the PDU packet. The mapping information is provisioned between the 5G provider and IP transport network and corresponding information should be carried in each IP packet on the F1-U, N3, N9 interface. To allow the IP transport edge nodes to inspect the transport context information efficiently, it should be carried in an IP header field that is easy to inspect. It may be noted that the F1-U, N3 and N9 interfaces in 5GS are IP interfaces. If the fronthaul, midhaul or backhaul IP path is bounded by an L2 network, one option may be to use VLANs to carry the MTNC-ID. 3GPP specifications for management plane defines transport end-points configuration in [TS.28.541-3GPP]. However, Layer 2 alternatives such as VLAN will fail in L3/routed networks on the F1-U, N3 or N9 path. GTP-U (F1-U, N3, N9 encapsulation header) field extensions offer a possibility, however these extensions are not always easy for a transport edge router to parse efficiently on a per packet basis. Other IP header fields like DSCP are not suitable as it only conveys the QoS aspects (but not other aspects like isolation, protection, etc.)

While IPv6 extension headers like SRv6 may be an option to carry the MTNC-ID that requires the end-to-end network to be IPv6 as well as the capability to lookup the extension header at line rate. To minimise the protocol changes and make this underlay transport independent (IPv4/IPv6/MPLS/L2), an option is to provision a mapping of MTNC-ID to a UDP port range of the GTP encapsulated user packet.
A simple mapping table between the MTNC-ID and the source UDP port number can be configured to ensure that ECMP /load balancing is not affected adversely by encoding the UDP source port with an MTNC-ID mapping. The UDP port information containing MTNC-ID is a simple extension that can be provisioned in 3GPP transport end-points defined in [TS.28.541-3GPP]. This mapping is configured in 3GPP user plane functions (5G-AN, UPF) and Provider Edge (PE) Routers that process MTNC-IDs.

PE routers can thus provision a policy based on the source UDP port number (which reflects the mapped MTNC-ID) to the underlying transport path and then deliver the QoS/slice resource provisioned in the transport network. The source UDP port that is encoded is the outer IP (corresponding to GTP-U header) while the inner IP packet (UE payload) is unaltered. The source UDP port is encoded by the node that creates the GTP-U encapsulation and therefore, this mechanism has no impact on UDP checksum calculations.

3GPP network operators may use IPSec gateways (SEG) to secure packets between two sites - for example over an F1-U, N3 or N9 segment. The MTNC identifier in the GTP-U packet should be in the outer IP source port even after IPSec encryption for PE transport routers to inspect and provide the level of service provisioned. Tunnel mode - which is the case for SEG/IPSec gateways - adds an outer IP header in both AH (Authenticated Header) and ESP (Encapsulated Security Payload) modes. The GTP-U / UDP source port with encoded MTNC identifier should be copied to the IPSec tunnel ESP header. One option is to use 16 bits from the SPI field of the ESP header to encode the MTNC identifier and use the remaining 16 bits in SPI field to identify an SA. Load balancing entropy for ECMP will not be affected as the MTNC encoding mechanism already accounts for this.

If the RAN uses O-RAN Alliance lower layer split architecture, then a fronthaul network is involved. On an Ethernet based fronthaul transport network, VLAN tag may be an option to carry the MTNC-ID. The VLAN ID provides a 12 bit space and is sufficient to support up to 4096 slices on the fronthaul transport network. The mapping of fronthaul traffic to corresponding network slices is based on the radio resource for which the fronthaul carries the I and Q samples. The mapping of fronthaul traffic to the VLAN tag corresponding to the network slice is specified in Section 2.1.2. On the UDP based fronthaul transport network, the UDP source port can be used to carry the MTNC-ID.
2.6. Functionality for E2E Management

With the TNF functionality in 5GS Service Based Interface, the following additional functionalities are required for end-to-end slice management including the transport network:

- The Specific Network Slice Selection Assistance Information (S-NSSAI) of PDU session SHOULD be mapped to the assigned transport VPN and the TE path information for that slice.

- For transport slice assignment for various SSTs (eMBB, URLLC, MIoT) corresponding underlay paths need to be created and monitored from each transport endpoint (CSR and PE@UPF).

- During PDU session creation, apart from radio and 5GC resources, transport network resources needed to be verified matching the characteristics of the PDU session traffic type.

- The TNF MUST provide an API that takes as input the source and destination 3GPP user plane element address, required bandwidth, latency and jitter characteristics between those user plane elements and returns as output a particular TE path’s identifier, that satisfies the requested requirements.

- Mapping of PDU session parameters to underlay SST paths need to be done. One way to do this is to let the SMF install a Forwarding Action Rule (FAR) in the UPF via N4 with the FAR pointing to a "Network Instance" in the UPF. A "Network Instance" is a logical identifier for an underlying network. The "Network Instance" pointed by the FAR can be mapped to a transport path (through L2/L3 VPN). FARs are associated with Packet Detection Rule (PDR). PDRs are used to classify packets in the uplink (UL) and the downlink (DL) direction. For UL procedures specified in Section 2.4, Section 2.5 can be used for classifying a packet belonging to a particular slice characteristic. For DL, at a PSA UPF, the UE IP address is used to identify the PDU session, and hence the slice a packet belongs to and the IP 5 tuple can be used for identifying the flow and QoS characteristics to be applied on the packet at UPF. If a PE is not co-located at the UPF then mapping to the underlying TE paths at PE happens based on the encapsulated GTP-U packet as specified in Section 2.5.

- In some SSC modes [I-D.chunduri-dmm-5g-mobility-with-ppr], if segmented path (CSR to PE@staging/ULCL/BP-UPF to PE@anchor-point-UPF) is needed, then corresponding path characteristics MUST be used. This includes a path from CSR to PE@UL-CL/BP UPF [TS.23.501-3GPP] and UL-CL/BP UPF to eventual UPF access to DN.
Continuous monitoring of the underlying transport path characteristics should be enabled at the endpoints (technologies for monitoring depend on traffic engineering technique used as described in Section 3.2). If path characteristics are degraded, reassignment of the paths at the endpoints should be performed. For all the affected PDU sessions, degraded transport paths need to be updated dynamically with similar alternate paths.

During UE mobility events similar to 4G/LTE i.e., gNB mobility (F1 based, Xn based or N2 based), for target gNB selection, apart from radio resources, transport resources MUST be factored. This enables handling of all PDU sessions from the UE to target gNB and this require co-ordination of gNB, AMF, SMF with the TNF module.

Integrating the TNF as part of the 5GS Service Based Interfaces, provides the flexibility to control the allocation of required characteristics from the TN during a 5GS signaling procedure (e.g., PDU Session Establishment). If TNF is seen as separate and in a management plane, this real time flexibility is lost. Changes to detailed signaling to integrate the above for various 5GS procedures as defined in [TS.23.502-3GPP] is beyond the scope of this document.

3. Transport Network Underlays

Apart from the various flavors of IETF VPN technologies to share the transport network resources and capacity, TE capabilities in the underlay network is an essential component to realize the 5G TN requirements. This section focuses on various transport underlay technologies (not exhaustive) and their applicability to realize Midhaul/Backhaul transport networks. Focus is on the user/data plane i.e., F1-U/N3/N9 interfaces as laid out in the framework Figure 1.

3.1. Applicability

For 3 different SSTs, 3 transport TE paths can be signaled from any node in the transport network. For Uplink traffic, the 5G-AN will choose the right underlying TE path of the UPF based on the S-NSSAI the PDU Session belongs to and/or the UDP Source port (corresponds to the MTNC-ID Section 2.4) of the GTP-U encapsulation header. Similarly in the Downlink direction matching Transport TE Path of the 5G-AN is chosen based on the S-NSSAI the PDU Session belongs to. The table below shows a typical mapping:
<table>
<thead>
<tr>
<th>GTP/UDP SRC PORT</th>
<th>SST in S-NSSAI</th>
<th>Transport Path Info</th>
<th>Transport Path Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Range Xx - Xy</td>
<td></td>
<td>MIOT</td>
<td>GBR (Guaranteed Bit Rate)</td>
</tr>
<tr>
<td>X1, X2 (discrete</td>
<td></td>
<td>PW ID/VPN info,</td>
<td>Bandwidth: Bx</td>
</tr>
<tr>
<td>values)</td>
<td></td>
<td>TE-PATH-A</td>
<td>Delay: Dx</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Jitter: Jx</td>
</tr>
<tr>
<td>Range Yx - Yy</td>
<td></td>
<td>URLLC (ultra-low</td>
<td>GBR with Delay Req.</td>
</tr>
<tr>
<td>Y1, Y2 (discrete</td>
<td></td>
<td>latency)</td>
<td>Bandwidth: By</td>
</tr>
<tr>
<td>values)</td>
<td></td>
<td>PW ID/VPN info,</td>
<td>Delay: Dy</td>
</tr>
<tr>
<td></td>
<td></td>
<td>TE-PATH-B</td>
<td>Jitter: Jy</td>
</tr>
<tr>
<td>Range Zx - Zy</td>
<td></td>
<td>EMBB (broadband)</td>
<td>Non-GBR</td>
</tr>
<tr>
<td>Z1, Z2 (discrete</td>
<td></td>
<td>PW ID/VPN info,</td>
<td>Bandwidth: Bx</td>
</tr>
<tr>
<td>values)</td>
<td></td>
<td>TE-PATH-C</td>
<td></td>
</tr>
</tbody>
</table>

Figure 2: Mapping of Transport Paths on F1-U/N3/N9

- It is possible to have a single TE Path for multiple input points through a MP2P TE tree structure separate in UL and DL direction.

- Same set of TE Paths are created uniformly across all needed 5G-ANs and UPFs to allow various mobility scenarios.

- Any modification of TE parameters of the path, replacement path and deleted path needed to be updated from TNF to the relevant ingress points. Same information can be pushed to the NSSF, and/or SMF as needed.

- TE Paths support for native L2, IPv4 and IPv6 data/user planes with optional TE features are desirable in some network segments. As this is an underlay mechanism it can work with any overlay encapsulation approach including GTP-U as defined currently for F1-U/N3/N9 interface.

In some E2E scenarios, security is desired granularly in the underlying transport network. In such cases, there would be a need to have separate sub-ranges under each SST to provide the TE path in preserving the security characteristics. The UDP Source Port range
captured in Figure 2 would be sub-divided to maintain the TE path for the current SSTs with the security. The current solution doesn’t provide any mandate on the UE traffic in selecting the type of security.

3.2. Transport Network Technologies

While there are many Software Defined Networking (SDN) approaches available, this section is not intended to list all the possibilities in this space but merely captures the technologies for various requirements discussed in this document.

RSVP-TE [RFC3209] provides a lean transport overhead for the TE path for MPLS user plane. However, it is perceived as less dynamic in some cases and has some provisioning overhead across all the nodes in N3 and N9 interface nodes. Also, it has another drawback with excessive state refresh overhead across adjacent nodes and this can be mitigated with [RFC8370].

SR-TE [RFC8402] does not explicitly signal bandwidth reservation or mechanism to guarantee latency on the nodes/links on SR path. But SR allows path steering for any flow at the ingress and particular path for a flow can be chosen. Some of the issues and suitability for mobile use plane are documented at Section 5.3 of [I-D.bogineni-dmm-optimized-mobile-user-plane]. However, [I-D.ietf-dmm-srv6-mobile-uplane] presents various options for optimized mobile user plane with SRv6 with or without GTP-U overhead along with traffic engineering capabilities. SR-MPLS allows reduction of the control protocols to one IGP (without needing for LDP and RSVP-TE).

Preferred Path Routing (PPR) is an integrated routing and TE technology and the applicability for this framework is described in [I-D.chunduri-dmm-5g-mobility-with-ppr]. PPR does not remove GTP-U, unlike some other proposals laid out in [I-D.bogineni-dmm-optimized-mobile-user-plane]. Instead, PPR works with the existing cellular user plane (GTP-U) for F1-U/N3 and N9. In this scenario, PPR will only help provide TE benefits needed for 5G slices from a transport domain perspective. It does so for any underlying user/data plane used in the transport network (L2/IPv4/IPv6/MPLS).

As specified with the integrated transport network function (TNF), a particular RSVP-TE path for MPLS or SR path for MPLS and IPv6 with SRH user plane or PPR with PPR-ID [I-D.chunduri-dmm-5g-mobility-with-ppr], can be supplied to SMF for mapping a particular PDU session to the transport path.
4. Acknowledgements

Thanks to Young Lee for discussions on this document including ACTN applicability for the proposed TNF. Thanks to Sri Gundavelli, Kausik Majumdar and 3GPP delegates who provided detailed feedback on this document.

5. IANA Considerations

This document has no requests for any IANA code point allocations.

6. Security Considerations

This document does not introduce any new security issues.

7. Contributing Authors

The following people contributed substantially to the content of this document and should be considered co-authors.

Richard Li
Futurewei
2330 Central Expressway
Santa Clara
CA 95050
USA
Email: richard.li@futurewei.com

Luis M. Contreras
Telefonica
Sur-3 building, 3rd floor
Madrid 28050
Spain
Email: luismiguel.contrerasmurillo@telefonica.com

Xavier De Foy
InterDigital Communications, LLC
1000 Sherbrooke West
Montreal
Canada

Email: Xavier.Defoy@InterDigital.com

8. References

8.1. Normative References


8.2. Informative References


Appendix A. New Control Plane and User Planes

A.1. Slicing Framework and RAN Aspects

The 3GPP architecture defines slicing aspects where the Network Slice Selection Function (NSSF) assists the Access Mobility Manager (AMF) and Session Management Function (SMF) to assist and select the right entities and resources corresponding to the slice requested by the User Equipment (UE). The User Equipment (UE) indicates information regarding the set of slices it wishes to connect, in the Network Slice Selection Assistance Information (NSSAI) field during network registration procedure (Attach) and the specific slice the UE wants to establish an IP session, in the Specific NSSAI (S-NSSAI) field during the session establishment procedure (PDU Session Establishment). The AMF selects the right SMF and the SMF in turn selects the User Plane Functions (UPF) so that the QoS and capabilities requested can be fulfilled.

The architecture for the Radio Access Network (RAN) is defined in [TS.38.300-3GPP] and [TS.38.401-3GPP]. The 5G RAN architecture allows disaggregation of the RAN into a Distributed Unit (DU) and a Centralized Unit (CU). The CU is further split into control plane (CU-CP) and user plane (CU-UP). The interface between CU-UP and the DU for the user plane traffic is called the F1-U and between the CU-CP and DU for the control plane traffic is called the F1-C. The F1-C and the F1-U together are called the mid-haul interfaces. The DU does not have a CP/UP split. Apart from 3GPP, O-RAN Alliance has specified further disaggregation of the RAN at the lower layer (physical layer). The DU is disaggregated into a ORAN DU (O-DU) which runs the upper part of the physical layer, MAC and RLC and the ORAN Radio Unit (O-RU) which runs the lower part of the physical layer. The interface between the O-DU and the O-RU is called the Fronthaul interface and is specified in [ORAN-WG4.CUS-O-RAN].
A.2. Slice aware Mobility: Discrete Approach

In this approach transport network functionality from the 5G-AN to UPF is discrete and 5GS is not aware of the underlying transport network and the resources available. Deployment specific mapping function is used to map the GTP-U encapsulated traffic at the 5G-AN (e.g. gNB) in UL and UPF in DL direction to the appropriate transport slice or transport Traffic Engineered (TE) paths. These TE paths can be established using RSVP-TE [RFC3209] for MPLS underlay, SR [RFC3209] for both MPLS and IPv6 underlay or PPR [I-D.chunduri-dmm-5g-mobility-with-ppr] with MPLS, IPv6 with SRH, native IPv6 and native IPv4 underlays.

As per [TS.23.501-3GPP] and [TS.23.502-3GPP] the SMF controls the user plane traffic forwarding rules in the UPF. The UPFs have a concept of a "Network Instance" which logically abstracts the underlying transport path. When the SMF creates the packet detection rules (PDR) and forwarding action rules (FAR) for a PDU session at the UPF, the SMF identifies the network instance through which the packet matching the PDR has to be forwarded. A network instance can be mapped to a TE path at the UPF. In this approach, TNF as shown in Figure 1 need not be part of the 5G Service Based Interface (SBI). Only management plane functionality is needed to create, monitor, manage and delete (life cycle management) the transport TE paths/transport slices from the 5G-AN to the UPF (on N3/N9 interfaces). The management plane functionality also provides the mapping of such TE paths to a network instance identifier to the SMF. The SMF uses this mapping to install appropriate FARs in the UPF. This approach provide partial integration of the transport network into 5GS with some benefits.

One of the limitations of this approach is the inability of the 5GS procedures to know, if underlying transport resources are available for the traffic type being carried in PDU session before making certain decisions in the 5G CP. One example scenario/decision could be, a target 5G-AN selection during a N2 mobility event, without knowing if the target 5G-AN is having a underlay transport slice resource for the S-NSSAI and 5QI of the PDU session. The Integrated approach specified below can mitigate this.

Authors' Addresses
Problems and Requirements of Satellite Constellation for Internet
draft-lhan-problems-requirements-satellite-net-02

Abstract

This document presents the detailed analysis about the problems and requirements of satellite constellation used for Internet. It starts from the satellite orbit basics, coverage calculation, then it estimates the time constraints for the communications between satellite and ground-station, also between satellites. How to use satellite constellation for Internet is discussed in detail including the satellite relay and satellite networking. The problems and requirements of using traditional network technology for satellite network integrating with Internet are finally outlined.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 17 August 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.
1. Introduction

Satellite constellation for Internet is emerging. Even there is no constellation network established completely yet at the time of the publishing of the draft (June 2021), some basic internet service has been provided and has demonstrated competitive quality to traditional broadband service.

This memo will analyze the challenges for satellite network used in Internet by traditional routing and switching technologies. It is based on the analysis of the dynamic characters of both ground-station-to-satellite and inter-satellite communications and its impact to satellite constellation networking.

The memo also provides visions for the future solution, such as in routing and forwarding.

The memo focuses on the topics about how the satellite network can work with Internet. It does not focus on physical layer technologies (wireless, spectrum, laser, mobility, etc.) for satellite communication.

2. Terminology

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>LEO</td>
<td>Low Earth Orbit with the altitude from 180 km to 2000 km.</td>
</tr>
<tr>
<td>VLEO</td>
<td>Very Low Earth Orbit with the altitude below 450 km</td>
</tr>
<tr>
<td>MEO</td>
<td>Medium Earth Orbit with the altitude from 2000 km to 35786 km</td>
</tr>
<tr>
<td>GEO</td>
<td>Geosynchronous orbit with the altitude 35786 km</td>
</tr>
<tr>
<td>GSO</td>
<td>Geosynchronous satellite on GEO</td>
</tr>
<tr>
<td>ISL</td>
<td>Inter Satellite Link</td>
</tr>
<tr>
<td>ISLL</td>
<td>Inter Satellite Laser Link</td>
</tr>
<tr>
<td>EIRP</td>
<td>Effective isotropic radiated power</td>
</tr>
<tr>
<td>P2MP</td>
<td>Point to Multiple Points</td>
</tr>
<tr>
<td>GS</td>
<td>Ground Station, a device on ground connecting the</td>
</tr>
</tbody>
</table>
satellite. In the document, GS will hypothetically provide L2 and/or L3 functionality in addition to process/send/receive radio wave. It might be different as the reality that the device to process/send/receive radio wave and the device to provide L2 and/or L3 functionality could be separated.

SGS
Source ground station. For a specified flow, a ground station that will send data to a satellite through its uplink.

DGS
Destination ground station. For a specified flow, a ground station that is connected to a local network or Internet, it will receive data from a satellite through its downlink and then forward to a local network or Internet.

PGW
Packet Gateway

UPF
User Packet Function

PE router
Provider Edge router

CE router
Customer Edge router

P router
Provider router

LSA
Link-state advertisement

LSP
Link-State PDUs

L1
Layer 1, or Physical Layer in OSI model [OSI-Model]

L2
Layer 2, or Data Link Layer in OSI model [OSI-Model]

L3
Layer 3, or Network Layer in OSI model [OSI-Model], it is also called IP layer in TCP/IP model

BGP
Border Gateway Protocol [RFC4271]

eBGP
External Border Gateway Protocol, two BGP peers have different Autonomous Number

iBGP
Internal Border Gateway Protocol, two BGP peers have same Autonomous Number
IGP

Interior gateway protocol, examples of IGPs include Open Shortest Path First (OSPF [RFC2328]), Routing Information Protocol (RIP [RFC2453]), Intermediate System to Intermediate System (IS-IS [RFC7142]) and Enhanced Interior Gateway Routing Protocol (EIGRP [RFC7868]).

3. Overview

The traditional satellite communication system is composed of few GSO and ground stations. For this system, each GSO can cover 42% Earth’s surface [GEO-Coverage], so as few as three GSO can provide the global coverage theoretically. With so huge coverage, GSO only needs to amplify signals received from uplink of one ground station and relay to the downlink of another ground station. There is no inter-satellite communications needed. Also, since the GSO is stationary to the ground station, there is no mobility issue involved.

Recently, more and more LEO and VLEO satellites have been launched, they attract attentions due to their advantages over GSO and MEO in terms of higher bandwidth, lower cost in satellite, launching, ground station, etc. Some organizations [ITU-6G][Surrey-6G][Ntttdocomo-6G] have proposed the non-terrestrial network using LEO, VLEO as important parts for 6G to extend the coverage of Internet. SpaceX has started to build the satellite constellation called StarLink that will deploy over 10 thousand LEO and VLEO satellites finally [StarLink]. China also started to request the spectrum from ITU to establish a constellation that has 12992 satellites [China-constellation]. European Space Agency (ESA) has proposed "Fiber in the sky" initiative to connect satellites with fiber network on Earth [ESA-HydRON].

