dmm K. Bogigeni
Internet-Draft Verizon
Intended status: Informational A. Akhavain
Expires: December 20, 2018 Huawei Canada Research Centre
T. Herbert
Quantonium
D. Farinacci
lispers.net
A. Rodriguez-Natal
G. Carofiglio
J. Auge
Cisco
S. Homma
NTT
June 18, 2018

Optimized Mobile User Plane Solutions for 5G
draft-bogineni-dmm-optimized-mobile-user-plane-01

Abstract

3GPP CT4 has approved a study item to study different mobility management protocols for potential replacement of GTP tunnels between UPFs (N9 Interface) in the 3GPP 5G system architecture.

This document provides an overview of 5G system architecture in the context of N9 Interface which is the scope of the 3GPP CT4 study item [TS.23.501-3GPP], [TS.23.502-3GPP], [TS.23.503-3GPP], [TS.23.203-3GPP], [TS.29.244-3GPP], [TS.29.281-3GPP], [TS.38.300-3GPP], and [TS.38.401-3GPP]. The requirements for the network functions and the relevant interfaces are provided.

Reference scenarios and criteria for evaluation of various IETF protocols are provided.

Several IETF protocols are considered for comparison: SRv6, LISP, ILA and several combinations of control plane and user plane protocols.

Status of This Memo

Table of Contents

1. Introduction and Problem Statement

3GPP CT4 WG has approved a study item [CT4SID] to study user-plane protocol for N9 in 5GC architecture as specified in [TS.23.501-3GPP] and [TS.23.502-3GPP] for Rel-15. This provides an opportunity to investigate potential limits of the existing user plane solution and potential benefits of alternative user plane solutions.

Include scope of CT4 study item (user plane)

Add SA2 study item and relevant scope of the work (control plane)

Scope of control plane and user plane considerations

IETF has some protocols for potential consideration as candidates. These protocols have the potential to simplify the architecture through reduction/elimination of encapsulation; use of native routing mechanisms; support of anchor-less mobility management; reduction of session state and reduction of signaling associated with mobility management.

This document comprehensively describes the various protocols and how they can be used in the 3GPP 5G architecture. Specifically Segment Routing v6 (SRv6), Locator Identifier Separation Protocol (LISP) and Identifier Locator Addressing (ILA) are described in the context of the 3GPP 5G architecture for several scenarios: as a replacement of GTP on N9; as a replacement of GTP in the whole system; integrated with transport; used in specific network slices, etc.

A comparison of the various protocols is also provided.

2. Conventions Used in This Document

In examples, "C:" and "S:" indicate lines sent by the client and server respectively.

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].

In this document, these words will appear with that interpretation only when in ALL CAPS. Lower case uses of these words are not to be interpreted as carrying significance described in RFC 2119.

In this document, the characters ">>" preceding an indented line(s) indicates a statement using the key words listed above. This convention aids reviewers in quickly identifying or finding the portions of this RFC covered by these keywords.

3. Overview of Existing Architecture and Protocol Stack

This section briefly describes the 5G system architecture as specified in 3GPP TS 23.501. The key relevant features for session management and mobility management are:

- Separate the User Plane (UP) functions from the Control Plane (CP) functions, allowing independent scalability, evolution and flexible deployments e.g. centralized location or distributed (remote) location.

   o Support concurrent access to local and centralized services. To
      support low latency services and access to local data networks, UP
      functions can be deployed close to the Access Network.

   o Support roaming with both Home routed traffic as well as Local
      breakout traffic in the visited PLMN.

```
         +------+ +------+     +------+
         | NSSF | | AUSF +-N13-+ UDM  |
         +------+ +------+     +------+
               \      |     /      \
                N22  N12   N8       N10
                 \    |    /         \
                 +----+----+       +--------+     +------+      +------+
     +-----------+   AMF   +- N11 -+  SMF   +- N7 -+  PCF +- N5 -+  AF  |
     |           +++-----+++       +---+----+     +--+---+      +------+
     |            ||      ||           |             |
     |            ||      |+-------------|----- N15 ---+
     N1          N2|+-N14-+           N4
     |            |                    |
  +--+--+         ++-------+       +---+----+       +------+
  |  UE +-- NR --+ (R)AN  +-- N3 --+  UPF   +-- N6 --+  DN  |
  +-----+         +--------+       ++------++       +------+
                                    |     |
                                    +--N9-+
```

    Figure 1: 5G System Architecture in Reference Point Representation

```
                        Service Based Interfaces
    ----+-----+-----+----+----+----+----+--------+-----+--------
        |     |     |    |    |    |    |        |     |
    +---+---+ |  +--+--+ | +--+--+ | +--+--+  +--+--+  |
    | NSSF  | |  | NRF | | | DSF | | | UDM |  | NEF |  |
    +-------+ |  +-----+ | +-----+ | +-----+  +-----+  |
          +---+----+  +--+--+  +---+--+  +-------------+--+
          |  AMF   |  | PCF |  | AUSF |  |      SMF        |
          +---+--+-+  +-----+  +------+  +-+-----------+--+
             N1  |                         |           |
   +-------+  |  |                         N4          N4
   | 5G UE |--+  |                         |           |
   +---+---+     N2                  +-----+-+     +---+---+      +----+
       |         |      +----N3------+  UPF  +-N9--+  UPF  +--N6--+ DN |
       |         |      |            ++----+-+     +-------+      +----+
       |         |      |             |    |
       |    +---+------+-+            +-N9-+
       +-----|    gNB    |
             +------------+
```

Figure 2: 5G Service Based Architecture

This document focuses on the N9 interface which represents the user data plane between UPFs in 5G architecture.

```
                               +-----------------+
                               |       SMF       |
                               +-+-------------+-+
                                 |             |
                                 N4            N4
                                 |             |
      +-----------+        +------+-+        +-+------+
      |  gNB/RAN  |---N3---+  UPF   |---N9----|  UPF   +---N6----
      +-----------+        +--------+        +--------+
```

Figure 3: N3, N4, N9, and N6 interfaces in 5G Service Based Architecture

3GPP specifies two roaming model namely the Local Break Out(LBO) and the Homre Routed (HR) model.

In LBO model, visted operator routes user traffic locally through UPFs that are local to the visted operator. In this model, the SMF and all UPF(s) involved by the PDU Session are located and are under the control of the Visited PLMN (VPLMN).

In HR model, user traffic is routed to the UPF in Home PLMN(HPLMN) via the UPF in the visited network. In this scenario, the SMF in HPLMN (H-SMF) selects the UPF(s) in the HPLMN and the SMF in VPLMN(V-SMF) selects the UPF(s) in the VPLMN. In this model, the UE obtains services from its home network. Here, the UPF acting as PGW resides in home network, and can directly communicate with policy and billing system.

A given UE can have multiple simultaneous PDU sessions with different roaming model. In these scenarios, the HPLMN uses subscription data per Data Netwrok Name(DNN) and per Single Network Slice Selection Assistance Information(S-NSSAI) to determine PDU sessions's roaming model.