When satellites on MEO, LEO and VLEO are deployed, the communication problem becomes more complicated than for GSO. This is because the altitude of MEO/LEO/VLEO satellites are much lower. As a result, the coverage of each satellite is much smaller than for GSO, and the satellite is not relatively stationary to the ground. This will lead to:

1. More satellites than GSO are needed to provide the global coverage. Section 4.2 will analyze the coverage area, and the minimum number of satellites required to cover the earth surface.

2. The point-to-point communication between satellite and ground station will not be static. Mobility issue has to be considered. Detailed analysis will be done in Section 5.1.
3. The inter-satellite communication is needed, and all satellites need to form a network. Details are described in Section 5.2.

In addition to above context, Section 7 will address the problem and requirements when satellite constellation is joining Internet.

As the 1st satellite constellation company in history, the SpaceX/StarLink will be inevitably mentioned in the draft. But it must be noted that all information about SpaceX/StarLink in the draft are from public. Authors of the draft have no relationship or relevant inside knowledge of SpaceX/Starlink.

4. Basics of Satellite Constellation

This section will introduce some basics for satellite such as orbit parameters, coverage estimation, minimum number of satellite and orbit plane required, real deployments.

4.1. Satellite Orbit

The orbit of a satellite can be either circular or ecliptic, it can be described by following Keplerian elements [KeplerianElement]:

1. Inclination (i)
2. Longitude of the ascending node (Omega)
3. Eccentricity (e)
4. Semimajor axis (a)
5. Argument of periapsis (omega)
6. True anomaly (nu)

For a circular orbit, two parameters, Inclination and Longitude of the ascending node, will be enough to describe the orbit.

4.2. Coverage of LEO and VLEO Satellites and Minimum Number Required

The coverage of a satellite is determined by many physical factors, such as spectrum, transmitter power, the antenna size, the altitude of satellite, the air condition, the sensitivity of receiver, etc. EIRP could be used to measure the real power distribution for coverage. It is not deterministic due to too many variants in a real environment. The alternative method is to use the minimum elevation angle from user terminals or gateways to a satellite. This is easier and more deterministic. [SpaceX-Non-GEO] has suggested originally
the minimum elevation angle of 35 degrees and deduced the radius of
the coverage area is about 435km and 1230km for VLEO (altitude
335.9km) and LEO (altitude 1150km) respectively. The details about
how the coverage is calculated from the satellite elevation angle can
be found in [Satellite-coverage].

Using this method to estimate the coverage, we can also estimate the
minimum number of satellites required to cover the earth surface.

It must be noted, SpaceX has recently reduced the required minimum
elevation angle from 35 degrees to 25 degrees. The following
analysis still use 35 degrees.

Assume there is multiple orbit planes with the equal angular interval
across the earth surface (The Longitude of the ascending node for
sequential orbit plane is increasing with a same angular interval).
Each orbit plane will have:

1. The same altitude.
2. The same inclination of 90 degree.
3. The same number of satellites.

With such deployment, all orbit planes will meet at north and south
pole. The density of satellite is not equal. Satellite is more
dense in the space above the polar area than in the space above the
equator area. Below estimations are made in the worst covered area,
or the area of equator where the satellite density is the minimum.

Figure 1 illustrates the coverage area on equator area, and each
satellite will cover one hexagon area. The figure is based on plane
geometry instead of spherical geometry for simplification, so, the
orbit is parallel approximately.

Figure 2 shows how to calculate the radius (Rc) of coverage area from
the satellite altitude (As) and the elevation angle (b).
**Figure 1:** Satellite coverage on ground

```
|--- 2*Rc --->|

+ Satellite

| b |

|---|---|---|

| 2*a |

|---|---|---|

| Earth center |

**Figure 2:** Satellite coverage estimation

x The vertical projection of satellite to Earth

Re The radius of the Earth, Re=6378(km)
As  The altitude of a satellite

Rc  The radius (arc length) of the coverage, or, the arc length of hexagon center to its 6 vertices. \( Rc = Re \cdot (a \cdot \pi) / 180 \)

a  The cap angle for the coverage area (the RC arc). \( a = \arccos((Re/(Re+As)) \cdot \cos(b)) - b. \)

b  The least elevation angle that a ground station or a terminal can communicate with a satellite, \( b = 35 \) degree.

Ns  The minimum number of satellites on one orbit plane, it is equal to the number of the satellite’s vertical projection on Earth, so, \( Ns = 180/(a \cdot \cos(30)) \)

No  The minimum number of orbit (with same inclination), it is equal to the number of the satellite orbit’s vertical projection, so, \( No = 360/(a \cdot (1+\sin(30))) \)

For a example of two type of satellite LEO and VEO, the coverages are calculated as in Table 1:

<table>
<thead>
<tr>
<th>Parameters</th>
<th>VLEO1</th>
<th>VLEO2</th>
<th>LEO1</th>
<th>LEO2</th>
</tr>
</thead>
<tbody>
<tr>
<td>As(km)</td>
<td>335.9</td>
<td>450</td>
<td>1100</td>
<td>1150</td>
</tr>
<tr>
<td>a(degree)</td>
<td>3.907</td>
<td>5.078</td>
<td>10.681</td>
<td>11.051</td>
</tr>
<tr>
<td>Rc(km)</td>
<td>435</td>
<td>565</td>
<td>1189</td>
<td>1230</td>
</tr>
<tr>
<td>Ns</td>
<td>54</td>
<td>41</td>
<td>20</td>
<td>19</td>
</tr>
<tr>
<td>No</td>
<td>62</td>
<td>48</td>
<td>23</td>
<td>22</td>
</tr>
</tbody>
</table>

Table 1: Satellite coverage estimation for LEO and VEO examples

4.3. Real Deployment of LEO and VEO for Satellite Network

Obviously, the above orbit parameter setup is not optimal since the sky in the polar areas will have the highest density of satellite.
In the real deployment, to provide better coverage for the areas with denser population, to get redundancy and better signal quality, and to make the satellite distance within the range of inter-satellite communication (2000km [Laser-communication-range]), more than the minimum number of satellites are launched. For example, different orbit planes with different inclination/altitude are used.

Normally, all satellites are grouped by orbit planes, each group has a number of orbit planes and each orbit plane has the same orbit parameters, so, each orbit in the same group will have:

1. The same altitude

2. The same inclination, but the inclination is less than 90 degrees. This will result in the empty coverage for polar areas and better coverage in other areas. See the orbit picture for phrase 1 for [StarLink].

3. The same number of satellites

4. The same moving direction for all satellites

The proposed deployment of SpaceX can be seen in [SpaceX-Non-GEO] for StarLink.

The China constellation deployment and orbit parameters can be seen in [China-constellation].

5. Communications for Satellite Constellation

Unlike the communication on ground, the communication for satellite constellation is much more complicated. There are two mobility aspects, one is between ground-station and satellite, another is between satellites.

In the traditional mobility communication system, only terminal is moving, the mobile core network including base station, front haul and back haul are static, thus an anchor point, i.e., PGW in 4G or UPF in 5G, can be selected for the control of mobility session. Unfortunately, when satellite constellation joins the static network system of Internet on ground, there is no such anchor point can be selected since the whole satellite constellation network is moving.

Another special aspect that can impact the communication is that the fast moving speed of satellite will cause frequent changes of communication peers and link states, this will make big challenges to the network side for the packet routing and delivery, session control and management, etc.
5.1. Dynamic Ground-station-Satellite Communication

All satellites are moving and will lead to the communication between ground station and satellite can only last a certain period of time. This will greatly impact the technologies for the satellite networking. Below illustrates the approximate speed and the time for a satellite to pass through its covered area.

In Table 2, VLEO1 and LEO3 have the lowest and highest altitude respectively, VLEO2 is for the highest altitude for VLEO. We can see that longest communication time of ground-station-satellite is less than 400 seconds, the longest communication time for VLEO ground-station-satellite is less than 140 seconds.

The "longest communication time" is for the scenario that the satellite will fly over the receiver ground station exactly above the head, or the ground station will be on the diameter line of satellite coverage circular area, see Figure 1.

Re  The radius of the Earth, Re=6378(km)
As  The altitude of a satellite
AL  The arc length(in km) of two neighbor satellite on the same orbit plane, AL=2*cos(30)*(Re+As)*(a*pi)/180
SD  The space distance(in km) of two neighbor satellite on the same orbit plane, SD=2*(Re+As)*sin(AL/(2*(Re+As))).
V   the velocity (in m/s) of satellite, V=sqrt(G*M/(Re+As))
G   Gravitational constant, G=6.674*10^(-11)(m^3/(kg*s^2))
M   Mass of Earth, M=5.965*10^24 (kg)
T   The time (in second) for a satellite to pass through its cover area, or, the time for the station-satellite communication. T=ALs/V
Table 2: The time for the ground-station-satellite communication

5.2. Dynamic Inter-satellite Communication

5.2.1. Inter-satellite Communication Overview

In order to form a network by satellites, there must be an inter-satellite communication. Traditionally, inter-satellite communication uses the microwave technology, but it has following disadvantages:

1. Bandwidth is limited and only up to 600M bps [Microwave-vs-Laser-communication].
2. Security is a concern since the microwave beam is relatively wide and it is easy for 3rd party to sniff or attack.
3. Big antenna size.
4. Power consumption is high.
5. High cost per bps.

Recently, laser is used for the inter-satellite communication, it has following advantages, and will be the future for inter-satellite communication.

1. Higher bandwidth and can be up to 10G bps [Microwave-vs-Laser-communication].
2. Better security since the laser beam size is much narrower than microwave, it is harder for sniffing.

3. The size of optical lens for laser is much smaller than microwave’s antenna size.

4. Power saving compared with microwave.

5. Lower cost per bps.

The range for satellite-to-satellite communications has been estimated to be approximately 2,000 km currently [Laser-communication-range].

From Table 2, we can see the Space Distance (SD) for some LEO (altitude over 1100km) are exceeding the ceiling of the range of laser communication, so, the satellite and orbit density for LEO need to be higher than the estimation values in the Table 1.

Assume the laser communication is used for inter-satellite communication, then we can analyze the lifetime of inter-satellite communication when satellites are moving. The Figure 3 illustrates the movement and relative position of satellites on three orbits. The inclination of orbit planes is 90 degrees.

```
+ North pole
 |\  
 | s |
 s | s
 / s \
 s | s
 | s1 |
 s4 | s6
 | s2 |
 \\
 s5 | s7
 | s3 |
 s | s
 \ s / \
 s | s
 | s |
 \|/  
 + South pole
```

Figure 3: Satellite movement

There are four scenarios:
1. For satellites within the same orbit
   The satellites in the same orbit will move to the same direction
   with the same speed, thus the interval between satellites is
   relatively steady. Each satellite can communicate with its front
   and back neighbor satellite as long as satellite’s orbit is
   maintained in its life cycle. For example, in Figure 3, s2 can
   communicate with s1 and s3.

2. For satellites between neighbor orbits in the same group at
   non-polar areas
   The orbits for the same group will share the same orbit altitude
   and inclination. So, the satellite speed in different orbit are
   also same, but the moving direction may be same or different.
   Figure 4 illustrates this scenario. When the moving direction is
   the same, it is similar to the scenario 1, the relative position
   of satellites in different orbit are relatively steady as long as
   satellite’s orbit is maintained in its life cycle. When the
   moving direction is different, the relative position of
   satellites in different orbit are un-steady, this scenario will
   be analyzed in more details in Section 5.2.2.

3. For satellites between neighbor orbits in the same group at
   polar areas
   For satellites between neighbor orbits with the same speed and
   moving direction, the relative position is steady as described in
   #2 above, but the steady position is only valid at areas other
   than polar area. When satellites meet in the polar area, the
   relative position will change dramatically. Figure 5 shows two
   satellites meet in polar area and their ISL facing will be
   swapped. So, if the range of laser pointing angle is 360 degrees
   and tracking technology supports, the ISL will not be flipping
   after passing polar area; Otherwise, the link will be flipping
   and inter-satellite communication will be interrupted.

4. For satellites between different orbits in the different group
   The orbits for the different group will have different orbit
   altitude, inclination and speed. So, the relative position of
   satellite is not static. The inter-satellite communication can
   only last for a while when the distance between two satellite is
   within the limit of inter-satellite communication, that is 2000km
   for laser [Laser-communication-range], this scenario will be
   analyzed in more details in Section 5.2.3
* The total number of orbit planes are $N$
* The number (i-1, i, i+1,...) represents the Orbit index
* The bottom numbers (i-1, i, i+1) are for orbit planes on which satellites (S1, S2, S3) are moving from bottom to up.
* The top numbers (i+N/2, i+1+N/2, i+2+N/2) are for orbit planes on which satellites (S4, S5, S6) are moving from up to bottom.

Figure 4: Two satellites with same altitude and inclination (i) move in the same or opposite direction

```
    \ /  \ /  \ /  \ /
   P3  P4
    \ /  \ /  \ /  \ /
     \    \    \    \    
      P1  P2
   /      \      /      \      /      \
```

* Two satellites S1 and S2 are at position P1 and P2 at time T1
* S1’s right facing ISL connected to S2’s left facing ISL
* S1 and S2 move to the position P4 and P3 at time T2
* S1’s left facing ISL connected to S2’s right facing ISL

Figure 5: Two satellites meeting in the polar area will change its facing of ISL

5.2.2. Satellites on Adjacent Orbit Planes with Same Altitude

For satellites on different orbit planes with same altitude, the estimation of the lifetime when two satellite can communicate are as follows.

Figure 6 illustrates a general case that two satellites move and intersect with an angle $A$. 
Figure 6: Two satellites (speed vector $V_1$ and $V_2$) intersect with angle $A$

More specifically, for orbit planes with the inclination angle $i$, Figure 7 illustrates two satellites move in the opposite direction and intersect with an angle $2i$.

Figure 7: Two satellites with same altitude and inclination ($i$) intersect with angle $A=2i$

Follows are the math to calculate the lifetime of communication. Table 3 are the results using the math for two satellites with different altitudes and different inclination angles.

- $D_l$ The laser communication limit, $D_l=2000$km
- $A$ The angle between two orbit’s vertical projection on Earth. $A=2i$
- $V_1$ The speed vector of satellite on orbit1
- $V_2$ The speed vector of satellite on orbit2

$|V|$ the magnitude of the difference of two speed vector $V_1$ and $V_2$, $|V|=|V_1-V_2| = \sqrt{((V_1-V_2\cos(A))^2+(V_2\sin(A))^2)}$. For satellites with the same altitude and inclination angle $i$, $V_1=V_2$, so, $|V|=V_1\sqrt{2-2\cos(2i)}=2V_1\sin(i)$
The lifetime two satellites can communicate, or the time of two satellites' distance is within the range of communication, \( T = \frac{2 \cdot D_l}{|V|} \).

| i (degree) | 80  | 80  | 65  | 65  | 50  | 50  |
| Alt (km)   | 500 | 800 | 500 | 800 | 500 | 800 |
| |V| (km/s) | 14.98 | 14.67 | 13.79 | 13.5 | 11.66 | 11.41 |
| T(s)       | 267 | 273 | 290 | 296 | 343 | 350 |

Table 3: The lifetime of communication for two LEOs (with two altitudes and three inclination angles)

5.2.3. Satellites on Adjacent Orbit Planes with Different Altitude

For satellites on different orbit planes with different altitude, the estimation of the lifetime when two satellite can communicate are as follows.

Figure 8 illustrates two satellites (with the altitude difference \( D_a \)) move and intersect with an angle \( A \).

\[ \text{Follows are the math to calculate the lifetime of communication} \]

\( D_l \) The laser communication limit, \( D_l=2000 \text{km} \)

\( D_a \) Altitude difference (in km) for two orbit planes
A   The angle between two orbit’s vertical projection on Earth

V1  The speed vector of satellite on orbit 1

V2  The speed vector of satellite on orbit 2

\[ |V| = |V1 - V2| = \sqrt{(V1 - V2 \cdot \cos(A))^2 + (V2 \cdot \sin(A))^2} \]

T   The lifetime two satellites can communicate, or the time of two satellites’ distance is within the range of communication, 

\[ T = \frac{2 \cdot \sqrt{D_1^2 - D_a^2}}{|V|} \]

Using formulas above, below is the estimation for the life of communication of two satellites when they intersect. Table 4 and Table 5 are for two VLEOs with the difference of 114.1km for altitude. (VLEO1 and VLEO2 on Table 2). Table 6 and Table 7 are for two LEOs with the difference of 175km for altitude (LEO2 and LEO3 on Table 2).

| Parameters   | VLEO1 | VLEO2 |
|--------------+-------+-------|
| As (km)      | 335.9 | 450   |
| V (km/s)     | 7.7   | 7.636 |

Table 4: Two VLEO with different altitude and speed

| A  (degree) | 0 | 10 | 45 | 90 | 135 | 180 |
|-------------+---+----+----+----+-----+-----|
| |V| (km/s) | 0.065 | 1.338 | 5.869 | 10.844 | 14.169 | 15.336 |
| T(s) | 61810 | 2984 | 680 | 368 | 282 | 260 |

Table 5: Two VLEO intersects with different angle and the life of communication
6. Use Satellite Network for Internet

Since there is no complete satellite network established yet, all following analysis is based on the predictions from the traditional GEO communication. The analysis also learnt how other type of network has been used in Internet, such as Broadband access network, Mobile access network, Enterprise network and Service Provider network.

As a criteria to be part of Internet, any device connected to any satellite should be able to communicate with any public IP4 or IPv6 address in Internet. There could be three types of methods to deliver IP packet from source to destination by satellite:

1. Data packet is relayed between ground station and satellite. For this method, there is no inter-satellite communication and networking. Data packet is bounced once or couple times between ground stations and satellites until the packet arrives at the destination in Internet.

2. Data packet is delivered by inter-satellite networking. For this method, the data packet traverses with multiple satellites and inter-satellite networking is used to deliver the packet to the destination in Internet.
3. Both satellite relay and inter-satellite networking are used. For this method, the data packet is relayed in some segments and traverse with multiple satellites in other segments. It is a combination of the method 1 and method 2.

Using the above methods, follows are typical deployment scenarios that a Satellite network is integrated with Internet:

1. The end user terminal access Internet through satellite relay (Figure 9 for one satellite relay, Figure 10 for multiple satellite relay).

2. The end user terminal access Internet through inter-satellite networking (Figure 11).

3. The local network access Internet through satellite relay (Figure 12 for one satellite relay, Figure 13 for multiple satellite relay).

4. The local network access Internet through inter-satellite networking (Figure 14).

5. The End user terminal or local network access Internet through satellite network and Mobile Access Network, From mobile access network to satellite network or From satellite network to mobile access network, Satellite network includes inter satellite network and relay network (Figure 15 for mobile access network to satellite network, Figure 16 for satellite netowrk to mobile access network).

Figure 9: End user terminal access Internet through one satellite relay

\[\begin{array}{c}
S1----\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\\\--\n\end{array}\]

Figure 10: End user terminal access Internet through multiple satellite relay
Figure 10: End user terminal access Internet through multiple satellite relay

```
S1-----S2-----S3--
 |       |       |
 T---GW--GS1--S4----S5----S6----GS2------PE Internet +
 |       |       |
 \-------S7----S8----S9/
```

Figure 11: End user terminal access Internet through intersatellite-networking

```
\-------S1--
 |       |
 + Local network CE------GS1--S4--GS2------PE Internet +
 |       |
 \-------S7/
```

Figure 12: Local network access Internet through one satellite relay

```
\-------S1----S4--
 |       |
 + Local network CE----GS1--S2--GS2--S5--GS3------PE Internet +
 |       |
 \-------S3/ \----S6/
```

Figure 13: Local network access Internet through multiple satellite relay

```
\-------S1----S2-----S3--
 |       |       |
 + Local network CE------GS1--S4----S5----S6----GS2------PE Internet +
 |       |       |
 \-------S7----S8----S9/
```

Figure 14: Local network access Internet through inter-satellite-networking