In the HR roaming model:

o The NAS SM terminates at the V-SMF.

o The V-SMF forwards SM related informaton to the SMF in the HPLMN.

o The V-SMF sends UE's Subscription Permanent Identifier(SUPI) to the H-SMF during the PDU session establishment procedure.

o The V-SMF sends the PDU Session Establishment Request message to the H-SMF along with the S-NSSAI with the value from the HPLMN.

o The H-SMF obtains subscription data directly from the Unified Data Management(UDM) and is responsible for checking the UE request with regard to the user subscription, and may reject the request in case of mismatch.

o The H-SMF may send QoS requirements associated with a PDU Session to the V-SMF. This may happen at PDU Session establishment and after the PDU Session is established. The interface between H-SMF and V-SMF is also able to carry (N9) User Plane forwarding information exchanged between H-SMF and V-SMF. The V-SMF may check QoS requests from the H-SMF with respect to roaming agreements.

o The AMF selects a V-SMF and a H-SMF, and provides the identifier of the selected H-SMF to the selected V-SMF.

o The H-SMF performs IP address management procedure based on the selected PDU session type.

Local Breakout and Home Routed roaming models are depicted in the two figures below.

```
                                   VPLMN      |     HPLMN
    ----------+-------+------+------+----    N32 -------+------+----
                                              |
                          +--+--+             |
                          | SMF |             |
                          +--+--+             |
                             |                |
   +-------+                N4                |
   | 5G UE +                 |                |
   +---+---+               +-----+--+         |
       |                   |        |         |
       |   +---+-+   +-+-+    +-+-+  +----+   |
       +---| gNB |---|UPF|-N9-|UPF|--| DN |   |
           +-----+   +-+-+    +---+  +----+   |
```

Figure 4: Roaming 5G System Architecture- Local Breakout Scenario

```
                              VPLMN   |      HPLMN
    ----------+------+------+-----  N32 --------+-------+------+-----+--
                            |         |         |
                         +--+--+      |      +--+--+
                         | SMF |      |      | SMF |
                         +--+--+      |      +--+--+
                            |         |         |
   +-------+                |         |         |
   | 5G UE |                |         |         |
   +---+---+               N4         |         N4
       |                    |                   |
       |     +-+---+     +--+--+             +--+--+      +----+
       +-----| gNB |-----| UPF |-----N9------| UPF |------| DN |
             +-----+     +--+--+             +-----+      +----+
```

Figure 5: Roaming 5G System Architecture- Home Routed Scenario

Figure 6 depicts the non-roaming architecture for UEs concurrently accessing two (e.g. local and central) data networks using multiple PDU Sessions, using the reference point representation. This figure shows the architecture for multiple PDU Sessions where two SMFs are selected for the two different PDU Sessions. However, each SMF may also have the capability to control both a local and a central UPF within a PDU Session.

```
                        Service Based Interfaces
    ---------+------------+------------------+-----------------------
                          |                  |
                       +--+--+            +--+--+
                       | SMF |            | SMF |
                       +--+--+            +--+--+
                          |                  |
   +--------+            N4                 N4
   | 5G UE  |             |                  |
   +---+----+          +--+--+    +-----+    +-----------+
       |           +---| UPF |----| DN  |    |           |
       |           |   +-----+    +-----+    |           |
       |     +-+----+-+                   +--+--+     +--+--+  +-----+
       +-----| gNB    |-------------------| UPF |--N9-| UPF |--| DN  |
             +--------+                   +-----+     +-----+  +-----+
```

Figure 6: Non-roaming 5G System Architecture for multiple PDU
Sessions Service Based Interface

```
                        Service Based Interfaces
    ---------+-----------------------+------------------------
                                     |
                                  +--+--+
                                  | SMF |
                                  +--+--+
                                     |
   +--------+                 +------+-------+
   | 5G UE  |                 |              |
   +---+----+                N4             N4
       |     +-+----+      +--+--+      +--+--+     +-----+
       +-----| gNB  |------| UPF |----N9--| UPF |----| DN  |
             +------+      +--+--+      +-----+     +-----+
                              |
                           +--+--+
                           |  DN |
                           +-----+
```

Figure 7: Non-roaming 5G System Architecture for Current Access to
Two (e.g. local and central) Data Networks (single PDU Session option

Figure 7 depicts the non-roaming architecture in case concurrent access to two (e.g. local and central) data networks is provided within a single PDU Session.

The User plane function (UPF) is the function relevant to this evaluation and the N9 interface between two UPFs [23501].

The User Plane Function (UPF) handles the user plane path of PDU sessions. The UPF transmits the PDUs of the PDU session in a single tunnel between 5GC and (R)AN. The UPF includes the following functionality. Some or all of the UPF functionalities may be supported in a single instance of a UPF. Not all of the UPF functionalities are required to be supported in an instance of user plane function of a Network Slice.

The following provides a breif list of main UPF functionalities.Please refer to section 6.2.3 3GPP TS 23.501 for detailed description of UPF and its functionalities.

o Anchor point for Intra-/Inter-RAT mobility (when applicable)"

o Sending and forwarding of one or more end marker to the source NG-RAN node

o External PDU Session point of interconnect to Data Network.

o PDU session type: IPv4, IPv6, Ethernet, Unstructured (type of PDU totally transparent to the 5GS)

o Activation and release of the UP connection of an PDU session, upon UE transition between the CM-IDLE and CM-CONNECTED states(i.e. activation and release of N3 tunnelling towards the access network)

o Data forwarding between the SMF and the UE or DN (e.g. IP address allocation or DN authorization during the establishment of a PDU session)

o Packet routing and forwarding (e.g. support of Uplink classifier to route traffic flows to an instance of a data network, support of Branching point to support IPv6 multi-homed PDU session>

o Branching Point to support routing of traffic flows of an IPv6 multi-homed PDU session to a data network, based on the source Prefix of the PDU

o User Plane part of policy rule enforcement (e.g. Gating, Redirection, Traffic steering)

o Uplink Classifier enforcement to support routing traffic flows to a data network, e.g. based on the destination IP address/Prefix of the UL PDU

o Lawful intercept (UP collection)

o Traffic usage reporting

o QoS handling for user plane including:

   * packet filtering, gating, UL/DL rate enforcement, UL/DL Session-AMBR enforcement (with the Session-AMBR computed by the UPF over the Averaging window provisioned over N4, see subclause 5.7.3 of 3GPP TS 23.501), UL/DL Guaranteed Flow Bit Rate (GFBR) enforcement, UL/DL Maximum Flow Bit Rate (MFBR) enforcement, etc

   * marking packets with the QoS Flow ID (QFI) in an encapsulation header on N3 (the QoS flow is the finest granularity of QoS differentiation in the PDU session)

   * enabling/disabling reflective QoS activation via the User Plane, i.e. marking DL packets with the Reflective QoS Indication (RQI) in the encapsulation header on N3, for DL packets matching a QoS Rule that contains an indication to activate reflective QoS

o Uplink Traffic verification (SDF to QoS flow mapping, i.e. checking that QFIs in the UL PDUs are aligned with the QoS Rules provided to the UE or implicitly derived by the UE e.g. when using reflective QoS)

o Transport level packet marking in the uplink and downlink, e.g. based on 5QI and ARP of the associated QoS flow.

o Downlink packet buffering and downlink data notification triggering: This includes the support and handling of the ARP priority of QoS Flows over the N4 interface, to support priority mechanism:

   * "For a UE that is not configured for priority treatment, upon receiving the "N7 PDU-CAN Session Modification" message from the PCF with an ARP priority level that is entitled for priority use, the SMF sends an "N4 Session Modification Request" to update the ARP for the Signalling QoS Flows, and sends an "N11 SM Request with PDU Session Modification Command" message to the AMF, as specified in clause 4.3.3.2 of TS 23.502.

   * "If an IP packet arrives at the UPF for a UE that is CM-IDLE over a QoS Flow which has an ARP priority level value that is entitled for priority use, delivery of priority indication during the Paging procedure is provided by inclusion of the ARP

in the N4 interface "Downlink Data Notification" message, as specified in clause 4.2.3.4 of TS 23.502."

o ARP proxying as specified in [RFC1027] and / or IPv6 Neighbour Solicitation Proxying as specified in [RFC4861] functionality for the Ethernet PDUs. The UPF responds to the ARP and / or the IPv6 Neighbour Solicitation Request by providing the MAC address corresponding to the IP address sent in the request.

o Packet inspection (e.g. Application detection based on service data flow template and the optional PFDs received from the SMF in addition)

o Traffic detection capabilities.

  * For IP PDU session type, the UPF traffic detection capabilities may detect traffic using traffic pattern based on at least any combination of:

    + PDU session

    + QFI

    + IP Packet Filter Set. Please refer to section 5.7.6.2 of 3GPP TS 23.501 for further details.

  * For Ethernet PDU session type, the SMF may control UPF traffic detection capabilities based on at least any combination of:

    + PDU session

    + QFI

    + Ethernet Packet Filter Set. Please refer to section 5.7.6.3 of 3GPP TS 23.501 for further details.

o Network slicing Requirements for different MM mechanisms on different slice. The selection mechanism for SMF to select UPF based on the selected network slice instance, DNN and other information e.g. UE subscription and local operator policies.

The following information is sent in an encapsulation header over the N3 interface. N9 needs to support that.

o QFI (QoS Flow Identifier), see subclause 5.7.1 of 3GPP TS 23.501.

o RQI (Reflective QoS Identifier), see subclause 5.7.5.4.2 of 3GPP TS 23.501.

o Support of RAN initiated QoS Flow mobility, when using Dual connectivity, also requires the QFI to be sent within End Marker packets. See subclause 5.11.1 of 3GPP TS 23.501 and subclause 4.14.1 of 3GPP TS 23.502 respectively.

GTPv1-U as defined in 3GPP TS 29.281 is used over the N3 and N9 interfaces in Release 15. Release 15 is still work-in-progress and RAN3 will specify the contents of the 5GS Container. It is to be decided whether CT4 needs to specify new GTP-U extension header(s) in 3GPP TS 29.281 for the 5GS Container.

A GTP-U tunnel is used per PDU session to encapsulate T-PDUs and GTP-U signaling messages (e.g. End Marker, Echo Request, Error Indication) between GTP-U peers.

A 5GS Container is defined as a new single GTP-U Extension Header over the N3 and N9 interfaces and the elements are added to this container as they appear with the forthcoming features and releases. This approach would allow to design the 5GS information elements independently from the tunneling protocol used within the 5GS, i.e. it would achieve the separation of the Transport Network Layer (TNL) and Radio Network Layer (RNL) as required in 3GPP TR 38.801 subclause 7.3.2. This would allow to not impact the RNL if in a future release a new transport network layer (TNL) other than GTP-U/UDP/IP (e.g. GRE/IP) was decided to be supported.

The protocol stack for the User Plane transport for a PDU session is depicted below in Figure 8.

```
   +-----+                     |                       |          |
   | App +---------------------|-----------------------|----------|
   +-----+                     |                       | +------+ |
   | PDU +---------------------|-----------------------|-+ PDU  | |
   +-----+  +---------------+  |  +-----------------+  | +------+ |
   |     |  |\             /|  |  |\               /|  | |      | |
   |     |  |  \  Relay  /  |  |  |  \    Relay  /  |  | |      | |
   |     |  |    \     /    |  |  |    \       /    |  | |5G UP | |
   | AN  |  |     --+--     |  |  |     ---+---     |  | | Enc  | |
   | Pro |  |AN Pro | GTP-U +--|--+ GTP-U  |5GUP Enc+--|-+      | |
   | Lyrs|  | Lyrs  +-------+  |  +--------+--------+  | +------+ |
   |     +--+       |UDP/IP +--|--+ UDP/IP | UDP/IP +--|-+UDP/IP| |
   |     |  |       +-------+  |  +--------+--------+  | +------+ |
   |     |  |       |  L2   +--|--+  L2    |   L2   +--|-+  L2  | |
   |     |  |       +-------+  |  +--------+--------+  | +------+ |
   |     |  |       |  L1   +--|--+  L1    |   L1   +--|-+  L1  | |
   +-----+  +-------+-------+  |  +--------+--------+  | +------+ |
     UE            AN          N3          UPF        N9          N6
                                                             UPF
                                                    (PDU Session Anchor)
```

Legend:
o PDU layer: This layer corresponds to the PDU carried between the UE
    and the DN over the PDU session. When the PDU session Type is
    IPV6, it corresponds to IPv6 packets; When the PDU session Type
    is Ethernet, it corresponds to Ethernet frames; etc.
o GPRS Tunnelling Protocol for the user plane (GTP U): This protocol
    supports multiplexing traffic of different PDU sessions (possibly
    corresponding to different PDU session Types) by tunnelling user
    data over N3 (i.e. between the AN node and the UPF) in the
    backbone network. GTP shall encapsulate all end user PDUs. It
    provides encapsulation on a per PDU session level. This layer
    carries also the marking associated with a QoS Flow.
o 5G Encapsulation: This layer supports multiplexing traffic of
    different PDU sessions (possibly corresponding to different PDU
    session Types) over N9 (i.e. between different UPF of the 5GC).
    It provides encapsulation on a per PDU session level. This layer
    carries also the marking associated with a QoS Flow.

Figure 8

Sections on GTP-U, PFCP for Release 15 and SBI for Release 16

Is there a need for deployment scenarios (to address regarding IPv4/
IPv6 end-to-end, backhaul and underlay for Mobile core)?

4. Objectives and architectural requirements

The objective of this draft is to propose a set of candidate protocols for possible replacement of GTP-U at N9 interface. This section first reviews a list of architectural requirements that candidate solution should address. Then some identified limitations of GTP are discussed, as emerging from the different proposals. Finally, a set of reference scenarios for evaluations are proposed.