```
+--------------+    +-------------+    +---------+
|    T or      |    |Mobile Access|    |Satellite|    |Internet|
| Local network |----+  Network    +----+ Network +----+
+--------------+    +-------------+    +---------+
```

Figure 15: End user terminal or local network access Internet through Mobile Access Network and Satellite Network
In above Figure 9 to Figure 16, the meaning of symbols are as follows:

T               The end user terminal
GW              Gateway router
GS1, GS2, GS3   Ground station with L2/L3 routing/switch functionality.
S1 to S9        Satellites
PE              Provider Edge Router
CE              Customer Edge Router

7. Problems and Requirements for Satellite Constellation for Internet

As described in Section 6, satellites in a satellite constellation can either relay internet traffic or multiple satellites can form a network to deliver internet traffic. More detailed analysis are in following sub sections. There might have multiple solutions for each method described in Section 6, following contexts only discuss the most plausible solution from networking perspectives.

Section 7.1 will list the common problems and requirements for both satellite relay and satellite networking.

Section 7.2 and Section 7.3 will describe key problems, requirement and potential solution from the networking perspective for these two cases respectively.

7.1. Common Problems and Requirements

For both satellite relay and satellite networking, satellite-ground-station must be used, so, the problems and requirements for the satellite-ground-station communication is common and will apply for both methods.
When one satellite is communicating with ground station, the satellite only needs to receive data from uplink of one ground station, process it and then send to the downlink of another ground station. Figure 9 illustrates this case. Normally microwave is used for both links.

Additionally, from the coverage analysis in Section 4.2 and real deployment in Section 4.3, we can see one ground station may communicate with multiple satellites. Similarly, one satellite may communicate with multiple ground stations. The characters for satellite-ground-station communication are:

1. Satellite-ground-station communication is P2MP. Since microwave physically is the carrier of broadcast communication, one satellite can send data while multiple ground stations can receive it. Similarly, one ground station can send data and multiple satellites can receive it.

2. Satellite-ground-station communication is in open space and not secure. Since electromagnetic fields for microwave physically are propagating in open space. The satellite-ground-station communication is also in open space. It is not secure naturally.

3. Satellite-ground-station communication is not steady. Since the satellite is moving with high speed, from Section 5.1, the satellite-ground-station communication can only last a certain period of time. The communication peers will keep changing.

4. Satellite-to-Satellite communication is not steady. For some satellites, even they are in the same altitude and move in the same speed, but they move in the opposite direction, from Section 5.2.2, the satellite-to-satellite communication can only last a certain period of time. The communication peers will keep changing.

5. Satellite-to-Satellite distance is not steady. For satellites with the same altitude and same moving direction, even their relative position is steady, but the distance between satellites are not steady. This will lead to the inter-satellite-communication’s bandwidth and latency keep changing.
6. Satellite physical resource is limited.
Due to the weight, complexity and cost constraint, the physical
resource on a satellite, such as power supply, memory, link
speed, are limited. It cannot be compared with the similar
device on ground. The design and technology used should consider
these factors and take the appropriate approach if possible.

The requirements of satellite-ground-station communication are:

R1. The bi-directional communication capability
Both satellites and ground stations have the bi-directional
communication capability

R2. The identifier for satellites and ground stations
Satellites and ground stations should have Ethernet and/or IP
address configured for the device and each link. More detailed
address configuration can be seen in each solution.

R3. The capability to decide where the IP packet is forwarded to.
In order to send Internet traffic or IP date to destination
correctly, satellites and ground station must have Ethernet hub
or switching or IP routing capability. More detailed capability
can be seen in each solution.

R4. The protocol to establish the satellite-ground-station
communication.
For security and management purpose, the satellite-ground-station
communication is only allowed after both sides agree through a
protocol. The protocol should be able to establish a secured
channel for the communication when a new communication peer comes
up. Each ground station should be able to establish multiple
channels to communicate with multiple satellites. Similarly,
each satellite should be able to establish multiple channels to
communicate to multiple ground stations.

R5. The protocol to discover the state of communication peer.
The discover protocol is needed to detect the state of
communication peer such as peer’s identity, the state of the peer
and other info of the peer. The protocol must be running
securely without leaking the discovered info.

R6. The internet data packet is forwarded securely.
When satellite or ground station is sending the IP packet to its
peer, the packet must be relayed securely without leaking the
user data.
The internet data packet is processed efficiently on satellite

Due to the resource constraint on a satellite, the packet may need more efficient mechanism to be processed on satellite. The process on satellite should be very minimal and offloaded to ground as much as possible.

7.2. Satellite Relay

One of the reasons to use satellite constellation for internet access is it can provide shorter latency than using the fiber underground. But using ISL for inter-satellite communication is the premise for such benefit in latency. Since the ISL is still not mature and adopted commercially, satellite relay is a only choice currently for satellite constellation used for internet access. In [UCL-Mark-Handley], detailed simulations have demonstrated better latency than fiber network by satellite relay even the ISL is not present.

7.2.1. One Satellite Relay

One satellite relay is the simplest method for satellite constellation to provide Internet service. By this method, IP traffic will be relayed by one satellite to reach the DGS and go to Internet.

The solution option and associated requirements are:

S1. The satellite only does L1 relay or the physical signal process.

For this solution, a satellite only receives physical signal, amplify it and broadcast to ground stations. It has no further process for packet, such as L2 packet compositing and processing, etc. All packet level work is done only at ground station. The requirements for the solution are:

R1-1. SGS and BGS are configured as IP routing node. Routing protocol is running in SGS and BGS

SGS and BGS is a IP peer for a routing protocol (IGP or BGP). SGS will send internet traffic to DGS as next hop through satellite uplink and downlink.

R1-2. DGS must be connected with Internet.

DGS can process received packet from satellite and forward the packet to the destination in Internet.

In addition to the above requirements, following problem should be solved:
P1-1.  IP continuity between two ground stations
This problem is that two ground stations are connected by one satellite relay. Since the satellite is moving, the IP continuity between ground stations is interrupted by satellite changing periodically. Even though this is not killing problem from the view point that IP service traditionally is only a best effort service, it will benefit the service if the problem can be solved. Different approaches may exist, such as using hands off protocols, multipath solutions, etc.

S2.  The satellite does the L2 relay or L2 packet process.
For this solution, IP packet is passing through individual satellite as an L2 capable device. Unlike in the solution S1, satellite knows which ground station it should send based on packet’s destination MAC address after L2 processing. The advantage of this solution over S1 is it can use narrower beam to communicate with DGS and get higher bandwidth and better security. The requirements for the solution are:

R2-1.  Satellite must have L2 bridge or switch capability
In order to forward packet to properly, satellite should run some L2 process such as MAC learning, MAC switching. The protocol running on satellite must consider the fast movement of satellite and its impact to protocol convergence, timer configuration, table refreshment, etc.

R2-2.  same as R1-1 in S1
R2-3.  same as R1-2 in S1

In addition to the above requirements, the problem P1-1 for S1 should also apply.

7.2.2.  Multiple Satellite Relay
For this method, packet from SGS will be relayed through multiple intermediate satellites and ground station until reaching a DGS.

This is more complicated than one satellite relay described in Section 7.2.1.

One general solution is to configure both satellites and ground-stations as IP routing nodes, proper routing protocols are running in this network. The routing protocol will dynamically determine forwarding path. The obvious challenge for this solution is that all links between satellite and ground station are not static, according to the analysis in Section 5.1, the lifetime of each link may last
only couple of minutes. This will result in very quick and constant
topology changes in both link state and IP adjacency, it will cause
the distributed routing algorithms may never converge. So this
solution is not feasible.

Another plausible solution is to specify path statically. The path
is composed of a serials of intermediate ground stations plus SGS and
DGS. This idea will make ground stations static and leave the
satellites dynamic. It will reduce the fluctuation of network path,
thus provide more steady service. One variant for the solution is
whether the intermediate ground stations are connected to Internet.
Separated discussion is as below:

S1. Manual configuring routing path and table

For this solution, the intermediate ground stations and DGS are
specified and configured manually during the stage of network
planning and provisioning. Following requirements apply:

R1-1. Specify a path from SGS to DGS via a list of intermediate
ground stations.
The specified DGS must be connected with internet. Other
specified intermediate ground stations does not have to

R1-2. All Ground stations are configured as IP routing node.
Static routing table on all ground stations must be pre-
configured, the next hop of routes to Internet destination in any
ground station is configured to going through uplink of satellite
to the next ground station until reaching the DGS.

R1-3. All Satellites are configured as either L1 relay or L2
relay.
The Satellite can be configured as L1 relay or L2 relay described
in S1 and S2 respectively in Section 7.2.1

In addition to the above requirements, the problem P1-1 in
Section 7.2.1 should also apply.

S2. Automatic decision by routing protocol.

This solution is only feasible after the IP continuity problem (P1-1
in Section 7.2.1) is solved. Following requirements apply:
R2-1. All Ground stations are configured as IP routing node. Proper routing protocols are configured as well. The satellite link cost is configured to be lower than the ground link. In such a way, the next hop of routes for the IP forwarding to Internet destination in any ground station will be always going through the uplink of satellite to the next ground station until reaching the DGS.

R2-2. All Satellites are configured as either L1 relay or L2 relay. The Satellite can be configured as L1 relay or L2 relay described in S1 and S2 respectively in Section 7.2.1

In addition to the above requirements, the problem P1-1 in Section 7.2.1 should also apply.

7.3. Satellite Networking

In the draft, satellite Network is defined as a network that satellites are inter-connected by inter-satellite links (ISL). One of the major difference of satellite network with the other type of network on ground (telephone, fiber, etc.) is its topology and links are not stationary, some new issues have to be considered and solved. Follows are the factors that impact the satellite networking.

7.3.1. L2 or L3 network

The 1st question to answer is should the satellite network be configured as L2 or L3 network? As analyzed in Section 4.2 and Section 4.3, since there are couple of hundred or over ten thousand satellites in a network, L2 network is not a good choice, instead, L3 or IP network is more appropriate for such scale of network.

7.3.2. Inter-satellite-Link Lifetime

If we assume the orbit is circular and ignore other trivial factors, the satellite speed is approximately determined by the orbit altitude as described in the Section 5.1. The satellite orbit can determine if the dynamic position of two satellites is within the range of the inter-satellite communication. That is 2000km for laser communication [Laser-communication-range] by Inter Satellite Laser Link (ISLL).

When two satellites’ orbit planes belong to the same group, or two orbit planes share the same altitude and inclination, and when the satellites move in the same direction, the relative positions of two satellites are relatively stationary, and the inter-satellite communication is steady. But when the satellites move in the
opposite direction, the relative positions of two satellites are not stationary, the communication lifetime is couple of minutes. The Section 5.2.2 has analyzed the scenario.

When two satellites’ orbit planes belong to the different group, or two orbit planes have different altitude, the relative position of two satellite are unstable, and the inter-satellite communication is not steady. As described in Section 5.2, The life of communication for two satellites depends on the following parameters of two satellites:

1. The speed vectors.
2. The altitude difference
3. The intersection angle

From the examples shown in Table 4 to Table 7, we can see that the lifetime of inter-satellite communication for the different group of orbit planes are from couple of hundred seconds to about 18 hours. This fact will impact the routing technologies used for satellite network and will be discussed in Section 7.3.3.

7.3.3. Problems for Traditional Routing Technologies

When the satellite network is integrated with Internet by traditional routing technologies, following provisioning and configuration (see Figure 17) will apply:

1. The ground stations connected to local network and internet are treated as PE router for satellite network (called PE_GS1 and PE_GS2 in the following context), and all satellites are treated as P router.
2. All satellites in the network and ground stations are configured to run IGP.
3. The eBGP is configured between PE_GS and its peered network’s PE or CE.

The work on PE_GS1 are:
* The local network routes are received at PE_GS1 from CE by eBGP. The routes are redistributed to IGP and then IGP flood them to all satellites. (Other more efficient methods, such as iBGP or BGP reflectors are hard to be used, since the satellite is moving and there is no easy way to configure a full meshed iBGP session for all satellites, or configure one satellite as BGP reflector in satellite network.)

* The internet routes are redistributed from IGP to eBGP running on PE_GS1, and eBGP will advertise them to CE.

The work on PE_GS2 are:

* The Internet routes are received at PE_GS2 from PE by eBGP. The routes are redistributed to IGP and then IGP flood them to all satellites. (Similar as in PE_GS1, Other more efficient methods, such as iBGP or BGP reflector cannot be used.)

* The local network routes are redistributed from IGP to eBGP running on PE_GS2, and eBGP will advertise them to Internet.

```
/----------\   S1---S2----S3-----\  /-------\  \
/           \  / IGP domain \  /  \  \
\ Local net CE--eBGP--PE_GS1---S4---S5---PE_GS2--eBGP--PE Internet +  \
 \--------/  /  \
 \-------/ \---S6---S7---S8/ \
```

Figure 17: Local access Internet through inter-satellite-networking

Local access Internet through inter-satellite-networking

On PE-GS1, due to the fact that IGP link between PE_GS1 and satellite is not steady; this will lead to following routing activity:

1. When one satellite is connecting with PE_GS1, the satellite and PE_GS1 form a IGP adjacency. IGP starts to exchange the link state update.

2. The local network routes received by eBGP in PE_GS1 from CE are redistributed to IGP, and IGP starts to flood link state update to all satellites.

3. Meanwhile, the Internet routes learnt from IGP in PE_GS1 will be redistributed to eBGP. eBGP starts to advertise to CE.

4. Every satellite will update its routing table (RIB) and forwarding table (FIB) after IGP finishes the SPF algorithm.
5. When the satellite is disconnecting with PE-GS1, the IGP adjacency between satellite and PE_GS1 is gone. IGP starts to exchange the link state update.

6. The routes of local network and satellite network that were redistributed to IGP in step 2 will be withdrawn, and IGP starts to flood link state update to all satellites.

7. Meanwhile, the Internet routes previously redistributed to eBGP in step 3 will also be withdrawn. eBGP starts to advertise route withdraw to CE.

8. Every satellite will update its routing table (RIB) and forwarding table (FIB) after the SPF algorithm.

Similarly on PE_GS2, due to the fact that IGP link between PE_GS2 and satellite is not steady; this will lead to following routing activity:

1. When one satellite is connecting with PE_GS2, the satellite and PE_GS2 form a IGP adjacency. IGP starts to exchange the link state update.

2. The Internet routes previously received by eBGP in PE_GS2 from PE are redistributed to IGP, IGP starts to flood the new link state update to all satellites.

3. Meanwhile, the routes of local network and satellite network learnt from IGP in PE_GS2 will be redistributed to eBGP. eBGP starts to advertise to Internet peer PE.

4. Every satellite will update its routing table (RIB) and forwarding table (FIB) after IGP finishes the SPF algorithm.

5. When the satellite is disconnecting with PE-GS2, the IGP adjacency between satellite and PE_GS2 is gone. IGP starts to exchange the link state update.

6. The internet routes previously redistributed to IGP in step 2 will be withdrawn, and IGP starts to flood link state update to all satellites.

7. Meanwhile, the routes of local network and satellite network previously redistributed to eBGP in step 3 will also be withdrawn. eBGP starts to advertise route withdraw to PE.

8. Every satellite will update its routing table (RIB) and forwarding table (FIB) after the SPF algorithm.
For the analysis of detailed events above, the estimated time interval between event 1 and 5 for PE_GS1 and PE_GS2 can use the analysis in Section 5.1. For example, it is about 398s for LEO and 103s for VLEO. Within this time interval, the satellite network including all satellites and two ground stations must finish the works from 1 to 4 for PE_GS1 and PE_GS2. The normal internet IPv6 and IPv4 BGP routes size are about 850k v4 routes + 100K v6 routes [BGP-Table-Size]. There are couple critical problems associated with the events:

P1. Frequent IGP update for its link cost
   Even for satellites in different orbit with the steady relative positions, the distance between satellites is keep changing. If the distance is used as the link cost, it means the IGP has to update the link cost frequently. This will make IGP keep running and update its routing table.

P2. Frequent IGP flooding for the internet routes
   Whenever the IGP adjacency changes (step 1 and 5 for PE_GS2), it will trigger the massive IGP flooding for the link state update for massive internet routes learnt from eBGP. This will result in the IGP re-convergency, RIB and FIB update.

P3. Frequent BGP advertisement for the internet routes
   Whenever the IGP adjacency changes (step 3 and 7 for PE_GS1), it will trigger the massive BGP advertisement for the internet routes learnt from IGP. This will result in the BGP re-convergency, RIB and FIB update. BGP convergency time is longer than IGP. The document [BGP-Converge-Time1] has shown that the BGP convergence time varies from 50sec to couple of hundred seconds. The analysis [BGP-Converge-Time2] indicated that per entry update takes about 150us, and it takes o(75s) for 500k routes, or o(150s) for 1M routes.
P4. More frequent IGP flooding and BGP update in whole satellite network

To provide the global coverage, a satellite constellation will have many ground stations deployed. For example, StarLink has applied for the license for up to one million ground stations [StarLink-Ground-Station-Fcc], in which, more than 50 gateway ground stations (equivalent to the PE_GS2) have been registered [SpaceX-Ground-Station-Fcc] and deployed in U.S. [StarLink-GW-GS-map]. It is expected that the gateway ground station will grow quickly to couple of thousands [Tech-Comparison-LEOs]. This means almost each satellite in the satellite network would have a ground station connected. Due to the fact that all satellites are moving, many IGP adjacency changes may occur in a shorter period of time described in Section 5.1 and result in the problem P1 and P2 constantly occur.

P5. Service is not steady

Due to the problems P1 to P3, the service provider of satellite constellation is hard to provide a steady service for broadband service by using inter-satellite network and traditional routing technologies.

As a summary, the traditional routing technology is problematic for large scale inter-satellite networking for Internet. Enhancements on traditional technologies, or new technologies are expected to solve the specific issues associated with satellite networking.

8. IANA Considerations

This memo includes no request to IANA.

9. Contributors

10. Acknowledgements

11. References

11.1. Normative References


11.2. Informative References


[ESA-HydRON]
"HydRON: Fiber in the sky",
<https://www.esa.int/ESA_Multimedia/Videos/2021/04/HydRON_Fibre_in_the_sky>.

[SpaceX-Non-GEO]
"FCC report: SPACEX V-BAND NON-GEOSTATIONARY SATELLITE SYSTEM",

[Satellite-coverage]
Alan R.Washburn, Department of Operations Research, Naval Postgraduate School, "Earth Coverage by Satellites in Circular Orbit",
<https://faculty.nps.edu/awashburn/Files/Notes/EARTHCOV.pdf>.

[Microwave-vs-Laser-communication]
International Journal for Research in Applied Science and Engineering Technology (IJRASET), "Comparison of Microwave and Optical Wireless Inter-Satellite Links",

[Laser-communication-range]
"Interferometric optical communications can potentially lead to robust, secure, and naturally encrypted long-distance laser communications in space by taking advantage of the underlying physics of quantum entanglement."

[BGP-Table-Size]
"BGP in 2020 - BGP table",
<https://blog.apnic.net/2021/01/05/bgp-in-2020-the-bgp-table/>.

[BGP-Converge-Time1]
"BGP in 2020 - BGP Update Churn",

[BGP-Converge-Time2]
"Bringing SDN to the Internet, one exchange point at the time",
Appendix A. Change Log

* Initial version, 07/03/2021

* 01 version, 10/20/2021

Authors’ Addresses

Lin Han (editor)
Futurewei Technologies, Inc.
2330 Central Expy
Santa Clara, CA 95050,
United States of America

Email: lhan@futurewei.com
Richard Li
Futurewei Technologies, Inc.
2330 Central Expy
Santa Clara, CA 95050,
United States of America
Email: rli@futurewei.com

Alvaro Retana
Futurewei Technologies, Inc.
2330 Central Expy
Santa Clara, CA 95050,
United States of America
Email: alvaro.retana@futurewei.com

Meiling Chen
China Mobile
32, Xuanwumen West
BeiJing 100053
China
Email: chenmeiling@chinamobile.com

Li Su
China Mobile
32, Xuanwumen West
BeiJing 100053
China
Email: suli@chinamobile.com

Ning Wang
University of Surrey
Guildford
Surrey, GU2 7XH
United Kingdom
Email: n.wang@surrey.ac.uk
Abstract

This document presents a method to do IP routing over satellite network that consists of LEO (Low Earth Orbit) satellites and ground-stations. The method uses the source routing mechanism. The whole routing info is obtained by path calculation. The routing path information is converted to be a list of instructions and embedded into user packet’s IPv6 extension header. At each hop or each satellite, the routing process engine will forward the packet based on the specified instruction for the satellite. Until the packet reaches the edge of satellite network, or the last satellite, the packet will be sent to a ground station.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 7 September 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document.
Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction .................................................. 3
2. Terminology .................................................... 4
3. Review of LEO satellite constellation for future Internet ..... 5
4. Basics of Instructive Routing ................................ 6
5. IPv6 Routing Header for Instructive Routing .................. 9
6. Instruction List for Instructive Routing ....................... 10
7. Instructive Routing Behaviors ................................ 11
   7.1. Fwd.Inc.Sat_ID .......................................... 12
   7.2. Fwd.Dec.Sat_ID .......................................... 13
   7.3. Fwd.Inc.Opb_ID ......................................... 14
   7.4. Fwd.Dec.Opb_ID ......................................... 15
   7.5. Fwd.Inc.Shl_ID ......................................... 16
   7.6. Fwd.Dec.Shl_ID ......................................... 17
   7.7. End.Intf_ID ............................................ 18
   7.8. End.Punt ................................................ 18
   7.9. End.Lookup ............................................. 19
   7.10. End.Lookup.IPv4 ....................................... 19
   7.11. End.Lookup.IPv6 ....................................... 20
   7.12. Fwd.Sat_Addr .......................................... 20
   7.13. Fwd.Sat_MacAddr ...................................... 21
   7.15. Forwarding_API_SAT(Packet,Input_Satellite,Sat_Addr) ....... 22
   7.16. Forwarding_API_MAC(Packet,Input_Satellite,Sat_MacAddr) .. 22
   7.17. Forwarding_GS_API(Packet,Input_Interface) ............... 23
8. IANA Considerations ........................................... 23
9. Security Considerations ....................................... 23
10. Contributors .................................................. 24
11. Acknowledgements ............................................ 24
12. References ................................................... 24
   12.1. Normative References .................................. 24
   12.2. Informative References ................................ 24
Appendix A. Change Log ........................................... 25
Authors’ Addresses .............................................. 25
1. Introduction

Massive LEO constellation is expected to be used for future Internet. It has raised challenges to the current IP networking technologies to support such super-fast-moving network. [I-D.lhan-problems-requirements-satellite-net] has analyzed the problems when using the regular routing protocols in such network.

Since all satellites in a LEO constellation are well organized and form a kind of multi-layered grid network, each satellite’s relative position in the satellite network will be steady during its lifetime. [I-D.lhan-satellite-semantic-addressing] has proposed to use couple of indexes to identify each satellite in the network. The combination of the indexes is called the satellite semantic address. The semantic address can be embedded into the field of the interface identifier (i.e., the rightmost 64 bits) of the IPv6 address, if IPv6 is used in the satellite network.

This memo proposes a method for routing for satellite network, it is based on the satellite semantic address. The routing information is embedded into the IPv6 packet as routing extension header defined in [RFC8200]. Unlike the segment routing [RFC8754] and programming [RFC8986], The new method will not use IPv6 SID (Segment Identifier) to represent the segments on the routing path. Instead, it will convert the segments on the path to be a list of instructions since each satellite could be represented by the semantic address. Each instruction can tell each satellite how to forward the packet to an adjacent satellite, either on the same orbit, or on the adjacent orbit.

Compared with the traditional IP forwarding, the new method will not use TCAM (Ternary Content-addressable Memory) lookup for IP prefix. Each satellite only needs to store a simple adjacency table. Therefore, the new method can save significant TCAM and the processing time for routing/forwarding tables.

It must be noted this memo just describes one aspect of the whole solution for satellite constellation used for Internet access and NTN (Non-Terrestrial Network) integration with 5G, following areas are not covered in this memo and will be addressed in other documents separately:

1. IP forwarding path calculation for a LEO constellation.
2. Data planes for different scenarios, such as Internet access and NTN integration.
3. Other protocols for control plane.
2. Terminology

LEO
Low Earth Orbit with the altitude from 180 km to 2000 km.

LEO constellation
LEO constellation consists of certain number of LEOs. Each LEO has pre-assigned orbit element.

ISL
Inter Satellite Link

GS
Ground Station, a device on ground connecting satellite. In the document, GS will hypothetically provide L2 and/or L3 functionality in addition to process/transmit/receive radio wave. It might be different as the reality that the device to process/transmit/receive radio wave and the device to provide L2 and/or L3 functionality could be separated.

L2
Layer 2, or Data Link Layer in OSI model [OSI-Model]

L3
Layer 3, or Network Layer in OSI model [OSI-Model], it is also called IP layer in TCP/IP model

OS
Operating System

NTN
Non-Terrestrial Network

SID
Segment Identifier

Sat-GS Links
Wireless links between satellites and ground-stations, it consists of uplink (from ground to satellite) and downlink (from satellite to ground).

Link Metrics
The cost of the outgoing interface for routing, typically, it may indicate the bandwidth, delay or other costs for the interface.

Sat_ID
Satellite Index, the Index for the satellite in an orbit plane, see [I-D.lhan-satellite-semantic-addressing]

Obp_ID
Orbit Plane Index, the Index for the orbit plane in a shell group of satellite, see [I-D.lhan-satellite-semantic-addressing]

Shl_ID
Shell Index, the Index for the shell group of
satellite in a satellite constellation, see [I-D.lhan-satellite-semantic-addressing]

Intf_ID          Interface Index
Sat_Addr         Satellite Semantic Address, it consists of indexes Shl_ID, Obp_ID and Sat_ID. It is 32-bit long and is defined in Section 5.4 in [I-D.lhan-satellite-semantic-addressing]
Sat_MacAddr      The MAC (Media Access Control) Address for a satellite

3. Review of LEO satellite constellation for future Internet

LEO satellite constellation is expected to be integrated with terrestrial network in future Internet. StarLink project [StarLink] has launched its satellites and provided the beta service in some areas. 3GPP [ThreeGPP] has studied the issues when NTN is integrated with Internet and 5G. 3GPP [TR38-821] has also proposed the Satellite-based NG-RAN architectures for NTN integration. The targets of LEO constellation for future Internet and NTN integration are as follows:

1. Global coverage: The Satellite network should cover all places on earth and any flying objects as long as the place or objects are below LEO attitude and within the coverage footprint of satellite constellation, the satellite network should be the complementary to terrestrial network.

2. Internet access: The Satellite network can provide the Internet access service for covered areas.

3. NTN integration: The Satellite network is fully integrated with Internet including Wireless such as 4G or 5G.

4. Competitive service: The Satellite network can provide the services that are competitive to terrestrial network in terms of service stability, Quality of Service, especially the latency for Satellite network is shorter.

As a new form of network, LEO constellation has lots of difference with the steady terrestrial network especially in the mobility. [I-D.lhan-problems-requirements-satellite-net] has analyzed the movement and coverage of satellite. For a massive LEO constellation, all satellites are moving on the allocated orbits, and form one or multiple layers of network. Finally, the massive LEO constellation will have the following unprecedented mobility:
1. Each LEO moves at the speed of 7.0 km/s.

2. Ground Stations move at the speed of 463 m/s due to earth rotation.

3. Half of LEOs move on the direction that is different with another half of LEOs.

4. Huge number of links between satellites and ground-stations, and all of them are constantly flipping within short period of time. All Link Metrics of Sat-GS Links are also constantly changing.

5. All Link Metrics of ISL on the Longitude direction are constantly changing.

6. All Links of ISL on the Longitude direction may be interrupted at two polar areas.

4. Basics of Instructive Routing

When using ISL for satellites in a LEO constellation, each layer of network will have satellite nodes connected by limited ISLs. A typical satellite will have about six ISL to connected to its adjacent satellites in 3D space. Additionally, there might have very few numbers of ISL working as un-steady link to connect to other satellites. Un-stead links are those between satellites moving to different directions, see [I-D.lhan-problems-requirements-satellite-net] for the detailed explanation. After using the semantic address for each satellite, the satellite relationship will be static. Figure 1 illustrates one satellite and its six direct connected adjacent satellites, it is easy to determine some indexes of its adjacent satellites:

1. S0, S1 and S2 have the same Shl_ID, the difference of Obp_ID between S0 and S1, S0 and S2 are both equal to one.

2. S0, S3 and S4 have the same Shl_ID and Obp_ID, the difference of Sat_ID between S0 and S3, S0 and S4 are both equal to one.

3. S0, S5 and S6 have different Shl_ID, and the difference of Shl_ID between S0 and S5, S0 and S6 are both equal to one.

Another benefit to use the semantic address is that the packet forwarding for routing and switching will be simplified significantly. There will be only six major forwarding directions to the directly connected adjacent satellites described above, plus one or few specified directions probably. The specified direction is to forward packet to a specified adjacent satellite through an un-steady
link. The un-steady link can connect to any satellite but only last for a short time. The usage of un-steady links are expected to be limited and are not major scenarios in a LEO constellation. Following are all directions for forwarding:

1. Forward to the Sat_ID Incremental or Decremental directions.
2. Forward to the Obp_ID Incremental or Decremental directions.
3. Forward to the Shl_ID Incremental or Decremental directions.
4. Forward to a specified satellite through an un-steady link.