4.1. Architectural requirements

A comprehensive summary of GTP architecture, as well as architectural requirements collected for the various 3GPP specifications, are covered by [I-D.hmm-dmm-5g-uplane-analysis]. We summarize here the main architectural requirements:

- o TODO

4.2. GTP limitations and desired benefits

Although being different in terms of architecture or implementations, common characteristics emerge from the different proposals in particular with how they are positioned with respect to GTP, and the advertised benefits. We describe the main aspects here that will be further discussed in further sections, clarifying some terminology at the same occasion.

Tunnel and encapsulations

- o Not clear if required by the specification, or something that can be optimized

Signalling overhead

- o

Anchorless mobility

- o Data plane anchor

- o Control plane anchor

4.3. Reference Scenario(s) for Evaluation

Different proposals will be described for the following scenarios:

1. Non-Roaming Scenarios

   * UE-Internet Connectivity (mobility cases)

   * UE-UE IP Packet Flow (mobility cases)

   * UE-2 DNs with multiple PDU sessions

   * UE-2 DNs Single PDU session

2. Roaming Scenarios

   * Local Break out

   * Home Routed

Flows will be provided for mobility cases (UE mobility, UPF mobility) and session continuity cases (SSC Mode 1/2/3).

1. UE mobility SSC Mode 1

   * Single UPF

   * Multiple UPF

2. UE mobility SSC Mode 2

   * Single UPF

   * Multiple UPF

3. UE mobility SSC Mode 3

   * Single UPF

   * Multiple UPF

Each proposal will also describe how network slicing will be supported for the following configurations:

- Support for independent slices using GTP and/or other protocol will be covered. Mobility Management will be within each slice.

- Support for one UE connected to multiple slices using different mobility protocols will be described.

The criteria for evaluation will be the ability to support the above scenarios and identifying the impacts to N2, N3, N4, gNB, AMF and SMF. Reference procedures/flows for above use cases from existing 3GPP specs.

5. Mobility management architectures

5.1. Overview

Based on their use of identifiers and locators, mobility approaches can be broadly categorized in the three following classes:

Locator-based, or Map-and-Encapsulate architectures

IP communication relies solely on locators (host interfaces' addresses) that are also used as node/service identifiers at network layer. Such semantic overloading of IP addresses as both identifiers and locators does not allow to disentangle locators from location-independent traffic identifiers, thus complexifying mobility management.

As a result, traffic anchors and tunnels have been introduced to handle mobility while preserving the identifier exposed to the transport layer.

Locator-ID separation architectures

To overcome the limitations of purely locator-based architectures, there have been proposed the so called "identifier/locator separation" (or Loc/ID split) schemes, using separate namespaces for identifiers and locators, and a mapping system allowing routers to bind them together.

Loc/ID split has been originally proposed by LISP to solve the scalability challenges of Internet routing, and further adapted as a mobility management solution. This category includes other approaches reviewed in this document, namely ILA, ILNP, ILSR and one SRv6-based solution, which all share the requirement for a mapping system. This service can be centralized/decentralized or distributed where a correspondence of ID and locators is stored and updated upon changes.

ID-based architectures

A third class of approaches exists that redefines IP communication principles (i.e. network and transport layers) around location-independent identifiers [I-D.vonhugo-5gangip-ip-issues].

Information-Centric Networking (ICN) approaches fall into such class of approaches that we refer to as purely ID-based, or in that specific case, as name-based [I-D.irtf-icnrg-terminology]. Previous work has highlighted the interest of ICN for mobility management [RFC7476].

The rest of this section details the set of reviewed approaches, namely SRv6, LISP, ILSR, ILNP, ILA and hICN, as summarized in Figure 9. Each proposal consists in an overview with pointers to reference material for a more in depth description. The focus is then given to a discussion on its integration at N9 interface in light of the architectural requirements discussed in Section 4, and the additional benefits it brings with respect to GTP-U. Extensions to N3 interface as well as alternative deployments preserving GTP tunnels as discussed later in this document in Section 7.

```
|_ Mobility Management
   |_ Locator-based / Map-and-Encapsulate
      |_ Tunnelling
         |_ 3GPP / GTP-U                Sec. XX
      |_ Packet steering
         |_ SRv6 (backwards-compatible) Sec. 5.2.1
   |_ Loc/ID split
      |_ Packet steering
         |_ SRv6                        Sec 5.2.2
      |_ Encapsulation
         |_ LISP, LISP-MN, ILSR         Sec. 5.3
      |_ Address rewrite
         |_ Host-based address change
            |_ ILNP, esp. ILNPv6        Sec. 5.4
         |_ Network-based translation
            |_ ILA                      Sec. 5.5
   |_ ID-based
      |_ Information-Centric Networking
         |_ ID-based mobility / IPv6
            |_ Hybrid ICN               Sec. 5.6
```

Figure 9: Overview of reviewed approaches

5.2. SRv6

5.2.1. SRv6 in "traditional mode"

5.2.2. SRv6 in "enhanced mode"

5.3. LISP

5.3.1. LISP-MN

5.3.2. ILSR

5.4. ILNP

5.5. ILA

5.6. Hybrid ICN (hICN)

5.6.1. Description

hCIN Anchorless Mobility Management (hICN-AMM) refers to a novel mobility management approach, introduced in [I-D.auge-hicn-mobility], that leverages routable location-independent identifiers (IDs) and an Information-Centric Networking (ICN) communication model integrated in IPv6, (also referred to as Hybrid ICN, or hICN) [I-D.muscariello-intarea-hicn].

Such approach belongs to the category of pure ID-based mobility management schemes whose objective is (i) to overcome the limitations of traditional locator-based solutions like Mobile IP (conf)using locators as identifiers, (ii) to remove the need for a global mapping system as the one required by locator-identifier separation solutions.

Consumer and Producer mobility

In ICN and hICN endpoints can act as consumers and/or producers. Consumers when they emit requests for named data packets (so called Interests), producers when they send data packets in response to consumers request (pull-based transport model). Clearly a node can be a consumer and a producer at the same time (e.g. in a voice conversation).

Consumer and producer mobility are handled in a different way due to the pull-based request model. More specifically, consumer mobility is natively supported: consumers pull traffic by sending Interest packets towards named content (wherever produced/stored, the source is a priori unknown by the consumer). Interests are named-based forwarded using the information found in traversed routers' FIBs.

In case of consumer mobility, i.e. mobility of the endpoint issuing the requests, selection of a new available output interface and retransmission of not-yet-satisfied Interests is sufficient for data delivery to continue, independently from the underlying change of locators. Consumer mobility is fully anchorless with hICN, and does not incur any signalization nor tunneling overhead.

Producer mobility is not natively supported by ICN architecture, rather handled in different ways according to the selected producer mobility management scheme.

Anchorless mobility support

The selected mobility management scheme for hICN is MAP-Me, an anchorless producer mobility management solution originally proposed for ICN [I-D.irtf-icnrg-mapme] [MAPME] and further extended to hICN in [I-D.auge-hicn-mobility].

MAP-Me belongs to the class of anchorless approaches that relies on scope-limited forwarding updates triggered by producer mobility events to keep locally up-to-date FIB information for a low-latency guaranteed reroute of consumer Interests towards changing location of the producer. Forwarding and mobility management operations in hICN are based only location-independent identifiers, preserving coexistence with IP locators whose existence may be required by non-hICN services and by control/management plane operations specific to the considered network architecture.

Signaling of mobility is only required upon producer movements and limited in scope to current-to-previous network hops. Unlike routing updates, it is not necessary to update all routers' FIBs after a node has moved, but only those located on the path between the new and a former position of the producer. Scalability of producer mobility is guaranteed by an efficient and secure FIB update process with minimal and bounded path stretch.

The difference w.r.t. to other classes of approaches is that it does not require an anchor neither in forwarding plane (no tunnel, traffic does not need to pass through a specific network node), nor in the control plane (no rendez-vous point, no mapping system).

5.6.2. Insertion in N9 interface

Insertion of hICN in 5G IP infrastructure is facilitated by its design allowing a selective insertion of hICN capabilities in a few network nodes at the edge (no need for pervasive fully hICN network enablement), and to guarantee a transparent interconnection with hICN-unaware IP nodes, without using overlays.

The deployment of hICN routers allow to avoid the reliance on GTP tunnels, and to provide an agile transport and native anchorless mobility natively. The resulting protocol stack is showin in Figure 10. We remark that in the protocol layer, hICN is associated to IPv6 PDU layer, transported in N9 directly over L2.

```
       UE            5G-AN        N3         UPF        N9   UPF    N6
                                  |                     |            |
   +--------+                     |                     |            |
   |  App.  |--------------------------------------------------------|
   +--------+                     |                     | +--------+ |
   | IP PDU |                     |                     | | IP PDU | |
   | (hICN) |---------------------------------------------| (hICN) | |
   +--------+ +-----------------+ | +-----------------+ | |        | |
   |        | |\     relay     /| | |\     decap     /  | |        | |
   |        | | \_____________/ |-|-| \_____________/   | |        | |
   |        | |        | GTP-U  | | | GTP-U  |          | |        | |
   |        | |        +--------+ | +--------+          | |        | |
   |   5G   | |   5G   |  UDP   |-|-|  UDP   |          | |        | |
   |   AN   |-|   AN   +--------+ | +--------+          | |        | |
   |protocol| |protocol|   IP   |-|-|   IP   |          | |        | |
   | layers | | layers +--------+ | +--------+--------+ | +--------+ |
   |        | |        |   L2   |-|-|   L2   |   L2   |-|-|   L2   | |
   |        | |        +--------+ | +--------+--------+ | +--------+ |
   |        | |        |   L1   |-|-|   L1   |   L1   |-|-|   L1   | |
   +--------+ +-----------------+ | +-----------------+ | +--------+ |
                                  |                     |            |
```

Figure 10: Replacement of N9 interface - Protocol layers

5.6.3. Benefits

The appeal of purely ID-based architectures is that they move Loc/ID split one step further by embedding ID-awareness in the network and transport layer by default and as such completely decoupling data delivery from underlying network connectivity. The resulting mobility management solution is fully anchorless for both consumer and producer mobility. Forwarding is performed directly based on identifiers stored in routers' FIBs and no mapping of ID into locators is required. In this way, purely ID-based architectures remove the need to maintain a global mapping system at scale, and its intrinsic management complexity.

Additional benefits brought specifically by ICN principles motivate the consideration of ICN solutions for next generation mobility architectures, like for instance:

- o the flexibility of multi-source/multi-path connectionless pull-based transport. An example is the native support for consumer mobility, i.e. the transparent emission of data requests over multiple and varying available network interfaces during node mobility;

- o the opportunity to define fine-grained per-application forwarding and security policies (in the network, and in-between UPFs);
- o low-latency and multicast capabilities by means of in-path edge caching;
- o network-assisted transport.

An in depth analysis of benefits originating from the coupling between a purely identifier-based approach and from specific hICN properties can be found in [I-D.auge-hicn-mobility-deployment-options] along with some illustrative examples.

5.6.4. Deployment considerations

Partial insertion

The benefits previously described can be obtained by an upgrade of only a few selected routers at the network edge. The design of hICN allows the rest of the infrastructure to remain unmodified, and to leverage existing management and monitoring tools. There exists thus a tradeoff between incremental deployment and benefits which are proportionally related to the degree of hICN penetration.

End-to-end deployment

The deployment of an hICN stack in endpoints is the preferred option and offers the full range of benefits. Both the hICN forwarder and the transport stack are available as reference implementations based on the CICN project [CICN]. They are both designed to facilitate insertion on routers and end-user devices thanks to implementation in user space, one targetting high-performance, the other aiming at wide support from major vendors including iOS, Android, Linux, MacOSX and Windows.

Network-contained deployment

It is not always possible nor desirable to affect endpoints, and a deployment fully contained in the network is possible through the deployment of proxies. An example would be the deployment of HTTP proxies at the ingress and egress (resp. first and last UPFs), in order to benefit from content awareness in the network. Such configuration however reduces the flexibility and dynamic forwarding capabilities in endpoints. In particular, existing transport protocols have limited support for dynamically changing paths or network conditions.

Traffic that is not handled though hICN mechanisms can still benefit from the lower overhead and anchorless mobility capabilities coming from the removal of GTP tunnels, as well as dynamic forwarding capabilities that are inherent to the forwarding pipeline. This results from the ability to assign location-independent identifiers to endpoints. It preserves the advantage of removing the mapping system, and of a lightweight FIB update process. No encapsulation is required and packet headers are not modified, which allows the network to have visibility in the source and/or destination identifiers.

hICN in a slice

The use of hICN does not impose any specific slicing of the network. Rather, it can assist a transition of services towards hICN, and/or the coexistence of different hICN deployment options.

As an example of use of hICN in a slice, a service provider might for instance decide to use an hICN-enabled slice dedicated to video delivery, with appropriate mobility management, and dedicated hICN nodes with appropriate caching/forwarding strategies at places aggregating considerable number of user requests.

Coupling between hICN and SRv6

The association of hICN with other data planes technologies, such as SRv6, is investigated as a possibility to overcome the above-mentioned tradeoff yielding to a selective, yet fully beneficial insertion of hICN in IP networks. This would inherit all SRv6 advantages presented in SR-dedicated section of this document - namely "underlay" management (fast reroute, etc.), service chaining or fine-grained TE for instance - but also extend the reach of hICN on regular IP routers with SRv6 functionality.

One realization being to create SRv6 domains in between hICN nodes. The hICN router (through forwarding strategies) would then act as a control plane for SRv6 by specifying the list of SIDs to insert in the packet.