```
^ Shl_ID Incremental direction
 /    /
S5    ^ Sat_ID Increment direction
 /  /
 /  /
S3    /  /
    /  /
    /  /
S2------S0------S1  -> Obp_ID Increment direction
    /  /
    /  /
    /  /
S4    /  /
    /  /
    /  /
    /  /
    /  /
S6
```

Figure 1: The LEO Satellite Relationship in 3D Space

Figure 2 illustrates a 2D example. It shows how a packet is forwarded in a grid satellite network. The forwarding path consists of a series of segments, and each segment consists of two satellites at its two ends. One segment could be on either the same orbit plane or crossing adjacent orbit plane. Intuitively, we can obtain the list of instructions to guide the packet and get the forwarding behaviors at different satellites. Following is an example:

1. At S1 to S2, forward packet to the Sat_ID Incremental direction, until the packet reaches S2
2. At S2 to S3, forward packet to the Obp_ID Incremental direction, until the packet reaches the orbit plane of S3
3. At S3 to S4, forward packet to the Sat_ID Incremental direction, until the packet reaches S4

4. At S4 to S5, forward packet to the Obp_ID Decremental direction, until the packet reaches the orbit plane of S5

5. At S5 to S6, forward packet to the Sat_ID Decremental direction, until the packet reaches S6

Obviously, at each satellite, the forwarding logic needs to check if the satellite reaches the end of a segment on the route path. In the regular segment routing, the SID is used to do such indication. But for satellite network, since satellite’s semantic address is embedded into the IPv6 address, it is not needed to include the long SID into the packet header. Instead, it will be much saving if we only embed one of three indexes information of the satellite semantic address in the instruction argument, and then we can further simplify the above instructions as:

1. At S1 to S2, forward packet to the Sat_ID Incremental direction, until the packet reaches a satellite and the satellite’s Sad_ID is equal to the given instruction argument (S2’s Satellite Index)

2. At S2 to S3, forward packet to the Obp_ID Incremental direction, until the packet reaches a satellite and the satellite’s Obp_ID is equal to the given instruction argument (S3’s Orbit Plane Index)

3. At S3 to S4, forward packet to the Sat_ID Incremental direction, until the packet reaches a satellite and the satellite’s Sat_ID is equal to the given instruction argument (S4’s Satellite Index)

4. At S4 to S5, forward packet to the Obp_ID Decremental direction, until the packet reaches a satellite and the satellite’s Obp_ID is equal to the given instruction argument (S5’s Orbit Plane Index)

5. At S5 to S6, forward packet to the Sat_ID Decremental direction, until the packet reaches a satellite and the satellite’s Sat_ID is equal to the given instruction argument (S6’s Satellite Index)
5. IPv6 Routing Header for Instructive Routing

For instructive routing, IPv6 routing header is used with a new routing type "Instructive Routing Type". The format of the new routing header is illustrated in Figure 3.

Figure 3: The IPv6 Routing Hdr for Instructive Routing
Routing Type    Instructive Routing Type

Inst. Offset   The offset in the number of octets from the start of Instruction List. The initial value is set to 0 and it points to the 1st instruction to be executed. The value is incremented by the number of octets of the total size of a instruction after the instruction is executed.

Remained Inst. Remained Number of Instructions. The initial value is set to the total number of instructions. The value will be decremented by one after one instruction is executed. The minimum number is one, and it indicates that the end of instruction stack is reached.

ST     The satellite address type, default is 0.

Inst. List    A list of instructions, the size is variable.

Paddings Pad1 or PadN options to make the packet extension header alignment, see [RFC8200]

6. Instruction List for Instructive Routing

For instructive routing, the instruction list is used to instruct each satellite how to do routing job. The format of the instruction list is illustrated in Figure 4. Each instruction consists of Function Code and Arguments.

\begin{verbatim}
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
\--------------/--------------/ \--------------/--------------/
\instruction[0]                   \instruction[1]...
\end{verbatim}

Figure 4: The Instruction List for Instructive Routing

Funct. Code    Function Code, size is 1 octet

Arguments    Arguments for the function, Variable length
7. Instructive Routing Behaviors

The behavior for each satellite for instructive routing is described here. Table 1 is the summary of the name, Hex values of all functions, arguments and size. New functions can be defined if needed.

The subsections below are the detailed explanation for each function.

<table>
<thead>
<tr>
<th>Func Name/Hex Value</th>
<th>Arguments/Size(Octet)</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fwd.Inc.Sat_ID/0X01</td>
<td>Sat_ID/1</td>
<td>Section 7.1</td>
</tr>
<tr>
<td>Fwd.Dec.Sat_ID/0X02</td>
<td>Sat_ID/1</td>
<td>Section 7.2</td>
</tr>
<tr>
<td>Fwd.Inc.Obp_ID/0X03</td>
<td>Obp_ID/1</td>
<td>Section 7.3</td>
</tr>
<tr>
<td>Fwd.Dec.Obp_ID/0X04</td>
<td>Obp_ID/1</td>
<td>Section 7.4</td>
</tr>
<tr>
<td>Fwd.Inc.Shl_ID/0X05</td>
<td>Shl_ID/1</td>
<td>Section 7.5</td>
</tr>
<tr>
<td>Fwd.Dec.Shl_ID/0X06</td>
<td>Shl_ID/1</td>
<td>Section 7.6</td>
</tr>
<tr>
<td>End.Intf_ID/0X07</td>
<td>Intf_ID/1</td>
<td>Section 7.7</td>
</tr>
<tr>
<td>End.Punt/0X08</td>
<td>0X0/1</td>
<td>Section 7.8</td>
</tr>
<tr>
<td>End.Lookup/0X09</td>
<td>0X0/1</td>
<td>Section 7.9</td>
</tr>
<tr>
<td>End.Lookup.IPv4/0X0A</td>
<td>IPv4_Addr/4</td>
<td>Section 7.10</td>
</tr>
<tr>
<td>End.Lookup.IPv6/0X0B</td>
<td>IPv6_Addr/16</td>
<td>Section 7.11</td>
</tr>
<tr>
<td>Fwd.Sat_Addr/0X0C</td>
<td>Sat_Addr/4</td>
<td>Section 7.12</td>
</tr>
<tr>
<td>Fwd.Sat_MacAddr/0X0D</td>
<td>Sat_MacAddr/6</td>
<td>Section 7.13</td>
</tr>
</tbody>
</table>

Table 1: Functions, Arguments and Reference

The functions in Section 7.1 to Section 7.6 are used for the instructions to forward packet to one of the six major directions discussed in Section 4. They will call API in Section 7.14 to forward the packet to the specified direction.
The functions in Section 7.12 and Section 7.13 are used for the instructions to forward packet to a specified adjacent satellite discussed in Section 4. They will call APIs in Section 7.15 and Section 7.16 respectively to forward the packet to the specified adjacent satellite.

In order to forward packet, each satellite should have an adjacency table stored locally; the table should contain the information about all adjacent satellites, it should at least store:

1. Each adjacent satellite’s semantic address.
2. The ID of local interface connecting to each adjacent satellite.
3. The MAC address for the remote interface of each adjacent satellite.

7.1. Fwd.Inc.Sat_ID

The definition of this function is "Forward the packet on the Satellite Index Incremental Direction until the packet reaches a Satellite whose Satellite Index is equal to the value specified in the argument"

This function is used for the instruction to forward packet to one of the six major directions discussed in Section 4.

When a satellite receives a packet with new routing header, assume the satellite indexes in the address are Shl_index, Obp_index, Sat_index respectively, the satellite does the following. During the forwarding, the Forwarding_API in Section 7.14 is called to forward the packet to the specified direction.
When an IRH is processed {
    If ((RI > 1) and (Argument != Sat_index)) {
        Input_Satellite = Current Satellite;
        Input_Direction = Satellite Index Incremental direction;
        Forwarding_API(Packet,Input_Satellite,Input_Direction);
    } else {
        IOF += 2;
        RI --;
        if (RI <= 0)
            Send an ICMP Parameter Problem to the Source Address
            with Code 0 (Erroneous header field encountered)
            and Pointer set to the RI field,
            interrupt packet processing, and discard the packet;
        Proceed to execute the next Instruction;
    }
}

7.2. Fwd.Dec.Sat_ID

The definition of this function is "Forward the packet on the
Satellite Index Decremental Direction until the packet reaches a
Satellite whose Satellite Index is equal to the value specified in
the argument"

This function is used for the instruction to forward packet to one of
the six major directions discussed in Section 4.

When a satellite receives a packet with new routing header, assume
the satellite indexes in the address are Shl_index, Obp_index,
Sat_index respectively, the satellite does the following. During the
forwarding, the Forwarding_API in Section 7.14 is called to forward
the packet to the specified direction.
S01. When an IRH is processed {
S02.   If ((RI > 1) and (Argument != Sat_index)) {
S03.     Input_Satellite = Current Satellite;
S04.     Input_Direction = Satellite Index Decremental direction;
S05.     Forwarding_API(Packet,Input_Satellite,Input_Direction);
S06. } else {
S07.     IOF += 2;
S08.     RI --;
S09.     if (RI <= 0)
            Send an ICMP Parameter Problem to the Source Address
            with Code 0 (Erroneous header field encountered)
            and Pointer set to the RI field,
            interrupt packet processing, and discard the packet;
S10.     Proceed to execute the next Instruction;
S11. } }
S12. }

7.3. Fwd.Inc.Opb_ID

The definition of this function is "Forward the packet on the Orbit Plane Index Incremental Direction until the packet reaches a Satellite whose Orbit Plane Index is equal to the value specified in the argument"

This function is used for the instruction to forward packet to one of the six major directions discussed in Section 4.

When a satellite receives a packet with new routing header, assume the satellite indexes in the address are Shl_index, Obp_index, Sat_index respectively, the satellite does the following. During the forwarding, the Forwarding_API in Section 7.14 is called to forward the packet to the specified direction.
S01. When an IRH is processed {
S02.   If ((RI > 1) and (Argument != Obp_index)) {
S03.      Input_Satellite = Current Satellite;
S04.      Input_Direction = Orbit Plane Index Incremental direction;
S05.      Forwarding_API(Packet,Input_Satellite,Input_Direction);
S06.   } else {
S07.      IOF += 2;
S08.      RI --;
S09.      if (RI <= 0)
          Send an ICMP Parameter Problem to the Source Address
              with Code 0 (Erroneous header field encountered)
              and Pointer set to the RI field,
              interrupt packet processing, and discard the packet;
S10.      Proceed to execute the next Instruction;
S11.   }
S12.}

7.4. Fwd.Dec.Opb_ID

The definition of this function is "Forward the packet on the Orbit Plane Index Decremental Direction until the packet reaches a Satellite whose Orbit Plane Index is equal to the value specified in the argument"

This function is used for the instruction to forward packet to one of the six major directions discussed in Section 4.

When a satellite receives a packet with new routing header, assume the satellite indexes in the address are Shl_index, Obp_index, Sat_index respectively, the satellite does the following. During the forwarding, the Forwarding_API in Section 7.14 is called to forward the packet to the specified direction.
S01. When an IRH is processed {
S02.   If ((RI > 1) and (Argument != Obp_index)) {
S03.      Input_Satellite = Current Satellite;
S04.      Input_Direction = Orbit Plane Index Decremental direction;
S05.      Forwarding_API(Packet,Input_Satellite,Input_Direction);
S06.   } else {
S07.      IOF += 2;
S08.      RI --;
S09.      if (RI <= 0)
           Send an ICMP Parameter Problem to the Source Address
           with Code 0 (Erroneous header field encountered)
           and Pointer set to the RI field,
           interrupt packet processing, and discard the packet;
S10.      Proceed to execute the next Instruction;
S11.   }
S12.}

7.5. Fwd.Inc.Shl_ID

The definition of this function is "Forward the packet on the Orbit Shell Index Incremental Direction until the packet reaches a Satellite whose Orbit Shell Index is equal to the value specified in the argument"

This function is used for the instruction to forward packet to one of the six major directions discussed in Section 4.

When a satellite receives a packet with new routing header, assume the satellite indexes in the address are Shl_index, Obp_index, Sat_index respectively, the satellite does the following. During the forwarding, the Forwarding_API in Section 7.14 is called to forward the packet to the specified direction.
When an IRH is processed {
    If ((RI > 1) and (Argument != Shl_index)) {
        Input_Satellite = Current Satellite;
        Input_Direction = Orbit Shell Index Incremental direction;
        Forwarding_API(Packet,Input_Satellite,Input_Direction);
    } else {
        IOF += 2;
        RI --;
        if (RI <= 0)
            Send an ICMP Parameter Problem to the Source Address
            with Code 0 (Erroneous header field encountered)
            and Pointer set to the RI field,
            interrupt packet processing, and discard the packet;
        Proceed to execute the next Instruction;
    }
}

7.6. Fwd.Dec.Shl_ID

The definition of this function is "Forward the packet on the Orbit Shell Index Decremental Direction until the packet reaches a Satellite whose Orbit Shell Index is equal to the value specified in the argument"

This function is used for the instruction to forward packet to one of the six major directions discussed in Section 4.

When a satellite receives a packet with new routing header, assume the satellite indexes in the address are Shl_index, Obp_index, Sat_index respectively, the satellite does the following. During the forwarding, the Forwarding_API in Section 7.14 is called to forward the packet to the specified direction.
S01. When an IRH is processed {
S02.   If ((RI > 1) and (Argument != Shl_index)) {
S03.      Input_Satellite = Current Satellite;
S04.      Input_Direction = Orbit Shell Index Decremental direction;
S05.      Forwarding_API(Packet,Input_Satellite,Input_Direction);
S06.   } else {
S07.      IOF += 2;
S08.      RI --;
S09.      if (RI <= 0)
           Send an ICMP Parameter Problem to the Source Address
           with Code 0 (Erroneous header field encountered)
           and Pointer set to the RI field,
           interrupt packet processing, and discard the packet;
S10.     Proceed to execute the next Instruction;
S11.   }
S12.}

7.7. End.Intf_ID

The definition of this function is "End of processing for the
Instructive routing, remove the Instructive Routing Header, Forward
the packet to the interface specified in the argument"

This function is normally used on the Dst_Sat to forward packet to
Dst_GS.

When a satellite receives a packet with new routing header, the
satellite does the following, Forwarding_GS_API in Section 7.17 is
called to forward the packet to the specified interface.

S01. When an IRH is processed {
S02.   Change the Next header in the packet header to be
       the Next Header field in the Instructive Routing header;
S03.   Remove the Instructive Routing Header;
S04.   Forwarding_GS_API(Packet, Argument);
S05.}

7.8. End.Punt

The definition of this function is "End of processing for the
Instructive routing, remove the Instructive Routing Header, Punt the
packet to the OS for process"

This function is normally used send packet to a satellite. At the
destination satellite, the packet is punted to the OS to be processed
further.
When a satellite receives a packet with new routing header, the satellite does the following:

S01. When an IRH is processed {
S02.   Change the Next header in the packet header to be the Next Header field in the Instructive Routing header;
S03.   Remove the Instructive Routing Header;
S04.   Punt packet to the local CPU for process;
S05.}

7.9.  End.Lookup

The definition of this function is "End of processing for the Instructive routing, remove the Instructive Routing Header, Lookup the destination address in packet header and forward the packet accordingly"

This function is normally used to send packet to Dst_GS. After the packet reaches the Dst_Sat, the packet is forwarded to Dst_GS by looking up the destination address in the IPv6 packet header.

When a satellite receives a packet with new routing header, the satellite does the following:

S01. When an IRH is processed {
S02.   Change the Next header in the packet header to be the Next Header field in the Instructive Routing header;
S03.   Remove the Instructive Routing Header;
S04.   Lookup the destination address in packet hdr and forward the packet;
S05.}

7.10.  End.Lookup.IPv4

The definition of this function is "End of processing for the Instructive routing, remove the Instructive Routing Header, Lookup the IPv4 address specified in the argument and forward the packet accordingly"

This function is normally used to send packet to Dst_GS. After the packet reaches the Dst_Sat, the packet is forwarded to Dst_GS by looking up the IPv4 destination address specified in the Function Argument.

When a satellite receives a packet with new routing header, the satellite does the following:
S01. When an IRH is processed {
S02.   Fetch the IPv4 addr in the argument;
S03.   Change the Next header in the packet header to be
       the Next Header field in the Instructive Routing header;
S04.   Remove the Instructive Routing Header;
S05.   Lookup the fetched IPv4 address and forward the packet;
S06.}

7.11.  End.Lookup.IPv6

The definition of this function is "End of processing for the
Instructive routing, remove the Instructive Routing Header, Lookup
the IPv6 address specified in the argument and forward the packet
accordingly"

This function is normally used to send packet to Dst_GS. After the
packet reaches the Dst_Sat, the packet is forwarded to Dst_GS by
looking up the IPv6 destination address specified in the Function
Argument.

When a satellite receives a packet with new routing header, the
satellite does the following:

S01. When an IRH is processed {
S02.   Fetch the IPv6 addr in the argument;
S03.   Change the Next header in the packet header to be
       the Next Header field in the Instructive Routing header;
S04.   Remove the Instructive Routing Header;
S05.   Lookup the fetched IPv6 address and forward the packet;
S06.}

7.12.  Fwd.Sat.Addr

The definition of this function is "Forward the packet to the
adjacent satellite with the address specified in the argument"

This function is normally used for the instruction to forward packet
to an adjacent satellite specified by its Satellite Semantic Address.
The Satellite Semantic Address is 32-bit long and is defined in
Section 5.4 in [I-D.lhan-satellite-semantic-addressing]

When a satellite receives a packet with new routing header, assume
the satellite semantic address is Sat_Addr, the satellite does the
following:
S01. When an IRH is processed {
S02.   If ((RI > 1) and (Argument != Sat_Addr)) {
S03.      Input_Satellite = Current Satellite;
S04.      SatAddr = Argument;
S05.      Forwarding_API_SAT(Packet,Input_Satellite,SatAddr);
S06.   } else {
S07.      IOF += 4;
S08.      RI --;
S09.      if (RI <= 0)
               Send an ICMP Parameter Problem to the Source Address
               with Code 0 (Erroneous header field encountered)
               and Pointer set to the RI field,
               interrupt packet processing, and discard the packet.
S10.     Proceed to execute the next Instruction;
S11.   }
S12.}

7.13. Fwd.Sat_MacAddr

The definition of this function is "Forward the packet to the
adjacent satellite with the MAC address specified as the argument"

This function is normally used for the instruction to forward packet
to an adjacent satellite specified by its MAC address.

When a satellite receives a packet with new routing header, assume
the satellite Mac address is Sat_MacAddr, the satellite does the
following:

S01. When an IRH is processed {
S02.   If ((RI > 1) and (Argument != Sat_MacAddr)) {
S03.      Input_Satellite = Current Satellite;
S04.      SatMacAddr = Argument;
S05.      Forwarding_API_Mac(Packet,Input_Satellite,SatMacAddr);
S06.   } else {
S07.      IOF += 6;
S08.      RI --;
S09.      if (RI <= 0)
               Send an ICMP Parameter Problem to the Source Address
               with Code 0 (Erroneous header field encountered)
               and Pointer set to the RI field,
               interrupt packet processing, and discard the packet.
S10.     Proceed to execute the next Instruction;
S11.   }
S12.}

This API will forward a packet to the specified direction. When a
satellite executes the API, it will do following:

S01. Forwarding_API(Packet, Input_Satellite, Input_Direction) {
S02.   Lookup the local adjacency table to find out
     1) The adjacent satellite of "Input_Satellite" on the
direction equal to "Input_Direction" (The adjacent
satellite’s semantic address can be inferred by
the "Input_Satellite" and "Input_Direction").
     2) The L2 address for the adjacent satellite;
     3) The local interface connecting to the adjacent
satellite;
S03.   Rewrite the L2 header of the Packet by the L2 address;
S04.   Send the Packet to the local interface;
S05.}

7.15. Forwarding_API_SAT(Packet, Input_Satellite, Sat_Addr)

This API will forward a packet to the specified adjacent satellite
with the semantic address as the argument. When a satellite executes
the API, it will do following:

S01. Forwarding_API_SAT(Packet, Input_Satellite, SatAddr) {
S02.   Lookup the local adjacency table to find out
     1) The adjacent satellite of "Input_Satellite"
     (The adjacent satellite address is SatAddr);
     2) The L2 address for the adjacent satellite;
     3) The local interface connecting to the adjacent
satellite;
S03.   Rewrite the L2 header of the Packet by the L2 address;
S04.   Send the Packet to the local interface;
S05.}

7.16. Forwarding_API_MAC(Packet, Input_Satellite, Sat_MacAddr)

This API will forward a packet to the specified adjacent satellite
with the MAC address as the argument. When a satellite executes the
API, it will do following:
S01. Forwarding_API_MAC(Packet, Input_Satellite, SatMacAddr) {
S02.   Lookup the local adjacency table to find out
   1) The adjacent satellite of "Input_Satellite"
      (The adjacent satellite MAC address is SatMacAddr);
   2) The L2 address for the adjacent satellite;
   3) The local interface connecting to the adjacent
      satellite;
S03.   Rewrite the L2 header of the Packet by the L2 address;
S04.   Send the Packet to the local interface;
S05.}

7.17. Forwarding_GS_API(Packet, Input_Interface)

This API will forward a packet to ground station the connected to the
specified interface. When a satellite executes the API, it will do
following:

S01. Forwarding_API(Packet, Input_Interface) {
S02.   Lookup the local adjacency table to find out
   1) The connected GS to the interface
      equal to "Input_Interface";
   2) The L2 address for the GS;
S03.   Rewrite the L2 header of the Packet by the L2 address;
S04.   Send the Packet to the "Input_Interface";
S05.}

8. IANA Considerations

This document defines a new IPv6 Routing Type: the "Instructive
Routing Header". It needs to be assigned a number by IANA.

This document also defines an 8-bit Function Name, for which IANA
will create and will maintain a new sub-registry entitled
"Instructive Routing Function Name" under the "Internet Protocol
values for the subtype registries are given in Table 1.

9. Security Considerations

The instructive routing is only applicable to a satellite network
that is using the satellite semantic address. It will add
instructive routing header at a GS and the header will be removed
before reaching another GS. Normally, a satellite network including
all GS is trusted domain. Traffic will be filtered at the domain
boundaries. Non-authorized users cannot access the satellite
network.
10. Contributors

11. Acknowledgements

12. References

12.1. Normative References


12.2. Informative References


[TR38-821] "Solutions for NR to support Non-Terrestrial Networks (NTN)",
<https://portal.3gpp.org/desktopmodules/Specifications/

[IPv6_Parameters]
IANA, "Internet Protocol Version 6 (IPv6) Parameters",
<https://www.iana.org/assignments/ipv6-parameters/
ipv6-parameters.xhtml#ipv6-parameters-3>.

Appendix A. Change Log

* Initial version, 02/28/2022

Authors’ Addresses

Lin Han (editor)
Futurewei Technologies, Inc.
2330 Central Express Way
Santa Clara, CA 95050,
United States of America
Email: lhan@futurewei.com

Alvaro Retana
Futurewei Technologies, Inc.
2330 Central Express Way
Santa Clara, CA 95050,
United States of America
Email: alvaro.retana@futurewei.com

Richard Li
Futurewei Technologies, Inc.
2330 Central Express Way
Santa Clara, CA 95050,
United States of America
Email: rli@futurewei.com
Satellite Semantic Addressing for Satellite Constellation
draft-lhan-satellite-semantic-addressing-01

Abstract

This document presents a semantic addressing method for satellites in satellite constellation connecting with Internet. The satellite semantic address can indicate the relative position of satellites in a constellation. The address can be used with traditional IP address or MAC address or used independently for IP routing and switching.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 7 September 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights
and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction ......................................................... 2
2. Terminology .......................................................... 3
3. Overview .............................................................. 4
4. Basics of Satellite Constellation and Satellite Orbit ......... 5
   4.1. Satellite Orbit ............................................... 5
   4.2. Satellite Constellation Compositions ......................... 6
   4.3. Communication between Satellites by ISL .................... 7
5. Addressing of Satellite ............................................ 9
   5.1. Indexes of Satellite .......................................... 9
   5.2. The Range of Satellite Indexes ............................... 12
   5.3. Other Info for satellite addressing ........................ 13
   5.4. Encoding of Satellite Semantic Address ..................... 14
   5.5. Link Identification by Satellite Semantic Address ....... 16
6. IANA Considerations ................................................ 18
7. Contributors .......................................................... 18
8. Acknowledgements ................................................... 18
9. References ............................................................ 18
   9.1. Normative References ........................................ 18
   9.2. Informative References ....................................... 19
Appendix A. Change Log .................................................. 21
Authors’ Addresses ..................................................... 21

1. Introduction

Satellite constellation technologies for Internet are emerging and expected to provide Internet service like the traditional wired network on the ground. A typical satellite constellation will have couple of thousands or over ten thousand of LEO and/or VLEO. Satellites in a constellation will be connected to adjacent satellites by Inter-Satellite-Links (ISL), and/or connected to ground station by microwave or laser links. ISL is still in research stage and will be deployed soon. This memo is for the satellite networking with the use of ISL.
The memo proposes to use some indexes to represent a satellite’s orbit information. The indexes can form satellite semantic address, the address can then be embedded into IPv6 address or MAC address for IP routing and switching. The address can also be used independently if the shorter than 128-bit length of IP address is accepted. As an internal address for satellite network, it only applies to satellites that will form a constellation to transport Internet traffic between ground stations and will not be populated to Internet by BGP.

2. Terminology

LEO  Low Earth Orbit with the altitude from 180 km to 2000 km.

VLEO  Very Low Earth Orbit with the altitude below 450 km

GEO  Geosynchronous orbit with the altitude 35786 km

ISL  Inter Satellite Link

ISLL  Inter Satellite Laser Link

3D  Three Dimensional

GS  Ground Station, a device on ground connecting the satellite. In the document, GS will hypothetically provide L2 and/or L3 functionality in addition to process/send/receive radio wave. It might be different as the reality that the device to process/send/receive radio wave and the device to provide L2 and/or L3 functionality could be separated.

SGS  Source ground station. For a specified flow, a ground station that will send data to a satellite through its uplink.

DGS  Destination ground station. For a specified flow, a ground station that is connected to a local network or Internet, it will receive data from a satellite through its downlink and then forward to a local network or Internet.

L1  Layer 1, or Physical Layer in OSI model [OSI-Model]

L2  Layer 2, or Data Link Layer in OSI model [OSI-Model]
L3  Layer 3, or Network Layer in OSI model [OSI-Model], it is also called IP layer in TCP/IP model

BGP  Border Gateway Protocol [RFC4271]

IGP  Interior gateway protocol, examples of IGPs include Open Shortest Path First (OSPF [RFC2328]), Routing Information Protocol (RIP [RFC2453]), Intermediate System to Intermediate System (IS-IS [RFC7142]) and Enhanced Interior Gateway Routing Protocol (EIGRP [RFC7868]).

3. Overview

For IP based satellite networking, the topology is very dynamic and the traditional IGP and BGP based routing technologies will face challenges according to the analysis in [I-D.lhan-problems-requirements-satellite-net]. From the paper, we can easily categorize satellite links as two types, steady and unsteady. For un-steady links, the link status will be flipping every couple of minutes.

Section 5.5 has more details about how to identify different links.

Some researches have been done to handle such fast changed topologies. One method to overcome the difficulties for routing with un-steady links is to only use the steady links, and get rid of un-steady links unless it is necessary. For example, for real deployment, only links between satellite and ground stations are mandatory to use, other un-steady links can be avoided in routing and switching algorithms. [Routing-for-LEO] proposed to calculate the shortest path by avoiding un-steady links in polar area and links crossing Seam line since satellites will move in the opposite direction crossing the Seam line.

Traditionally, to establish an IP network for satellites, each satellite and its interface between satellites and to ground stations have to be assigned IP addresses (IPv4 or IPv6). The IP address can be either private or public. IP address itself does not mean anything except routing prefix and interface identifier [RFC8200].

To utilize the satellite relative position for routing, it is desired that there is an easy way to identify the relative positions of different satellites and identify un-steady links quickly. The traditional IP address cannot provide such functionality unless we have the real-time processing for 3D coordinates of satellites to figure out the relative positions of each satellite, and some math calculation and dynamic database are also needed in routing algorithm.
to check if a link is steady or not. This will introduce extra data exchanged for routing protocols and burden for the computation in every satellite. Considering the ISL link speed (up to 10G for 2000km) and hardware cost (Radiation-hardened semiconductor components are needed) in satellite are more constraint than for network device on ground, it is expected to simplify the routing algorithm, reduce the requirement of ISL, onboard CPU and memory.

The document proposes to form a semantic address by satellite orbit information, and then embedded it into a proper IP address. The IP address of IGP neighbors can directly tell the relative position of different satellites and if links between two satellites are steady or not.

The document does not describe the details how the semantic address is used to improve routing and switching or new routing protocols, those will be addressed in different documents.

4. Basics of Satellite Constellation and Satellite Orbit

This section will introduce some basics for satellite such as orbit parameters.

4.1. Satellite Orbit

The orbit of a satellite can be either circular or ecliptic, it can be described by following Keplerian elements [KeplerianElement]:

1. Inclination (i)
2. Longitude of the ascending node (Omega)
3. Eccentricity (e)
4. Semimajor axis (a)
5. Argument of periapsis (omega)
6. True anomaly (nu)

The circular orbit is widely used by proposals of satellite constellation from different companies and countries. For a circular orbit, we will have:

* Eccentricity $e = 0$
* Semimajor axis $a = $ Altitude of satellite
* Argument of periapsis omega = 90 degree

So, three parameters, Altitude, Inclination and Longitude of the ascending node, will be enough to describe the orbit. The satellite will move in a constant speed and True anomaly (nu) can be easily calculated after the epoch time is defined.

4.2. Satellite Constellation Compositions

One satellite constellation may be composed of many satellites (LEO and VLEO), but normally all satellites are grouped in a certain order that is never changed during the life of satellite constellation. Each satellite constellation’s orbits parameters described in Section 4.1 must be approved by regulator and cannot be changed either. Follows are characters of one satellite constellation:

1. One Satellite Constellation is composed of couple of shell groups of satellites.

2. Same shell group of satellite will have the same altitude and inclination.

3. The total N orbit planes in the same shell group of satellites will be evenly distributed by the same interval of Longitude of the ascending node (Omega). The interval equals to (360 degree/ N). As a result, all orbit planes in the same shell group will effectively form a shell to cover earth (there will be a coverage hole for the shell if the inclination angle is less than 90 degree).

4. Each orbit plane in the same shell group will have the same number of satellites, all satellites in the same orbit plane will be evenly distributed angularly in the orbit plane.

5. All satellites in the same shell group are moving in the same circular direction within the same orbit plane. As a result, at any location on earth, we can see there will have two group of satellites moving on the opposite direction. One group moves from south to north, and another group moves from north to south. Section 5.5 has more details.
4.3. Communication between Satellites by ISL

When ISL is used for the communication between satellites, each satellite will have a fixed number of links to connect to its neighbor. Due to the cost of ISL and the constraints of power supply on satellite, the number of ISL is normally limited to connect to its closest neighbors. In 3D space, each satellite may have six types of adjacent satellites, each type represents one direction. The number of adjacent neighbors in one direction is dependent on the number of deployment of ISL device on satellites, for example, the laser transmitter and receiver for ISLL. Figure 1 illustrates satellite S0 and its adjacent neighbors.

```
/           /           /
/           /           /
/           /           /  S7 S8 S9
/           /           /
/           /           /
/           /           /
/           S1           /
S5           /           S3
/           /           /
/           /           /
/           S0           /
/           /           /
S6           /           S4
/           S2           /
/           /           /
/           /           /
/           /           /
S10 S11 S12
/           /           /
/           /           /
/           /           /
/orbit     orbit     orbit  ^ Moving direction
```

Figure 1: Satellite S0 and its adjacent neighbors

All adjacent satellites of S0 in Figure 1 are listed below:

1. The front adjacent satellite S1 that is on the same orbit plane as S0.
2. The back adjacent satellite S2 that is on the same orbit plane as S0.
3. The right adjacent satellites S3 and S4 that are on the right orbit plane of S0.
4. The left adjacent satellites S5 and S6 that are on the left orbit plane of S0

5. The above adjacent satellites S7 to S9 that are on the above orbit plane of S0

6. The below adjacent satellite S10 to S12 that are on the below orbit of plane S0

The relative position of adjacent satellites will directly determine the quality of ISL and communication. From the analysis in [I-D.lhan-problems-requirements-satellite-net], The speed of satellite is only related to the altitude of the satellite (on circular orbit), all satellites with a same altitude will move with the same speed. So, in above adjacent satellites, some adjacent satellite’s relative positions are steady and the ISL can be alive without interruption caused by movement. Some adjacent satellites relative positions are changing quickly, the ISL may be down since the distance may become out of reach for the laser of ISL, or the quick changed positions of two satellite make the tracking of laser too hard. Below are details:

* The relative position of satellites in the same orbit plane will be the steadiest.

* The relative position of satellites in the direct neighbor orbit planes in the same shell group and moving in the same direction will be steady at equator area, but will be changing when two orbits meet on the polar area. Whether the link status will be flipping depends on the tracking technology and the range of laser pointing angle of ISL. See Figure 2.

* The relative position of satellites in the neighbor orbit planes in the same shell group but moving in the different direction will not be steady at all times. More details are explained in Figure 8

* The relative position of satellites in the neighbor orbit planes in the different shell group will be dependent on the difference of altitude and inclination. This has been analyzed in [I-D.lhan-problems-requirements-satellite-net].
Two satellites S1 and S2 are at position P1 and P2 at time T1. S1’s right facing ISL is connected to S2’s left facing ISL. S1 and S2 move to the position P4 and P3 at time T2. S1’s left facing ISL is connected to S2’s right facing ISL. So, if the range of laser pointing angle is 360 degree and tracking technology supports, the ISL will not be flipping after passing polar area; Otherwise, the link will be flipping.

Figure 2: Satellite’s Position and ISL Change at Polar Area

5. Addressing of Satellite

When ISL is deployed in satellite constellation, all satellites in the constellation can form a network like the wired network on ground. Due to the big number of satellites in a constellation, the network could be either L2 or L3. The document proposes to use L3 network for better scalability.

When satellites form a L3 network, it is expected that IP address is needed for each satellite and its ISLs.

While the traditional IP address can still be used for satellite network, the document proposes an alternative new method for satellite’s addressing system. The new addressing system can indicate a satellite’s orbit info such as shell group index, orbit plane index and satellite index. This will make the adjacent satellite identification for link status easier and benefit the routing algorithms.

5.1. Indexes of Satellite

As described in Section 4.2, one satellite has three important orbit related information as described below.

1. Index for the shell group of satellites in a satellite constellation

2. Index for the orbit plane in a shell group of satellites
3. Index for the satellite in an orbit plane

It should be noted that for all type of indexes, it is up to the owner to assign the index number. There is no rule for which one should be assigned with which number. The only important rule is that all index number should be in sequential to reflect its relative order and position with others. Below is an example of assignment rules:

1. The 1st satellite launched in an orbit plane can be assigned for the 1st satellite index (0), the incremental direction of the satellite index in the same orbit plane is the incremental direction of "Argument of periapsis (omega)"

2. The 1st orbit plane established can be assigned for the 1st orbit plane index (0), the incremental direction of the orbit plane index is the incremental direction of "Longitude of the ascending node (Omega)".

3. The shell group of satellites with the lowest altitude can be assigned for the 1st shell group index (0), the incremental direction of shell group index is the incremental direction of altitude.

Figure 3 and Figure 4 illustrate three types of indexes for satellite.
Figure 3: Shell Group and Orbit Plane Indexes for Satellites

Shell Group and Orbit Plane Indexes for Satellites

Figure 4

Three type of Index for satellites
5.2. The Range of Satellite Indexes

The ranges of different satellite indexes will determine the range the dedicated field for semantic address. The maximum indexes depend on the number of shell group, orbit plane and satellite per orbit plane. The number of orbit plane and satellite per orbit plane have relationship with the coverage of a satellite constellation. There are minimum numbers required to cover earth. [I-D.lhan-problems-requirements-satellite-net] has given the detailed math to estimate the minimal number required to cover the earth. There are two key parameters that determine the minimal number of satellite required. One is the elevation angle, another is the altitude. Spacelink has proposed two elevation angles, 25 and 35 degrees [SpaceX-Non-GEO]. The lowest LEO altitude can be 160km according to [Lowest-LEO-ESA]. The Table 1 and Table 2 illustrate the estimation for different altitude (As), the coverage radius (Rc), the minimal required number of orbit planes (No) and satellite per orbit plane (Ns). The elevation angle is 25 degree and 35 degrees respectively.

<table>
<thead>
<tr>
<th>Parameters</th>
<th>VLEO1</th>
<th>VLEO2</th>
<th>LEO1</th>
<th>LEO2</th>
<th>LEO3</th>
<th>LEO4</th>
<th>LEO5</th>
</tr>
</thead>
<tbody>
<tr>
<td>As(km)</td>
<td>160</td>
<td>300</td>
<td>600</td>
<td>900</td>
<td>1200</td>
<td>1500</td>
<td>2000</td>
</tr>
<tr>
<td>Rc(km)</td>
<td>318</td>
<td>562</td>
<td>1009</td>
<td>1382</td>
<td>1702</td>
<td>1981</td>
<td>2379</td>
</tr>
<tr>
<td>Ns</td>
<td>73</td>
<td>42</td>
<td>23</td>
<td>17</td>
<td>14</td>
<td>12</td>
<td>10</td>
</tr>
<tr>
<td>No</td>
<td>85</td>
<td>48</td>
<td>27</td>
<td>20</td>
<td>16</td>
<td>14</td>
<td>12</td>
</tr>
</tbody>
</table>

Table 1: Satellite coverage (Rc), minimal number of orbit plane (No) and satellite (Ns) per orbit plane for different LEO/VLEOs, Elevation angle = 25 degree
The real deployment may be different as above analysis. Normally, more satellites and orbit planes are used to provide better coverage. So far, there are only two proposals available, one is StarLink, another is from China Constellation. For proposals of [StarLink], there are 7 shell groups, the number of orbit plane and satellites per orbit plane in all shell groups are 72 and 58; For proposals of [China-constellation], there are 7 shell groups, the number of orbit plane and satellites per orbit plane in all shell groups are 60 and 60;

It should be noted that some technical parameters, such as the inclination and altitude of orbit planes, in above proposals may be changed during the long-time deployment period, but the total numbers for indexes normally do not change.

From the above analysis, to be conservative, it is safe to conclude that the range of all three satellite indexes are less than 256, or 8-bit number.

5.3. Other Info for satellite addressing

In addition to three satellite indexes described in Section 5.1, other information is also important and can also be embedded into satellite address:

1. The company or country code, or the owner code. In the future, there may have multiple satellite constellations on the sky from different organizations, and the inter-constellation communication may become as normal that is similar to the network on the ground. This code will be useful to distinguish different satellite constellation and make the inter-constellation communication possible. One satellite constellation will have

Han, et al. Expires 7 September 2022 [Page 13]
one code assigned by international regulator (IANA or ITU).
Considering the limit of LEO orbits and the cost of satellite
constellations, the total number of satellite constellation is
very limited. So, the size of code is limited.

2. The Interface Index. This index is to identify the ISL or ISLL
for a satellite. As described in Section 4.3, the total number
of ISL is limited. So, the size of interface index is also
limited.

5.4. Encoding of Satellite Semantic Address

The encoding for satellite semantic address is dependent on what
routing and switching (L2 or L3 solution) technologies are used for
satellite networking, and finally dependent on the decision of IETF
community.

Follows are some initial proposals:

1. When satellite network is using L3 solution, the satellite
semantic address is encoded as the interface identifier (i.e.,
the rightmost 64 bits) of the IPv6 address for IPv6. Figure 5
shows the format of IPv6 Satellite Address.

2. When satellite network is using L2 solution, the satellite
semantic address can be embedded into the field of "Network
Interface Controller (NIC) Specific" in MAC address
[IEEE-MAC-Address]. But due to shorter space for NIC, the "Index
for the shell group" and "Index for Interface" will only have
4-bit. This is illustrated in Figure 6. This encoded MAC
address can also be used for L3 solution where the interface MAC
may be also needed to be configured for each ISL.

3. Recently, some works suggested to use Length Variable IP address
for routing and switching [Length-Variable-IP] or use flexible IP
address [I-D.jia-flex-ip-address-structure] or shorter IP address
[I-D.li-native-short-addresses] to solve some specific problems
that regular IPv6 is not very suitable. Satellite network also
belongs to such specific network. Due to the resource and cost
constraints and requirement for radiation hardened electronic
components, the ISL speed, on-board processor and memory are
limited in performance, power consumption and capacity compared
with network devices on ground. So, using IPv6 directly in
satellite network is not an optimal solution because IPv6 header
size is too long for such small network. From above analysis,
32-bit to 64-bit length of IP address is enough for satellite
networking. Using 128-bit IPv6 will consume more resource
especially the ISL bandwidth, processing power and memory, etc.
If shorter than 128-bit IP address is accepted as IETF work, the satellite semantic address can be categorized as a similar use case. Figure 7 illustrates a 32-bit Semantic Satellite Address format. The final coding for the shorter IP address can be decided by the community. How to use the 32-bit Semantic Satellite address can be addressed later on in different document.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
˜                     Subnet Prefix (64 bits)                   ˜
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Owner Code  |  Shell_Index  |  Orbit_Index  |   Sat_Index   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   Intf_Index  |                    Reserved                   |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Owner Code: Identifier for the owner of the constellation
Shell_Index: Index for the shell group of satellite in a satellite constellation
Orbit_Index: Index for the orbit plane in a shell group of satellite
Sat_Index: Index for the satellite in an orbit plane
Intf_Index: Index for interface on a satellite
Reserved: 24-bits reserved

Figure 5: The IPv6 Satellite Address
Using above satellite semantic addressing scheme, to identify steady and un-steady links is as simple as below:
Assuming:

1. The total number of satellites per orbit plane is M
2. The total number of orbit planes per shell group is N.
3. Two satellites have:
   * Satellite Indexes as: Sat1_Index, Sat2_Index
   * Orbit plane Indexes as: Orbit1_Index, Orbit2_Index
   * Shell group Indexes as: Shell1_Index, Shell2_Index

Steady links:

1. The links between adjacent satellites on the same orbit plane, or, the satellite indexes satisfy:
   * Sat2_Index = Sat1_Index + 1, when Sat1_Index < M-1; Sat2_Index = 0, when Sat1_Index = M-1; and
   * Orbit1_Index = Orbit2_Index, Shell1_Index = Shell2_Index.

2. The links between satellites on adjacent orbit planes on the same altitude. and two satellites are moving to the same direction, or, the satellite indexes satisfy:
   * Orbit2_Index = Orbit1_Index + 1, when Orbit1_Index < N-1; Orbit2_Index = 0, when Orbit1_Index = N-1; and
   * Shell1_Index = Shell2_Index.
   * Sat1_Index and Sat2_Index may be equal or have difference, depend on how the link is established.

Un-Steady links:

1. The links between satellite and ground stations.
2. The links between satellites on adjacent orbit planes on the same altitude. Two satellites are moving to the different direction. Or, the satellite indexes do not satisfy conditions described in above #2 for Steady links.
3. The links between satellites on adjacent orbit planes on different altitude. Or, the satellite indexes satisfy:
Figure 8 illustrates the links for adjacent orbit planes (#2 for Steady Link and Un-steady Link above). From the figure, it can be noticed that some links may have shorter distance than steady link, but they are unsteady. For example, the links between S1 and S4; S4 and S2; S2 and S5, etc.