5.6.5. Summary

hICN proposes a general purpose architecture that combines the benefits of a pure-ID architecture with those of ICN. While a full deployment is recommended to make efficient use of available network resources, it is still possible to opt for a partial or phased deployment, with the associated tradeoffs that we have reviewed here.

An hICN enabled network offers native offloading capabilities thanks to the anchorless properties resulting from the pure-ID communication scheme. It does so without the need for a third party mapping system, and further requires no change in the 5G architecture nor in its control plane. The architecture will further leverage the incremental insertion of information centric functionalities through proxies or direct insertion in user devices as the technology gets adopted and deployed.

6. Integration of new data planes into the 5G framework

6.1. Integration of ID-LOC Architecture

```
                                 +----+----+
         +-------------N4--------+   SMF   +---------N4-----------+
         |                       +----+----+                      |
         |                            |                           |
         |                       +----+----+                      |
         |                       | ID-LOC  |                      |
         |                       | Mapping |      ID-LOC          |
         |               +------>| System  |<--control-plane      |
         |               |       +----+----+     |                |
         |               V                       V                |
     +---+---+      +----+----+             +----+----+      +---+---+
--N3-+ UPF-A +--N9--+ID-L Node+<--ID-LOC--->+ID-L Node+--N9--+ UPF-B +-N6--
     +-------+  GTP +----+----+ data-plane  +----+----+  GTP +-------+
```

Figure 11: 5G Integration with ID-LOC (GTP integration)

```
                                 +----+----+
         +-------------N4--------+   SMF   +---------N4-----------+
         |                       +----+----+                      |
         |                            |                           |
         |                       +----+----+                      |
         |                       | ID-LOC  |                      |
         |                       | Mapping |      ID-LOC          |
         |  +------------------->| System  |<--control-plane--+   |
         |  |                    +----+----+                  |   |
         |  V                                                 V   |
     +---+---+                                               +---+---+
--N3-+ UPF-A +<---------- N9 - ID-LOC data-plane ----------->+ UPF-B +-N6--
     +-------+                                               +-------+
```

Figure 12: 5G Integration with ID-LOC (GTP replacement)

An ID-LOC network architecture is able to decouple the identity of endpoints (ID) from their location in the network (LOC). Common ID-

LOC architectures are based on two main components, ID-LOC data-plane nodes and an ID-LOC mapping system.

ID-LOC data-plane nodes act upon received data traffic and perform ID-LOC data-plane operation. The specific operation that these ID-LOC data-plane nodes perform is based on the particular ID-LOC data-plane protocol that they implement. ID-LOC data-plane protocols are usually divided in two categories, (1) those that encapsulate ID-based data-plane packets into LOC-based data-plane packets and (2) those that transform the addresses on the data-plane packets from ID-based addresses to LOC-based addresses. SRv6 and LISP-DP protocols are examples of the former while the ILA protocol is an example of the latter.

The ID-LOC mapping system is a database that provides mappings of Identity to Location for ID-LOC data-plane nodes to use. Usually, ID-LOC architectures use an ID-LOC control-plane protocol to make available at the data-plane nodes the ID-LOC mappings that they need to operate. Examples of such ID-LOC control-plane protocols are LISP-CP and ILAMP.

When integrating ID-LOC architecture into the 5G framework there are several aspects to take into account. First the ID-LOC data-plane function needs to be performed in the data-plane path as the packets enter and leave the ID-LOC domain. On option for this is to deploy ID-LOC data-plane nodes adjacent to UPFs to perform the ID-LOC operation on the traffic as it leaves or enters the UPFs (as shown in Fig. Figure 11). In this case the ID-LOC data-plane protocol will be part of the N9 interface along with current GTP. Another option is to implement the ID-LOC data-plane function directly in the UPFs (as shown in Fig. Figure 12). In this case, these ID-LOC enabled UPFs will directly generate packets encapsulated or transformed and will be able to directly process packets encapsulated or transformed. In this case the ID-LOC protocol will completely replace GTP in the N9 interface.

Second, the Mapping System needs to contain the appropriate ID-LOC mappings in coordination with the SMF. In order to do so, the mappings in the Mapping System are populated either by the SMF directly or by the ID-LOC nodes that should be in synch with the SMF. In the former case, an interface from the SMF to the Mapping System is needed (as shown in Figs. Figure 11 and Figure 12).

See also section [REF] for discussion on an approach for incremental deployment of ID-LOC solutions in the 5G framework.

6.2. Existing control planes for ID-Loc separation architectures

- LISP-CP
  * for ILA
  * for SRv6
- ILAMP
- BGP

6.3. Integration of hICN

By operating directly on routers' FIBs for mobility updates, dynamic hop-by-hop forwarding strategies etc., hICN inherits the simplicity of IP forwarding and reuses IP routing protocols for ID prefixes advertisement and routing. In this way it removes the challenges of managing a distributed mapping service at scale (cache update/refresh, etc.). In addition it remains compatible with the exiting control plane architecture as proposed in the 3GPP standard, with no change required to N1, N2 or N4.

MAP-Me anchorless producer mobility management does not imply SMF interaction, but does not exclude neither to use SMF signaling to trigger MAP-Me updates or to handle FIB updates, at the condition to follow the same procedure described for MAP-Me. However, the absence of SMF interaction might be beneficial in case of dense deployments or failure of the central control entities (infrastructure-less communication scenarios) to empower distributed control of local mobility within an area.

6.4. Slicing

- Coexistence of several solutions in different slices

7. Alternative deployment options

7.1. Extensions to N3 interface

Although not strictly the object of study by 3GPP, previous solutions can (and would gain to) be extended beyond N9 to cover N3 interface too.

The immediate benefit is the complete removal of all GTP tunnels, along with associated mangement complexity and traffic overhead. In particular, this removes the need for internetworking between N3 and

N9 technologies, and offers a uniform user plane as recommended in the specification.

Potential gains can result for an early handling of traffic right from the RAN and thus possibly closer to the UE. The result is a simpler and lighter architecture, allowing convergence with other non-3GPP accesses.

7.1.1. hICN

This option ensures that forwarding beyond the radio access is directly managed through hICN. As a consequence, no additional state nor signaling is required for static and mobile consumers, nor for static producers. The impact of producer mobility is low because of the small number of impacted routers.

Dynamic forwarding capabilities are extended in this configuration to the selection of the first UPF, with the potential of additional performance improvement and higher traffic offload because of the deployment of hICN functionalities closer to the UE. A significant advantage arises in dense deployments scenarios where it becomes possible to isolate the core network from the locally-management mobility (a design objective of the mobile architecture), while allowing distributed selection of ingress UPFs, and dynamic per-packet load balancing of traffic across them.

7.1.2. Any technology specific comment here ?

7.2. Coexistence with GTP-based architecture

An alternative vision, although not recommended, would be to preserve the current architecture as is, and deploy alternative data planes on top.