```
  i+N/2       i+1+N/2       i+2+N/2
  \  \         \  \         \  \ 
    \    \       \    \       \    
  S1 ............ S2 ............ S3 \
     / \\          / \\          / \\ \
    /   S4 ............ S5 ............ S6
   /     /          /     /          / \
  /     /          /     /          / \\
 i-1    i           i    i+1
```

* The total number of orbit planes are N  
* The number (i-1, i, i+1,...) represents the Orbit index  
* The bottom numbers (i-1, i, i+1) are for orbit planes on which satellites (S1, S2, S3) are moving from bottom to up.  
* The top numbers (i+N/2, i+1+N/2, i+2+N/2) are for orbit planes on which satellites (S4, S5, S6) are moving from up to bottom.  
* Dot lines are the steady links

Figure 8: The links between satellites on adjacent orbit planes

6. IANA Considerations

This memo may include request to IANA for owner code, see Section 5.4.

7. Contributors

8. Acknowledgements

9. References

9.1. Normative References

9.2. Informative References

[I-D.lhan-problems-requirements-satellite-net]

[I-D.jia-flex-ip-address-structure]

[I-D.li-native-short-addresses]
[Routing-for-LEO]

[Length-Variable-IP]

[IEEE-MAC-Address]

[Lowest-LEO-ESA]
"Lowest LEO by ESA", <https://www.esa.int/ESA_Multimedia/Images/2020/03/Low_Earth_orbit#:~:text=A%20low%20Earth%20orbit%20(LEO,very%20far %20above%20Earth’s%20surface)>.

[KeplerianElement]

[OSI-Model]


[China-constellation]

[SpaceX-Non-GEO]
Appendix A. Change Log

* Initial version, 10/19/2021

* 01 version, 02/28/2022

Authors’ Addresses

Lin Han (editor)
Futurewei Technologies, Inc.
2330 Central Express Way
Santa Clara, CA 95050,
United States of America
Email: lhan@futurewei.com

Richard Li
Futurewei Technologies, Inc.
2330 Central Express Way
Santa Clara, CA 95050,
United States of America
Email: rli@futurewei.com

Alvaro Retana
Futurewei Technologies, Inc.
2330 Central Express Way
Santa Clara, CA 95050,
United States of America
Email: alvaro.retana@futurewei.com

Meiling Chen
China Mobile
32, Xuanwumen West
Beijing 100053
China
Email: chenmeiling@chinamobile.com

Ning Wang
University of Surrey
Guildford
Surrey, GU2 7XH
United Kingdom
Email: n.wang@surrey.ac.uk
Abstract

This document defines the Segment Routing IPv6 Mobile User Plane (SRv6 MUP) architecture for Distributed Mobility Management. The requirements for Distributed Mobility Management described in [RFC7333] can be satisfied by routing fashion.

Mobile services are deployed over several parts of IP networks. A Segment Routing over IPv6 (SRv6) network can accommodate all, or part of those networks thanks to the large address space of IPv6 and the network programming capability described in [RFC8986].

Segment Routing IPv6 Mobile User Plane Architecture can incorporate existing session based mobile networks. By leveraging SRv6 network programmability, mobile user plane can be integrated into the SRv6 data plane. In that routing paradigm, session information between the entities of the mobile user plane is turned to routing information.
Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 20 September 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

Table of Contents

1. Introduction ........................................... 3
   1.1. Requirements Language ............................ 5
2. Terminology ........................................... 5
3. Architecture Overview .................................. 5
4. Mobile User Plane Segment ............................. 6
5. Distribution of Mobile User Plane Segment Information ... 7
   5.1. Direct Segment Discovery Route ................... 7
   5.2. Interwork Segment Discovery Route ............... 8
6. Distribution of Session Transformed Route ............... 8
   6.1. Type 1 Session Transformed Route ............... 9
   6.2. Type 2 Session Transformed Route ............... 9
   6.3. MUP Controller .................................. 9
7. Illustration .......................................... 10
8. IANA Considerations .................................. 15
9. Security Considerations .............................. 16

1. Introduction

Mobile services require IP connectivity for communication between the entities of mobile service architecture [RFC5213][TS.23501]. To provide the IP connectivity, Segment Routing (SR) [RFC8402] can be a candidate solution.

In PMIPv6 [RFC5213], IP connectivity between LMA and MAG can be provided over SR networks, as well as LMA and Internet. In 3GPP 5G [TS.23501], IP connectivity for N3 interface between gNodeB(es) and UPFs can also be provided by SR, as well as for N6 interface between UPFs and DNs (Data Network).

These IP connectivities may be covered by multiple SR networks, or just one SR network, depending on the size of the deployment. In the latter case, it is expected that the address space of the SR network should be large enough to cover a vast number of nodes, such as millions of base stations. For this reason, use of IPv6 for the SR dataplane looks sufficiently suitable.

SRv6 is an instantiation of SR over IPv6 dataplane in which a single network can accommodate all entities of mobile services thanks to the huge available address space and network programming capability described in [RFC8986].

Meanwhile, SRv6 network programmability enhances SRv6 dataplane to be integrated with mobile user plane [I-D.ietf-dmm-srv6-mobile-uplane]. It will make an entire SRv6 network support the user plane in a very efficient distributed routing fashion.

On the other hand, the requirements for Distributed Mobility Management (DMM) described in [RFC7333] can be satisfied by session management based solutions. [RFC8885] defines protocol extension to PMIPv6 for the DMM requirements. 3GPP 5G defines an architecture in which multiple session anchors can be added to one mobility session by the session management.

As a reminder, the user plane related requirements in [RFC7333] are reproduced here:

REQ1: Distributed mobility management
IP mobility, network access solutions, and forwarding solutions provided by DMM MUST enable traffic to avoid
traversing a single mobility anchor far from the optimal route. It is noted that the requirement on distribution applies to the data plane only.

REQ3: IPv6 deployment
DMM solutions SHOULD target IPv6 as the primary deployment environment and SHOULD NOT be tailored specifically to support IPv4, particularly in situations where private IPv4 addresses and/or NATs are used.

REQ4: Existing mobility protocols
A DMM solution MUST first consider reusing and extending IETF standard protocols before specifying new protocols.

REQ5: Coexistence with deployed networks/hosts and operability across different networks
A DMM solution may require loose, tight, or no integration into existing mobility protocols and host IP stacks. Regardless of the integration level, DMM implementations MUST be able to coexist with existing network deployments, end hosts, and routers that may or may not implement existing mobility protocols. Furthermore, a DMM solution SHOULD work across different networks, possibly operated as separate administrative domains, when the needed mobility management signaling, forwarding, and network access are allowed by the trust relationship between them.

This document defines the Segment Routing IPv6 Mobile User Plane (SRv6 MUP) architecture for Distributed Mobility Management. SRv6 MUP is not a mobility management system itself, but an architecture to integrate mobile user plane into the SRv6 data plane.

In this routing paradigm, session information from a mobility management system will be transformed to routing information. It means that mobile user plane specific nodes for the anchor or intermediate points are no longer required. The user plane anchor and intermediate functions can be supported by SR throughout an SR domain (REQ1), not to mention that SRv6 MUP will naturally be deployed over IPv6 networks (REQ3).

SRv6 MUP architecture is independent from the mobility management system. For the requirements (REQ4, 5), SRv6 MUP architecture is designed to be pluggable user plane part of existing mobile service architectures. Those existing architectures are for example defined in [RFC5213], [TS.23501], or if any.
The level of SRv6 MUP integration for mobile networks running based on the existing architecture will be varied and depending on the level of SRv6 awareness of the control and user plane entities.

Specifying how to modify the existing architecture to integrate SRv6 MUP is out of scope of this document. What this document provides for the existing architecture is an interface for SRv6 MUP which the existing or future architectures can easily integrate.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

2. Terminology

MUP: Mobile User Plane
MUP Segment: Representation of mobile user plane segment
PE: Provider Edge node in an SR network
MUP Controller: Controller node for an SR network
UE: User Equipment, as per [TS.23501]
MN: Mobile Node, as per [RFC5213]

3. Architecture Overview

SRv6 MUP architecture defined in this document introduces new segment types of MUP segment called "Direct segment", and "Interwork Segment". An SR node of PE accommodates an Interwork Segment and/or a Direct Segment. Figure 1 depicts the overview.
This document also defines new routing information called "Segment Discovery route" and "Session Transformed route". To carry these new routing information, this architecture requires extending the existing routing protocols. Any routing protocol can be used to carry this information but this document recommends using BGP. Thus, this document describes extensions on BGP as an example.

4. Mobile User Plane Segment

This document defines two new types of Mobile User Plane (MUP) segment. A MUP segment represents a network segment consisting of a mobile service. The MUP segment can be created by a PE which provides connectivity for the mobile user plane.

Direct Segment is a type of MUP segment that provides connectivity between MUP segments through the SRv6 network. Interwork Segment is another type of MUP segment. It provides connectivity between a user plane protocol of existing or future mobile service architecture and other MUP segments through the SRv6 networks.

An SRv6 SID (Segment Identifier) can represents a MUP segment. The SID can be any behavior defined in [RFC8986], [I-D.ietf-dmm-srv6-mobile-uplane], or any other extensions for further use cases. The behavior of the MUP segment will be chosen by the role of the representing MUP segment.
For example, in case of a PE interfaces to 5G user plane on the access side defined as "N3" in [TS.23501], the PE accommodates the N3 network as Interwork Segment in a routing instance and then the behavior of created segment SID by the PE will be "End.M.GTP4.E", or "End.M.GTP6.E". In this case, the PE may associate the SID to the routing instance for the N3 access network (N3RAN).

Another example here is that a PE interfaces to 5G DN on the core side defined as "N6" in [TS.23501], the PE accommodates the N6 network in a routing instance as Direct Segment and then the behavior of the created segment SID by the PE will be "End.DT4", "End.DT6", or "End.DT2". In this case, the PE may associate the SID to the routing instance for the N6 data network (N6DN).

5. Distribution of Mobile User Plane Segment Information

Distribution of MUP segment information can be done by advertising routing information with the MUP segment for mobile service. A PE distributes MUP segment information when a MUP segment is connected to the PE.

A MUP Segment Discovery route is routing information that associates the MUP segment with network reachability. This document defines the basic discovery route types, Direct Segment Discovery route, and Interwork Segment Discovery route. Other types of segment discovery route may be mobile service architecture specific. Defining the architecture specific network reachability is out of scope of this document and it will be specified in another document.

5.1. Direct Segment Discovery Route

When a PE accommodates a network through an interface or a routing instance as a Direct Segment, the PE advertises the corresponding Direct Segment Discovery route for the interface or the routing instance. The Direct Segment Discovery route includes an address of the PE in the network reachability information with an extended community indicating the corresponding Direct Segment, and SID of the routing instance to the SR domain.

For example in 3GPP 5G specific case, an PE may connect to N6 interface on a DN side, an MUP Segment Discovery route for the DN will be advertised with an address of the PE, corresponding SID and Direct Segment extended community to the routing instance for the DN from the PE.

When a PE receives a Direct Segment Discovery route from other PEs, the PE keeps the received Direct Segment Discovery route in the RIB. The PE uses the received Direct Segment Discovery route to resolve
Type 2 session transformed routes reachability, described in Section 6.2. If the Direct Segment Discovery route resolves reachability for the endpoints, and match the Direct Segment extended community of the Type 2 session transformed routes, the PE updates the FIB entry for the Type 2 session transformed route with the SID of the matched Direct Segment Discovery route.

5.2. Interwork Segment Discovery Route

When a PE accommodates a network through an interface or a routing instance for the user plane protocol of the mobile service architecture as an Interwork Segment, the PE advertises the corresponding Interwork Segment Discovery route with the prefixes of the Interwork Segment and the corresponding SID of the prefixes to the SR domain.

For example in 3GPP 5G specific case, an Interwork Segment Discovery route for N3 network accommodating RAN will be incorporated in an N3RAN segment discovery route associated with a RAN segment SID.

When a PE receives an Interwork Segment Discovery route, the PE keeps the received Interwork Segment Discovery routes in the RIB. The PE uses the received Interwork Segment Discovery routes to resolve the reachability for remote endpoint of Type 1 session transformed routes, described in Section 6.1. If the Interwork Segment Discovery route resolves the reachability for Type 1 session transformed routes, the PE updates the FIB entry for the prefix of Type 1 session transformed route with the SID of the matched MUP segment discovery route.

The received Interwork Segment Discovery routes MUST be used only to resolve reachability for the remote endpoints of Type 1 session transformed routes. The connectivity among the routing instances for Interwork Segments may be advertised as VPN routes. This is to avoid forwarding entries to the prefixes of Interwork Segment mingled in the other type of routing instance. A PE may discard the received Interwork segment discovery route if the Route Target extended communities of the route does not meet the PE's import policy.

6. Distribution of Session Transformed Route

SRv6 MUP architecture defines two types of session transformed route.
6.1. Type 1 Session Transformed Route

First type route, called Type 1 Session Transformed route, encodes IP prefix(es) for a UE or MN in a BGP MP-NLRI attribute with associated session information of the tunnel endpoint identifier on the access side. The MUP controller advertises the Type 1 Session Transformed route with the Route Target extended communities for the UE or MN to the SR domain.

A PE may receive the Type 1 Session Transformed routes from the MUP Controller in the SR domain. The PE may keep the received Type 1 Session Transformed routes advertised from the MUP Controller. The receiving PE will perform the importing of the received Type 1 Session Transformed routes in the configured routing instances based on the Route Target extended communities. A PE may discard the received Type 1 Session Transformed route if the PE fails to import the route based on the Route Target extended communities.

6.2. Type 2 Session Transformed Route

Second type route, called Type 2 Session Transformed route, encodes the tunnel endpoint identifier of the session on the core side in a BGP MP-NLRI attribute with the nature of tunnel decapsulation. Longest match algorithm for the prefix in this type of session transformed route should be applicable to aggregate the routes for scale. The MUP controller advertises the Type 2 Session Transformed route with the Route Target and Direct Segment extended communities for the endpoint to the SR domain.

A PE may receive the Type 2 Session Transformed routes from the MUP Controller in the SR domain. The PE may keep the received Type 2 Session Transformed routes advertised from the MUP Controller. The receiving PE will perform the importing of the received Type 2 Session Transformed routes in the configured routing instances based on the Route Target extended communities. A PE may discard the received Type 2 Session Transformed route if the PE fails to import the route based on the Route Target extended communities.

6.3. MUP Controller

A MUP controller provides a northbound API. A consumer of the API inputs session information for a UE or a MN from mobility management system. The MUP controller transforms the received session information to routing information and will advertise the session transformed routes with the corresponding extended communities to the SR domain.
The received session information is expected to include the UE or MN IP prefix(es), tunnel endpoint identifiers for both ends, and any other attributes for the mobile networks. For example in a 3GPP 5G specific case, the tunnel endpoint identifier will be a pair of the F-TEIDs on both the N3 access side (RAN) and core side (UPF).

7. Illustration

This section shows an illustration of SRv6 MUP deployment. The example deployment cases here is 3GPP 5G.

Before enabling SRv6 MUP, how SRv6 networks can accommodate existing mobile network service shown in Figure 2. The PEs of S1, S2, and S3 join an SR network. A routing instance is configured to each network of the mobile service. N6DN in S1 and S2 are supposed to provide connectivity to edge servers and the Internet respectively.

VRF (Virtual Routing Forwarding) is the routing instance to accommodate MUP segments in this section. All example cases in this section follow the typical routing policy control using the BGP extended community described in [RFC4360] and [RFC4684].

The following routing instances are configured:

* N3RAN in S1
  - export route V/v with route-target (RT) community C1
  - import routes which have route-target (RT) community C1 and C2

* N6DN in S1
- export route E/e with RT C4
- import routes which have RT C3 and C4

* N6DN in S2
- export route W/w with RT C4
- import routes which have RT C3 and C4

* N3UPF in S3
- export route X/x with RT C2
- import routes which have RT C1

* N6UPF in S3
- export route Y/y with RT C3
- import routes which have RT C4

Note: The above configurations are just to provide typical IP connectivity for 3GPP 5G. When the above configurations have been done, each endpoint in V/v and X/x can communicate through S1 and S3, but they can not communicate with nodes in E/e, W/w and Y/y.

Here, the PEs are configured to enable SRv6 MUP as following:

* S1
  - advertises Interwork type discovery route: V/v with SID S1::
  - set S1:: behavior End.M.GTP4.E or End.M.GTP6.E

* S1
  - advertise Direct type discovery route: MUP Direct Segment community D1 and SID S1:1::
  - set S1:1:: behavior End.DT4 or End.DT6 for the N6DN in S1

* S2
  - advertise Direct type route: MUP Direct Segment community D1 and SID S2::
- set S2:: behavior End.DT4 or End.DT6 for the N6DN in S2

S1 here adopts the local N6DN to prioritize closer segment for the same Direct Segment. Another PE may adopt D1 from S2, if the PE has no local N6DN for D1 and closer to S2 than S1.

![Diagram](image-url)

Figure 3

Now, session information U1 is put to a MUP Controller, MUP-C, and MUP-C is configured to transforms U1 to the routes as follows:

* MUP-C
  - attach the RT C3 to the DN in U1
  - transforms UE’s prefix U/u, the F-TEID on access side (gNB) and QFI in U1 to the Type 1 session transformed route for the prefix U/u with the F-TEID, the QFI, and RT C3
  - transforms F-TEID on core side (UPF) X in U1 to the Type 2 session transformed route for X with MUP segment-ID D1 and RT C2

Then N3RAN and N6DN import route X and U/u respectively. S1 and S2 resolves U/u’s remote endpoint with V/v and then install SID S1:: for U/u in FIB. S1:: will not be appeared in the packet from E/e to U/u over the wire.
As S1 adopts local N6DN for D1, N3RAN in S1 decapsulates GTP-U packets from V/v to X and then lookup the inner packets from U/u in N6DN after the decapsulation.

Note: When the above configurations have been done, SRv6 MUP is applied only to the packets from/to U/u. Each endpoint in U/u, W/w and E/e can communicate through S1 and S2. The rest of traffic from/to other UEs go through the usual 3GPP 5G user plane path using UPF via S3.

Another case shown in Figure 4 is that S4 joins the SR network and accommodates edge servers in the N6DN in S4.

The following routing instances are configured:

* N3RAN in S1 (same with the previous case)
  - export route V/v with RT C1
  - import routes which have RT C1 and C2

* N6DN in S1
  - export no route
  - import routes which have RT C4

* N6DN in S2 (same with the previous case)
Here, the PEs are configured to enable SRv6 MUP as following:

* S1 (same with the previous case)
  - advertises Interwork type route: V/v with SID S1::
  - set S1:: behavior End.M.GTP4.E or End.M.GTP6.E

* S2 (same with the previous case)
  - advertise Direct type route: MUP Direct Segment community D1 and SID S2::
  - set S2:: behavior End.DT4 or End.DT6 for the N6DN in S2

* S4
  - advertise Direct type route: MUP Direct Segment community D2 and SID S4::
- set S4:: behavior End.DT4 or End.DT6 for the N6DN in S4

As same as the previous case, S1 adopts the local N6DN for D1 as long as S1 prioritizes closer segment for the same MUP Direct Segment. The Direct type route from S4 for D2 with SID S4:: will be kept in S1.

* MUP-C (same with the previous case)

- attach the RT C3 to the DN in U1
- transforms UE’s prefix U/u, the F-TEID on access side (gNB) and QFI in U1 to the Type 1 session transformed route for the prefix U/u with the F-TEID, the QFI, and RT C3
- transforms F-TEID on core side (UPF) X in U1 to the Type 2 session transformed route for X with MUP Direct Segment community D1 and RT C2

Then N3RAN and N6DN import route X and U/u respectively. S2 and S4 resolve U/u’s remote endpoint with V/v and then install SID S1:: for U/u in FIB.

As same as the previous case, S1 adopts local N6DN for D1, N3RAN in S1 decapsulates GTP-U packets from V/v to X and then lookup the inner packets from U/u in N6DN after the decapsulation.

For D2 on the other hand, no corresponding N6DN existed in S1. However E/e with RT C4 from S4 is imported into N6DN in S1 as a vpn route, E/e is reachable from U/u via N6DN for D1 in S1.

If a session U1’ includes DN corresponding to D2, MUP-C advertises Type 2 session transformed route X’ with MUP Direct Segment community D2, and then N3RAN in S1 instantiates H.M.GTP4.D or End.M.GTP6.D for X with S4:: as the last SID in the received Direct type route from S4.

Note: When the above configurations have been done, SRv6 MUP is applied only to the packets from/to U/u. Each endpoint in U/u, W/w and E/e can communicate through S1, S2 and S4. The rest of traffic from/to other UEs go through the usual 3GPP 5G user plane path using UPF via S3.

8. IANA Considerations

This memo includes no request to IANA.
9. Security Considerations

TBD.

10. References

10.1. Normative References

[I-D.ietf-dmm-srv6-mobile-uplane]
Matsushima, S., Filsfils, C., Kohno, M., Garvia, P. C.,
Voyer, D., and C. E. Perkins, "Segment Routing IPv6 for
Mobile User Plane", Work in Progress, Internet-Draft,
draft-ietf-dmm-srv6-mobile-uplane-18, 18 February 2022,
<https://datatracker.ietf.org/doc/html/draft-ietf-dmm-
srv6-mobile-uplane-18>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
Requirement Levels", BCP 14, RFC 2119,
DOI 10.17487/RFC2119, March 1997,

Korhonen, "Requirements for Distributed Mobility
Management", RFC 7333, DOI 10.17487/RFC7333, August 2014,

Decraene, B., Litkowski, S., and R. Shakir, "Segment
Routing Architecture", RFC 8402, DOI 10.17487/RFC8402,

D., Matsushima, S., and Z. Li, "Segment Routing over IPv6
(SRv6) Network Programming", RFC 8986,
DOI 10.17487/RFC8986, February 2021,

10.2. Informative References

Communities Attribute", RFC 4360, DOI 10.17487/RFC4360,


Authors’ Addresses

Satoru Matsushima
SoftBank
Japan
Email: satoru.matsushima@g.softbank.co.jp

Katsuhiro Horiba
SoftBank
Japan
Email: katsuhiro.horiba@g.softbank.co.jp

Ashiq Khan
SoftBank
Japan
Email: ashiq.khan@g.softbank.co.jp

Yuya Kawakami
SoftBank
Japan
Email: yuya.kawakami01@g.softbank.co.jp

Tetsuya Murakami  
Arrcus, Inc.  
United States of America  
Email: tetsuya@arrcus.com

Keyur Patel  
Arrcus, Inc.  
United States of America  
Email: keyur@arrcus.com

Miya Kohno  
Cisco Systems, Inc.  
Japan  
Email: mkohno@cisco.com

Teppei Kamata  
Cisco Systems, Inc.  
Japan  
Email: tkamata@cisco.com

Pablo Camarillo Garvia  
Cisco Systems, Inc.  
Spain  
Email: pcamaril@cisco.com

Daniel Voyer  
Bell Canada  
Canada  
Email: daniel.voyer@bell.ca

Shay Zadok  
Broadcom  
Israel  
Email: shay.zadok@broadcom.com

Israel Meilik  
Broadcom  
Israel  
Email: israel.meilik@broadcom.com
Ashutosh Agrawal
Intel
United States of America
Email: ashutosh.agrawal@intel.com

Kumaresh Perumal
Intel
United States of America
Email: kumaresh.perumal@intel.com

Jakub Horn
Cisco Systems, Inc.
Czech Republic
Email: jakuhorn@cisco.com
User Plane Message Encoding
draft-murakami-dmm-user-plane-message-encoding-05

Abstract

This document defines the encoding of User Plane messages into Segment Routing Header (SRH). The SRH carries the User Plane messages over SRv6 Network.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on September 6, 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect
to this document. Code Components extracted from this document must include Simplified BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Simplified BSD License.

Table of Contents

1. Introduction .............................................. 2
2. Requirements Language ...................................... 3
3. Conventions and Terminology ................................ 3
4. Motivation .................................................. 3
5. User Plane Message encoding into SRH ...................... 3
   5.1. GTP-U Header format .................................... 4
   5.2. Args.Mob.Upmsg ......................................... 4
   5.3. Encoding of Tags Field ................................... 5
   5.4. User Plane message Information Element Support ........ 6
   5.5. SID flavor consideration ................................ 7
6. Security Considerations ..................................... 8
7. IANA Consideration .......................................... 8
8. Acknowledgements ........................................... 8
9. References .................................................. 8
   9.1. Normative References .................................... 8
   9.2. Informative References .................................. 9
Authors’ Addresses ................................................ 9

1. Introduction

3GPP defines User Plane function (UPF) and the protocol messages that it supports. The User Plane messages support in-band signalling for path and tunnel management. Currently, User Plane messages are defined in TS 29.281 [TS29281].

When applying SRv6 (Segment Routing IPv6) to the user plane of mobile networks, based on draft-ietf-dmm-srv6-mobile-uplane [I-D.ietf-dmm-srv6-mobile-uplane]. User Plane messages must be carried over SRv6 network. This document defines which User Plane message must be encoded to SRv6 and also defines how to encode the User Plane messages into SRH.

In addition, SRH is mandatory at the ultimate segment upon carrying the User Plane messages because User Plane message is encoded into SRH. Hence, this document considers how to deal with the encoding of User Plane messages into SRH when PSP is applied that SRH is popped out at the penultimate segment.
2. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in RFC 2119 [RFC2119].

3. Conventions and Terminology

UPF: User Plane Function.
SRH: IPv6 Segment Routing Header.
PSP: Penultimate Segment POP of the SRH.
USP: Ultimate Segment Pop of the SRH.

4. Motivation

3GPP User Plane needs to support the user plane messages associated with a GTP-U tunnel defined in [TS29281]. In the case of SRv6 User Plane [I-D.ietf-dmm-srv6-mobile-uplane], those messages are also required when the user plane interworks with GTP-U.

IPv6 Segment Routing Header (SRH) [RFC8754] is used for SRv6 User Plane. SRH is able to associate additional information to the segments. The Tag field of SRH is capable to indicate different properties within a SID. SRH TLV is capable to provide meta-data to the endpoint node.

The above capability of SRH motivates us to map the user plane messages into it because of the same encapsulation with the packets of carrying client packets. It introduces no additional headers or extension headers to be chained in the packet just for carrying the user plane messages.

5. User Plane Message encoding into SRH

This section defines how to encode the User Plane messages into SRH in order to carry the User Plane messages over SRv6 network.
5.1. GTP-U Header format

3GPP defines GTP-U Header format as shown below.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Ver |P|R|E|S|N| Message Type |           Length              |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                Tunnel Endpoint Identifier                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|      Sequence Number          | N-PDU Number  |  Next-Ext-Hdr |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 1: GTP-U Header format

User Plane message type is encoded in Message Type field of GTP-U Header. The following User Plane messages must be carried over SRv6 network at least. The value of each User Plane message type is defined as shown below.

- **Echo Request**: 1
- **Echo Reply**: 2
- **Error Indication**: 26
- **End Marker**: 254

5.2. Args.Mob.Upmsg

draft-ietf-dmm-srv6-mobile-uplane [I-D.ietf-dmm-srv6-mobile-uplane] defines the format of Args.Mob.Session argument which is used in SRv6 SID Mobility Functions in order to carry the PDU Session identifier. The format of Args.Mobs.Session is defined as shown below.

```
0                   1                   2                   3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   QFI     |R|U|                PDU Session ID                 |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|PDU Sess(cont’)|
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

Figure 2: Args.Mob.Session format

In case of Echo Request, Echo Reply and Error Indication, Sequence Number in GTP-U header needs to be carried. Similar to draft-ietf-
dmm-srv6-mobile-uplane [I-D.ietf-dmm-srv6-mobile-uplane], the new arguments to carry sequence number for Echo Request, Echo Reply and Error Indication message needs to be defined. For this, the following Args.Mobs.Upmsg should be defined newly to carry sequence number.

```
+----------------------------------+
|   QFI    |R|U|      Sequence Number      |      Pad      |
+----------------------------------+
|   Pad(cont') |      +---------------------------+
```

Figure 3: Args.Mob.Upmsg format for Echo Request, Echo Reply and Error Indication

QFI bit, R bit, U bit and 16-bit Sequence Number is encoded in Args.Mobs.Upmsg. The remaining bits followed by Sequence Number must be padded in 0.

In case of End Marker, TEID shall be used as PDU Session ID same as draft-ietf-dmm-srv6-mobile-uplane [I-D.ietf-dmm-srv6-mobile-uplane]. Hence, for End Marker, Args.Mobs.Session should be used to carry TEID as PDU Session ID.

5.3. Encoding of Tags Field

The Segment Routing Header is defined in IPv6 Segment Routing Header (SRH) [RFC8754]. This draft defines 16 bits Tag field but does not define the format or use of this Tag field in the Segment Routing Header.

The User Plane message type encoding is defined in TS 29.281 [TS29281]. Based on this definition, the User Plane message type must be encoded into the Tag field in the Segment Routing Header in order to indicate the type of the user plane messages for at least Echo Request, Echo Reply, Error Indication or End Marker.

Only UPF must process the Tag field where the user plane message is encoded. In addition, when the user plane message is encoded in the Tag field, the UF should not encode any segments in the Segment Routing Header whose function modifies the Tag field value. Any other transport router implementing SRv6 must ignore the Tag field upon processing the Segment Routing Header.

The user plane messages must be encoded into the Tag field as shown below.
End Marker, Echo Request and Echo reply messages do not require any additional information elements. However, Error Indication message requires the additional information elements like Tunnel Endpoint Identifier Data IE, GSN Address, etc. These additional information elements can be encoded into the SRH TLV that is defined in the next section.

5.4. User Plane message Information Element Support

End Marker, Echo Request and Echo Reply messages do not require any additional information elements. However, Error Indication message requires additional 3GPP IEs (Information Element). These additional information elements must be carried over SRv6 network as well. However SRv6 SID has limited space only. Hence it cannot carry a lot of information elements.

In order to carry more information elements, SRH TLV shall be leveraged. SRH TLV is defined in IPv6 Segment Routing Header (SRH) [RFC8754] in order to carry the meta-data for the segment processing. In order to carry additional User Plane messages like 3GPP IEs, the new type named as "User Plane Container" must be defined as the new SRH TLV. The "User Plane Container" can carry additional User Plane messages which includes multiple 3GPP IEs with 1 sub-TLV.
User Plane Container TLV

Type: to be assigned by IANA
Length: Length of User Plane message sub-TLV
User Plane message sub-TLV: User Plane message sub-TLV defined below

User Plane message sub-TLV

Type: Type of User Plane message sub-TLV

3GPP IE sub-TLV: 0x01
Length: Length of Value
Value: User Plane Message data

3GPP IE sub-TLV: multiple 3GG IEs

5.5. SID flavor consideration

This section considers SID flavor of where the SRH is popped out at either the penultimate or the ultimate segment.

In order to carry User Plane message over SRv6 network, SRH must be sustained over entire SRv6 network because User Plane message type and required information elements are encoded into SRH. If the
penultimate segment is popping out SRH, i.e., PSP, User Plane message can not be carried in entire SRv6 network.

In order to avoid this problem, USP is recommended in SRv6 Mobile network. In this case, SRH is never popped out and User Plane message can be sustained over entire SRv6 network.

However, if PSP needs to be enabled in SRv6 network, it is also a possible solution to encap another SRH which carries User Plane message along with the outer IPv6 or SRH.

6. Security Considerations

This document does not raise any additional security issues. This document just define the mechanisms for mapping between user plane message (GTP-U message) and SRH in SRv6. Basically, since this document is using SRH defined in [RFC8754] to carry user plane message, same security consideration stated in [RFC8754] shall be applied.

7. IANA Consideration

The type value of SRH TLV for User Plane Container must be assigned by IANA.

8. Acknowledgements

9. References

9.1. Normative References


9.2. Informative References

[I-D.ietf-dmm-srv6-mobile-uplane]
Matsushima, S., Filsfils, C., Kohno, M., Garvia, P. C.,
Voyer, D., and C. E. Perkins, "Segment Routing IPv6 for
Mobile User Plane", draft-ietf-dmm-srv6-mobile-uplane-18
(work in progress), February 2022.

[RFC1918]  Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G.,
and E. Lear, "Address Allocation for Private Internets",
BCP 5, RFC 1918, DOI 10.17487/RFC1918, February 1996,

(IPv6) Addressing Architecture", RFC 3513,
DOI 10.17487/RFC3513, April 2003,

Matsushima, S., and D. Voyer, "IPv6 Segment Routing Header
(SRH)", RFC 8754, DOI 10.17487/RFC8754, March 2020,

[TS29281]  3GPP, "General Packet Radio System (GPRS) Tunnelling
Protocol User Plane (GTPv1-U)", 2019,

Authors’ Addresses

Tetsuya Murakami (editor)
Arrcus, Inc
2077 Gateway Place, Suite 400
San Jose
USA

Email: tetsuya@arrcus.com

Satoru Matsushima
SoftBank
1-9-1 Higashi-Shinbashi, Munato-ku
Tokyo
Japan

Email: satoru.matsushima@g.softbank.co.jp
Abstract

This document specifies mechanisms that allow a node to monitor an OSPF network actively without influencing the topology or affecting its stability.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 8 September 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document. Code Components extracted from this document must include Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.
1. Introduction

Monitoring the control plane activity in a network is essential to designing and maintaining a robust and stable network. Passive (listen-only) devices deployed in broadcast or non-broadcast multi-access (NBMA) networks have typically satisfied the need. However, passive devices depend on more than two routers being present in the network and are not visible to the network operator — anyone can listen.

An alternative implementation, primarily used in point-to-point interfaces, or in cases where the listening device is the only other node on the interface, is to participate fully in the protocol: create a full adjacency with the closest router, participate in designated router (DR) election, etc. The node is now visible in the network, can advertise control plane information, and any changes in its status are flooded throughout the network. Many link state advertisements (LSA) or state changes can cause instability in the network, and additional configuration is usually needed to avoid the device becoming a transit node.

This document specifies mechanisms that allow a node to monitor OSPF activity without influencing the topology or affecting its stability while being fully adjacent and known to the network operator. These nodes are referred to as a Monitor Node. Two such mechanisms are introduced:

Section 3 describes a local implementation to be used in the case where the Monitor Node is the only other router on an interface.

Section 4 specifies signaling in the Hello message for a node to communicate its intention to become a Monitor Node.
The mechanisms presented apply to both OSPFv2 [RFC2328] and OSPFv3 [RFC5340]. The term OSPF is used to refer to both versions.

1.1. Requirements Language

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

2. Router Interface Parameters

This document defines the following router interface configurable parameters:

DoNotAdvertiseLink
Indicates whether or not the link is advertised on the local router-LSA. If set to "enabled," the router MUST NOT include a corresponding interface description in its router-LSA. The router MUST NOT originate other LSAs related to the link or its addresses. Enabling this interface parameter overrides the setting of LinkLSASuppression [RFC5340].