7.2.1. Coexistence of Id/Loc Separation architecture with current 5GC

7.2.1.1. Overview of the Low Impact Approach

ID-Locator separation architecture can be implemented by control plane of a dedicated protocol such as LISP, ILNP, ILA, etc., however, it may cause major impact to the specifications of 3GPP 5GS. This approach, described in [I-D.homma-dmm-5gs-id-loc-coexistence], enables to introduce such ID-Locator separation protocols into 5GS with no or low impacts. It would also support a migration path toward a network which an ID-Locator separation protocol is completely incorporated.

This approach establishes an individual domain/slice in which an ID-Locator separation protocol works as packet forwarding mechanism, and divert appropriate packets to the domain by using Up-Link Classifier (ULCL) which is a fundamental function of UPF. An overview of this architecture is shown in Figure 13.

```
                         .--.
                        (    )-.
                      .'  cDN/  '
                     (  Internet )
                      (        -'
                       '-(     )
                          '---'
                            |N6            ,---------.
                      +-----+-----+        | Mapping |
                      |   cUPF    |        | System  |
                      +-----+-----+        `---------'
                            |N9                .
   ,------------------------+----------------.---------.
  / Transport Network   . . . . . . . . . . . . . . .  \
  |                     .                          .   |
  \                  #============================#===  /
   `----+------------#--.-----------+-------------#--.-'
        |N9          #  .           |N9           #  .
  +-----+-----+   +--------+   +-----+-----+   +--------+
  |   dUPF#1  |N6 | LOC-   |   |   dUPF#2  |N6 | LOC-   |
  |       [UL]+---+ Node#1 |   |       [UL]+---| Node#2 |..
  |       [CL]|   |        |   |       [CL]|   |        |
  +-----+-----+   +---+----+   +-----+-----+   +---+----+
        |N3           |              |N3           |
                   ,-----.                      ,-----.
     (( o ))      /       \     (( o ))        /       \
        A         | dDN#A |        A           | dDN#B |
       /-\  RAN   \       /       /-\  RAN     \       /
      /-|-\        `-----'       /-|-\          `-----'

        |                           |

     [ UE ] ..                   [ UE ] ..


                    dUPF/cUPF: Distributed/Central UPF
                     dDN/cDN : Distributed/Central DN
                       ===== : Connection between LOC nodes
                       . . . : IF to Mapping System
```

Figure 13: Overview of Proposed Architecture

LOC-node is a node which has a locator and forwards packets to appropriate destination based on looking up of destination ID. It is defined as xTR in LISP, and defined ILA-Node in ILA. Mapping System manages IDs of end points (e.g., UE, NF in dDN) and their bineded Locators which each ID is connected.

7.2.1.2. Data Plane

GTP-U or any forwarding protocol described in this document can be used as data plane mechanism of this approach. However, each LOC-Node must be connected to the all other LOC-Nodes and thus it may cause complexity of path management if you use a protocol which needs session establishment.

7.2.1.3. Control Plane

A control plane of every dedicated ID-Locator separation protocol described in this document can be used for this approach. For management of mobility of UEs in ID-Locator separation domain, some cooperation between SMF and mapping system is needed. In this approach, a UE is attached to a LOC-Node only when it communicates to another UE or an NF in a dDN. In 5GS, SMF manages sessions, and thus SMF may be required to update mapping database when an UE moves to under another UPF or an NF is moved to another dDN. The impact caused by such cooperation can be reduced by using Naf interface which is defined in 5GS specifications.

7.2.1.4. Features

This approach provides a mechanism for introducing ID-Locator separation architecture into 5GS with no or nominal impact, and achieves optimization of forwarding path and session continuity. Moreover, this can keep scalability on forwarding on down link from cDN/Internet because it can use the current GTP-based mechanism.

On the other hand, this approach causes an extra hop when diverting packets to ID-Locator separation domain, and it may leads to increase of latency.

7.2.2. Coexistence of hICN with current 5GC

This section discusses the insertion of hICN-AMM in an unmodified 3GPP 5G reference architecture, where GTP tunnels are preserved. As previously stated, maintaining GTP tunnels does not allow to overcome limitations of anchor-based approaches. However, a transparent integration of hICN-AMM limits to the minimum deployment costs and already brings advantages over the baseline architecture presented earlier.

7.2.2.1. Local breakout in Mobile Edge Computing (MEC)

The first option shares some similarities with the previous situation and proposes to deploy hICN-AMM within Mobile Edge Computing (MEC) platforms, as illustrated in Figure 14.

```
   +-------------------+          +-------------------+
   |        AMF        |          |        SMF        |
   +-+--------------+-+          +++--------------+-+
     |              |             .|              |
     | N1           | N2       N4 .| N4           | N4
     |              |             .|              |
 +---+---+      +---+---+  N3  +---+---+  N9  +---+---+  N6  +--------+
 |  UE   +-------+  gNB  +-------+ UL/CL +-------+  UPF  +-------+  DN    |
 +-------+      +-------+      +---+---+      +-------+      +-------+
                                  .|
                                  .| N9
                                  .|                            _
                               +--++---+  N9  +-------+  MEC  _( )___
                               |  UPF  +-------+   DN   +-------(  hICN )_
                               +-------+      +-------+      (_________)
                                  ^               ^              ^
                                  |_______________|              |
                                    local breakout        hICN insertion
```

Figure 14: hICN insertion in MEC

It relies on the local breakout capability introduced in 5G, i.e. a specific UPF denoted UL/CL (uplink classifier) that locally diverts specific flows to an alternative DN, filtering packets based for instance on information carried in packet headers. In the hICN-AMM case, this function is used to realize the hICN punting function described in [I-D.muscariello-intarea-hicn], i.e. to identify hICN traffic (Interest and Data packets) and forward it to the local MEC hICN instance.

Although it preserves tunnels and anchor points, this option permits an early termination of tunnels and the distribution of hICN capabilities close to the edge like in path caching and rate/loss/ congestion control which may be leveraged for efficient low-latency content distribution especially in presence of consumer mobility.

7.2.2.2. hICN as a UPF

The second option consists in the deployment of hICN-AMM as User Plane Function (UPF) inside mobile user plane. It has the advantage of preserving the advantage of hICN in terms of consumer mobility and flexible transport.

A more in depth presentation of those alternative deployments can be found in [I-D.auge-hicn-mobility-deployment-options].

7.3. Discussion

Motivate the introduction of optimized solutions with respect to :

- the gap between both "coexistence" and "integrated" deployment strategies;
- how well architectural requirements are addressed
- how well we address GTP shortcomings (provided they are detailed in the beginning of the document)

complex management
out of 3GPP control plane as we do an early termination
keep overhead and complexity of tunnel

8. No Protocol Option

In this option, mobility is handled nomadically by the app.

9. Comparison of Protocols

This section will compare the different protocols with reference to how they will support the requirements for UPF and N9 interface; how the various scenarios identified in Sections 3 and 4 will be supported and impacts to other interfaces and functions of the architecture (e.g. N3, N4, SMF, AMF, etc).

What is the criteria for comparison?