DoNotRequestAndIgnoreLSAs
Indicates whether or not the router should request and use LSAs from other routers on this interface. If set to "enabled," the router MUST consider its Link state request list empty. Also, the router MUST consider the LS age of any received LSA to be equal to MaxAge and process it according to Section 13 of [RFC2328].

3. Monitoring Interface

By using the interface parameters specified in Section 2, a router can treat all neighbors on the interface as Monitor Nodes. To do so, DoNotAdvertiseLink and DoNotRequestAndIgnoreLSAs SHOULD be configured simultaneously. If either parameter is configured on a broadcast or NBMA interface, the router MUST NOT participate in the Designated Router (DR) selection process.

Enabling DoNotAdvertiseLink by itself results in any LSAs originated by the Monitor Node not being resolved in the routing table.

If only DoNotRequestAndIgnoreLSAs is enabled, the router MUST treat the link as a stub network. Note that the neighbor information (corresponding to the Monitor Node) is not advertised.
4. The Monitor Node Option

This document defines a new Option in the Extended Options and Flags (EOF) Link-Local Signaling (LLS) TLV [RFC5613]. The new option is called Monitor (M-bit) and has a value of TBD.

When set, the M-bit indicates that the originating router is a Monitor Node. Other routers on the same link MUST:

* Consider the Monitor Node ineligible for the DR selection process.
* Consider its Link state request list empty with respect to the Monitor Node.
* Consider the LS age of any LSA received from the Monitor Node is equal to MaxAge.

If the Monitor Node is one of only two routers on an interface, the other router MUST NOT include a corresponding interface description in its router-LSA. Furthermore, other LSAs related to the link or its addresses MUST NOT be originated. This situation overrides the setting of LinkLSASuppression.

5. Operational Considerations

The use of the monitoring interface (Section 3) applies to all other routers on the same interface. While the Monitor Node option (Section 4) applies to only the router signaling the M-bit. Network administrators should use the Monitor Node option in transit interfaces where one router is a Monitor Node.

If the Monitor Node is the only other router on an interface, the link information can be advertised (as a stub link) if only DoNotRequestAndIgnoreLSAs is enabled.

The deployment of the Monitoring Interface (Section 3) requires that only the non-Monitor Node supports this specification. On the other hand, the Monitor Node Option (Section 4) requires all nodes on the interface to support the functionality. If support is not present in all the routers on the link, the Monitor Node will be eligible to be a DR, and its information may be flooded through the network.

6. Acknowledgements

TBD
7. IANA Considerations

IANA is requested to allocate a value (TBD) from the "LLS Type 1 Extended Options and Flags" registry for the M-bit (Section 4).

8. Security Considerations

The security considerations documented in [RFC2328], [RFC5340], and [RFC5613] apply to this extension.

This document defines a new type of node, called a Monitor Node, intended only to receive information from its neighbors and not send any. If the LSAs from the Monitor Node are not ignored, they will be flooded throughout the network. A rouge Monitor Node may advertise LSAs with an Advertising Router field that doesn’t correspond to its router ID. This type of vulnerability is not new, but it is already present in the base specification.

Even though it is expected that the local network operator deploys any Monitor Node, authentication mechanisms such as those specified in [RFC5709], [RFC7474], [RFC4552], or [RFC7166] SHOULD be used.

9. References

9.1. Normative References


9.2. Informative References


Authors’ Addresses

Alvaro Retana
Futurewei Technologies, Inc.
2330 Central Expressway
Santa Clara, CA 95050
United States of America
Email: aretana@futurewei.com

Lin Han
Futurewei Technologies, Inc.
2330 Central Expressway
Santa Clara, CA 95050
United States of America
Email: lin.han@futurewei.com
5G Distributed UPFs for 5G Multicast and Broadcast Services (5MBS)
draft-tjiang-dmm-5g-dupf-5mbs-00

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

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 8 September 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.
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.zhang-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:
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:

![Network Architecture Diagram]

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


6.2. Informative References


Author’s Address

Tianji Jiang
China Mobile
Email: tianjijiang@chinamobile.com
Abstract

This document describes evolution of mobile user plane in 5G, including distributed UPFs and alternative user plane implementations that some vendors/operators are pushing without changing 3GPP architecture/signaling. This also sets the stage for discussions in a companion document about potentially integrating UPF and Acess Node (AN) in a future generation (xG) of mobile network.

Status of This Memo

This Internet-Draft is submitted in full conformance with the provisions of BCP 78 and BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF). Note that other groups may also distribute working documents as Internet-Drafts. The list of current Internet-Drafts is at https://datatracker.ietf.org/drafts/current/.

Internet-Drafts are draft documents valid for a maximum of six months and may be updated, replaced, or obsoleted by other documents at any time. It is inappropriate to use Internet-Drafts as reference material or to cite them other than as "work in progress."

This Internet-Draft will expire on 7 September 2022.

Copyright Notice

Copyright (c) 2022 IETF Trust and the persons identified as the document authors. All rights reserved.

This document is subject to BCP 78 and the IETF Trust’s Legal Provisions Relating to IETF Documents (https://trustee.ietf.org/license-info) in effect on the date of publication of this document. Please review these documents carefully, as they describe your rights and restrictions with respect to this document.
1. Current User Plane in 5G

Mobile User Plane (MUP) in 5G [3GPP-23.501] has two distinct parts: the Access Network part between UE and AN/gNB, and the Core Network part between AN/gNB and UPF.
For the core network (CN) part, N3 interface extends the PDU layer from AN/gNB towards the PSA UPF, optionally through I-UPFs and in that case N9 interface is used between I-UPF and PSA UPF. Traditionally, UPFs are deployed at central locations and the N3/N9 tunnels extend the PDU layer to them. The N3/N9 interface uses GTP-U tunnels that are typically over a VPN over a transport infrastructure. While N6 is a 3GPP defined interface, it is for reference only and there is no tunneling or specification involved – it is simply a direct IP (in case of IP PDU session) or Ethernet (in case of Ethernet PDU session) connection to the DN.

At the AN/gNB, relay is done between the radio layer and the GTP-U layer. At the PSA UPF, routing/switching is done for IP/Ethernet before GTP-U encapsulation (for downlink traffic) or after GTP-U decapsulation (for uplink traffic).

2. MUP Evolution in 5G: Distributed UPFs

With MEC, ULCL UPFs are deployed closer to gNBs, while centralized PSA UPFs are still used to provide persistent IP addresses to UEs.

In fact, even PSA UPFs could be distributed closer to gNBs and then the N3 interface becomes very simple – over a direct or short transport connection between gNB and UPF (or even an internal connection if the gNB and UPF are hosted on the same server). On the other hand, since the UPF to DN connection is direct, the DN becomes a VPN (e.g., IP VPN in case of IP PDU sessions or EVPN in case of Ethernet PDU sessions) over a transport infrastructure, most likely the same transport infrastructure for the VPN supporting the N3/N9 tunneling in centralized PSA UPF case, as shown in the following picture:
The central PSA UPF is no longer needed in this case. Distributed UPF1/UPF2 connect to VRF1 on PE1/PE2 and VRF1 is for the VPN of the DN that UE1/UE2 access. There is also a PE3 for other sites of the VPN, which could be wireline sites including sites providing Internet access.

UEs may keep their persistent IP addresses even when they re-anchor from one PSA UPF to another. In that case, for downlink traffic to be sent to the right UPF, when a UE anchors at a UPF the UPF advertises a host route for the UE and when a UE de-achors from a UPF the UPF withdraws the host route.

While this relies on host routes to direct to-UE traffic to the right UPF, it does not introduce additional scaling burden compared to centralized PSA UPF model, as the centralized UPFs need to maintain
per-UE forwarding state (in the form of PDRs/FARs) and the number is the same as the number of host routes that a hub DN router (e.g. vrf1 on PE3 for internet access) need to maintain in the distributed PSA UPFs model. Since the host routes may be lighter-weighted than the PDRs/FARs, the total amount of state may be actually smaller in the distributed model.

For UE-UE traffic, the distributed PSA UPFs may maintain host routes that they learn from each other. With that the UE-UE traffic may take direct UPF-UPF path instead of going through a hub router in the DN (equivalent of central UPF). That is important in LAN-type services that require low delay. Alternatively, the distributed UPFs may maintain only a default route pointing to the hub router like PE3 (besides the host routes for locally anchored UEs). That way, they don’t need to maintain many host routes though UPF-UPF traffic has to go through the hub router (and that is similar to all traffic going through a central PSA UPF).

Optionally, even the host routes for locally anchored UEs can be omitted in the FIB of local UPF. Traffic among local UEs can be simply routed to the hub router following the default route, who will then send back to local UPF using VPN tunnels (MPLS or SRv6) that are stitched to GTP tunnels for destination UEs.

2.1. Advantages of Distributed PSA UPFs

Distributed PSA UPFs have the following advantages:

* MEC becomes much simpler - no need for centralized PSA UPF plus ULCL UPFs, and no need for special procedures for location based edge server discovery.

* For LAN-type services, UE-UE traffic can be optimized (no need to go through centralized PSA UPFs) when UPFs maintain host routes. It also allows seamless integration of services across wireline/wireless-connected customer sites.

* N3/N9 tunneling is simplified

In particular, there is now only short/simple N3 tunneling between AN/gNB and local UPFs in proximity. Among the distributed UPFs and other DN sites, versatile IETF/wireline VPN technologies are used instead. For example:
* Any tunneling technology — MPLS, SR-MPLS or SRV6 — with any traffic engineering/differentiation capabilities can be used. Removal of the GTP/UDP header (and IPv4/IPv6 header in case of MPLS data plane) brings additional bandwidth savings in the transport infrastructure.

* Any control plane model for VPN can be used — traditional distributed or newer controller based route advertisement.

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

Notice that, since UPF has routing functions, depending on the capability of a UPF device, it may even be possible for a UPF device to act as a VPN PE. That can be done in one of the two models:

* The UPF function and VPN PE function are separate but co-hosted on the same device with a logical/internal N6 connection between them.

* The UPF and VPN PE function are integrated and the PDU sessions become VPN PE-CE links.

The second model is especially useful when a UE is multi-homed to different EVPN PEs in case of Ethernet PDU sessions — EVPN’s all-active multihoming procedures can be utilized.

2.2. Enablers of Distributed PSA UPFs

To distribute PSA UPFs, if persistent addresses must be used for UEs, the SMF must be able to allocate persistent IP addresses from a central pool even when a UE re-anchors at different PSA UPFs (e.g. due to mobility). If DHCPv4 is used, either the SMF acts as a central DHCP server or it relays DHCP requests to a central DHCP server on the DN.

The distributed PSA UPFs must be able to advertise host routes in the DN. This should not be a problem since a UPF is essentially a router in that it routes traffic between DN and UEs (that are connected via PDU sessions).

Notice that, advertising host routes for persistent IP addresses is no different from advertising MAC addresses in case of Ethernet PDU sessions.
3. MUP Evolution in 5G: Alternative Implementation Options

3.1. GTP vs. SRv6 vs. MPLS tunneling

3GPP specifies that all tunneling (e.g. N3/N9) use GTP, whose encapsulation includes IP header, UDP header and GTP header. The tunnel is between 3GPP NFs (e.g. gNBs and UPFs) over an IP transport, and the IP transport may be a VPN over the multi-service transport infrastructure of an operator.

There have been proposals to replace GTP with SRv6 tunnels for the following benefits:

* Traffic Engineering (TE) and Service Function Chaining (SFC) capability provided by SRv6
* Bandwidth savings because UDP and GTP headers are no longer needed

While 3GPP has not adopted the proposal, and GTP can be transported over SRv6 (as overlay, instead of SRv6 replacing GTP), some operators still prefer to replace GTP with SRv6 "under the hood". That is, while RAN/UPF still use N2/N4 signaling, the actual tunnel are no longer GTP but SRv6 based on GTP parameters signaled by N2/N4. The SRv6 tunnel could be between two NFs, or a GW could be attached to an NF that still use traditional GTP and the GW will convert GTP to/from SRv6. This is specified in [I-D.ietf-dmm-srv6-mobile-uplane].

Similarly, if an operator prefers to use MPLS, a GTP tunnel can also be replaced with an MPLS PW instead of an SRv6 tunnel. Compared with SRv6, it is even more bandwidth efficient (no need for a minimum 40-byte IPv6 header) and SR-MPLS can also provide TE/SFC capabilities. This is specified in [I-D.zzhang-pals-pw-for-ip-udp-payload].

Note that, While only IPv6 can scale to the 5G requirements for the transport infrastructure, it does not mean MPLS can not be used as data plane in the IPv6 network.

3.2. Routing Based UPF

Traditionally, a UPF is implemented to follow 3GPP specifications. Specifically, N4 signaling is used for SMF to instruct a UPF to set up its session state in terms of PDRs/FARs. On N6 side, a UPF receives downlink traffic with destination addresses that are covered by the UPF’s address range for its anchored UEs. The packet is matched against the installed PDRs and forwarded according to the associated FARs. On N3 side, a UPF decapsulates GTP+UDP+IP header of uplink traffic and uses the TEID to identify the DN where inner IP
routing or Ethernet switching is done.

[I-D.mhkk-dmm-srv6mup-architecture] specifies a new SRv6 based MUP architecture. When it is applied to a 3GPP based mobile architecture:

* BGP signaling from a MUP Controller replaces N4 signaling from SMF. N4 signaling is still used between the MUP Controller and SMF - from SMF’s point of view it is just interacting with a traditional UPF as usual.

* A MUP GW becomes a distributed UPF for uplink traffic.

* A MUP PE, which is different from a usually central PSA UPF, becomes a UPF for downlink traffic, in that traffic to each UE is placed into a different tunnel that is stitched to a GTP tunnel for that UE by a MUP GW (no route lookup is needed on the MUP GW for the downlink traffic).

In this approach UE to UE traffic may still optionally go through the central PSA UPF. This is similar to that a hub router may be used in Section 2.

This approach can be viewed as a specific way of implementing/deploying distributed UPFs discussed in Section 2. It does have the advantage that from SMF’s point of view, nothing is different from before – both from N4 signaling and deployment model point of view.

While the above is specific to SRv6, a similar MPLS based approach will be specified separately for operators who prefer MPLS data plane, and it can even be SR-agnostic.

4. Security Considerations

To be provided.

5. Informative References

[I-D.ietf-dmm-srv6-mobile-uplane]
Authors’ Addresses

Zhaohui Zhang
Juniper Networks

Email: zzhang@juniper.net

Keyur Patel
Arrcus

Email: keyur@arrcus.com

Tianji Jiang
China Mobile

Email: tianjiwang@chinamobile.com
Abstract

[I-D.zzhang-dmm-5g-distributed-upf] describes evolution of mobile user plane in 5G, including distributed UPFs and alternative user plane implementations that some vendors/operators are pushing without changing 3GPP architecture/signaling. Building on top of that, this document further discusses potentially integrating UPF and Acess Node (AN) in a future generation (xG) of mobile network.

This document is not an attempt to do 3GPP work in IETF. Rather, it discusses potential integration of IETF/wireline and 3GPP/wireless technologies - first among parties who are familiar with both areas and friendly with IETF/wireline technologies. If the ideas in this document are deemed reasonable, feasible and desired among these parties, they can then be brought to 3GPP for further discussions.
1. MUP Evolution in xG

[I-D.zzhang-dmm-5g-distributed-upf] describes evolution of mobile user plane in 5G [3GPP-23.501], including distributed UPFs and alternative user plane implementations that some vendors/operators are pushing without changing 3GPP architecture/signaling.

This section discusses potential MUP evolution in a future generation (referred to as xG) of mobile networks. It does involve changes in 3GPP architecture and signaling, so the purpose of this section is to share the ideas in IETF community first. If it gains consensus within IETF community especially among mobile operators, then the proposal may be brought to 3GPP community for further discussions.

1.1. Integrated AN/UP Function

In the distributed UPF model for 5G [I-D.zzhang-dmm-5g-distributed-upf], AN and UPF are separate functions connected by N3 tunneling over a short/internet transport connection. Routing happens on the UPF between the DN and UEs over the N3 tunnels, and relay happens on the AN between the N3 tunnels and AN protocol stack.
With AN and UPF functions more and more disaggregated and virtualized even in 5G, it is becoming more and more feasible and attractive to integrate the AN and UPF functions, eliminating the N3 tunneling and the relay on AN entirely. The combined function is referred to as ANUP in this document, which does routing between DN and UEs over the AN protocol stack directly:

With this architecture, 3GPP and IETF technologies are applied where they are best applicable: 3GPP technologies responsible for radio access and IETF technologies for the rest. As IETF technologies continue to evolve, they can be automatically applied in mobile networks without any changes in 3GPP architecture/specification.

Some advantages of this new architecture include:
Any kind of tunnels can be used for the DN VPN, whether it is MPLS or SRv6, w/o the overhead of UDP/GTP encapsulation compared to GTP tunneling. Network slicing function is still supported (even with current GTP tunneling the transport network need to instantiate slices for N3/N9 tunnels as well).

* 5G-LAN and MEC become native applications (PDU sessions terminate into the closest ANUP and routed/switched to various DNAs).

* MBS becomes very simple - the ANUP gets the multicast traffic from the DN and then use either shared radio bearer or individual bearers to send to interested UEs.

Because the ANUP already implement the routing/switching functions, even the PE functions (for the DN VPN) could be optionally integrated into it, further streamlining end-to-end communication by reducing NFs and connections between them.

1.2. Separate AN/UP Functions Connected by Pseudo Wires

There are still cases where separate AN/UP functions are desired/required:

* An MNO may want to deploy one UPF for a cluster of ANs in proximity in some scenarios/locations

* An MNO may support MVNOs who have their own UP functions but make use of the hosting MNO’s ANs

* Home Routed roaming requires separate HPLMN UPs and VPLMN ANs

All these still require tunneling between ANs and UPs, but the tunneling can be achieved via IETF’s Pseudo Wire technology [RFC3985] as shown in the following diagram. Note that, using PW is just an option - GTP can still be used if that is desired.
1.2.1. Details on Pseudo Wire

This section provides some details on how PWs are used for the AN-UP tunneling.

From [RFC3985]:

On the AN, relay happens between the AN protocol stack and PW protocol stack. On the UP (at the right side of the above diagram), routing happens at PDU layer (over the PW that is stitched to the AN protocol stack on the AN) between UE1 and N6 connection to VRF1 on PE3. The UP is either one in HPLMN in Home Routed Roaming case (and the AN is in the VPLMN), or one in VMNO (and the AN is in the hosting MNO), or one for a cluster of ANs.
This document an architecture for Pseudo Wire Emulation Edge-to-Edge (PWE3) in support of [RFC3916]. It discusses the emulation of services such as Frame Relay, ATM, Ethernet, TDM, and SONET/SDH over packet switched networks (PSNs) using IP or MPLS. It presents the architectural framework for pseudo wires (PWs), defines terminology, and specifies the various protocol elements and their functions.

PWs provide the following functions in order to emulate the behavior and characteristics of the native service.

- Encapsulation of service-specific PDUs or circuit data arriving at the PE-bound port (logical or physical).
- Carriage of the encapsulated data across a PSN tunnel.
- Establishment of the PW, including the exchange and/or distribution of the PW identifiers used by the PSN tunnel endpoints.

The payload is classified into the following generic types of native data units:

- Packet
- Cell
- Bit stream
- Structured bit stream

When applied to tunneling between AN and UP, the PW payload type is "Packet" - IP packet or Ethernet frame (that is the over the SDAP layer between UE and AN) for IP or Ethernet PDU session respectively. In case of Unstructured PDU session type, the PW payload type would be "Bit stream" or "Structured bit stream".

Also from [RFC3985]:

Zhang & Patel Expires 7 September 2022 [Page 6]
Figure 2 illustrates the network reference model for point-to-point PWs.

The following explains the mapping to AN-UP tunneling:

* CE1 corresponds to a UE and PE1 corresponds to the AN

* The radio link between CE1/UE and PE1/AN is the AC in PW architecture. PDU session is the Emulated Service. Pseudo Wire corresponds to the AN-UP tunnel. TSN tunnel corresponds to the UDP tunnel that transports N3/N9 in 5G.

* PE2 and CE2 together correspond to the UP. It could be viewed that the PE2 provides AN function (with the PW corresponding to the radio link) and CE2 provides the UP function.

* PE1 takes the PDU packet from UE (after decapsulate the SDAP stack), which is treated as PW payload, and sends to PE2 over the PW. PE2 decapsulates the PW encapsulation and exposes the PDU (like that a gNB decapsulates the SDAP stack), which is then terminated by CE2 (though PE2 and CE2 are integrated into a single UP).

In 5G Home Routed roaming architecture, there is a pair of I-UPFs between the two PLMNs - the N3 tunnel does not extend directly from a VPLMN’s AN to a HPLMN’s UPF. The same concept also exists in VPN/PW technology - the I-UPFs are comparable to a pair of ASBRs that provide Option-B inter-AS VPN/PW services.
2. Security Considerations

To be provided.

3. Acknowledgements

The authors thank Arda Akamn, Constantine Polychronopoulos, Sandeep Patel and Shraman Adhikary for their review, comments and suggestions to make this document and solution more complete.

4. References

4.1. Normative References


4.2. Informative References


Authors’ Addresses

Zhaohui Zhang
Juniper Networks
Email: zzhang@juniper.net

Keyur Patel
Arrcus
Email: keyur@arrcus.com