10. Summary

This document summarized the various IETF protocol options for GTP replacement on N9 interface of 3GPP 5G architecture.

11. Formal Syntax

The following syntax specification uses the augmented Backus-Naur Form (BNF) as described in [RFC2234].

12. Security Consideration

TBD

13. IANA Considerations

TBD

14. Acknowledgement

The authors would like to thank Farooq Bari, Devaki Chandramouli, Ravi Guntupalli, Sri Gundavelli, Peter Ashwood Smith, Satoru Matsushima, Michael Mayer, Vina Ermagan, Fabio Maino, Albert Cabellos, Cameron Byrne, Giovanna Carofiglio and Luca Muscariello for reviewing various iterations of the document and for providing content into various sections.

15. References

15.1. Normative References

[RFC1027] Carl-Mitchell, S. and J. Quarterman, "Using ARP to implement transparent subnet gateways", RFC 1027, DOI 10.17487/RFC1027, October 1987, <https://www.rfc-editor.org/info/rfc1027>.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

[RFC2234] Crocker, D., Ed. and P. Overell, "Augmented BNF for Syntax Specifications: ABNF", RFC 2234, DOI 10.17487/RFC2234, November 1997, <https://www.rfc-editor.org/info/rfc2234>.

[RFC4861] Narten, T., Nordmark, E., Simpson, W., and H. Soliman, "Neighbor Discovery for IP version 6 (IPv6)", RFC 4861, DOI 10.17487/RFC4861, September 2007, <https://www.rfc-editor.org/info/rfc4861>.

[RFC6830] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "The Locator/ID Separation Protocol (LISP)", RFC 6830, DOI 10.17487/RFC6830, January 2013, <https://www.rfc-editor.org/info/rfc6830>.

[RFC6831] Farinacci, D., Meyer, D., Zwiebel, J., and S. Venaas, "The Locator/ID Separation Protocol (LISP) for Multicast Environments", RFC 6831, DOI 10.17487/RFC6831, January 2013, <https://www.rfc-editor.org/info/rfc6831>.

[RFC6832] Lewis, D., Meyer, D., Farinacci, D., and V. Fuller, "Interworking between Locator/ID Separation Protocol (LISP) and Non-LISP Sites", RFC 6832, DOI 10.17487/RFC6832, January 2013, <https://www.rfc-editor.org/info/rfc6832>.

[RFC6833] Fuller, V. and D. Farinacci, "Locator/ID Separation Protocol (LISP) Map-Server Interface", RFC 6833, DOI 10.17487/RFC6833, January 2013, <https://www.rfc-editor.org/info/rfc6833>.

[RFC6835] Farinacci, D. and D. Meyer, "The Locator/ID Separation Protocol Internet Groper (LIG)", RFC 6835, DOI 10.17487/RFC6835, January 2013, <https://www.rfc-editor.org/info/rfc6835>.

[RFC6836] Fuller, V., Farinacci, D., Meyer, D., and D. Lewis, "Locator/ID Separation Protocol Alternative Logical Topology (LISP+ALT)", RFC 6836, DOI 10.17487/RFC6836, January 2013, <https://www.rfc-editor.org/info/rfc6836>.

[RFC7215] Jakab, L., Cabellos-Aparicio, A., Coras, F., Domingo-Pascual, J., and D. Lewis, "Locator/Identifier Separation Protocol (LISP) Network Element Deployment Considerations", RFC 7215, DOI 10.17487/RFC7215, April 2014, <https://www.rfc-editor.org/info/rfc7215>.

[RFC7476] Pentikousis, K., Ed., Ohlman, B., Corujo, D., Boggia, G., Tyson, G., Davies, E., Molinaro, A., and S. Eum, "Information-Centric Networking: Baseline Scenarios", RFC 7476, DOI 10.17487/RFC7476, March 2015, <https://www.rfc-editor.org/info/rfc7476>.

[RFC8060] Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", RFC 8060, DOI 10.17487/RFC8060, February 2017, <https://www.rfc-editor.org/info/rfc8060>.

[RFC8061] Farinacci, D. and B. Weis, "Locator/ID Separation Protocol (LISP) Data-Plane Confidentiality", RFC 8061, DOI 10.17487/RFC8061, February 2017, <https://www.rfc-editor.org/info/rfc8061>.

[RFC8111] Fuller, V., Lewis, D., Ermagan, V., Jain, A., and A. Smirnov, "Locator/ID Separation Protocol Delegated Database Tree (LISP-DDT)", RFC 8111, DOI 10.17487/RFC8111, May 2017, <https://www.rfc-editor.org/info/rfc8111>.

[RFC8112] Farinacci, D., Jain, A., Kouvelas, I., and D. Lewis, "Locator/ID Separation Protocol Delegated Database Tree (LISP-DDT) Referral Internet Groper (RIG)", RFC 8112, DOI 10.17487/RFC8112, May 2017, <https://www.rfc-editor.org/info/rfc8112>.

15.2. Informative References

[CICN]   io, Linux., "CICN project", 2018.

[I-D.auge-hicn-mobility]
Auge, J., Carofiglio, G., Muscariello, L., and M. Papalini, "Anchorless mobility through hICN principles", 2018.

[I-D.auge-hicn-mobility-deployment-options]
Auge, J., Carofiglio, G., Muscariello, L., and M. Papalini, "Anchorless mobility through hICN principles: Deployment options", 2018.

[I-D.farinacci-lisp-mobile-network]
Farinacci, D., Pillay-Esnault, P., and U. Chunduri, "LISP for the Mobile Network", draft-farinacci-lisp-mobile-network-03 (work in progress), March 2018.

[I-D.filsfils-spring-srv6-network-programming]
Filsfils, C., Li, Z., Leddy, J., daniel.voyer@bell.ca, d., daniel.bernier@bell.ca, d., Steinberg, D., Raszuk, R., Matsushima, S., Lebrun, D., Decraene, B., Peirens, B., Salsano, S., Naik, G., Elmalky, H., Jonnalagadda, P., and M. Sharif, "SRv6 Network Programming", draft-filsfils-spring-srv6-network-programming-04 (work in progress), March 2018.

[I-D.gundavelli-dmm-mfa]
Gundavelli, S., Liebsch, M., and S. Matsushima, "Mobility-aware Floating Anchor (MFA)", draft-gundavelli-dmm-mfa-00 (work in progress), February 2018.

[I-D.herbert-ila-ilamp]
Herbert, T., "Identifier Locator Addressing Mapping Protocol", draft-herbert-ila-ilamp-00 (work in progress), December 2017.

[I-D.herbert-ila-mobile]
Herbert, T. and K. Bogineni, "Identifier Locator Addressing for Mobile User-Plane", draft-herbert-ila-mobile-01 (work in progress), March 2018.

[I-D.herbert-intarea-ila]
Herbert, T. and P. Lapukhov, "Identifier-locator addressing for IPv6", draft-herbert-intarea-ila-01 (work in progress), March 2018.

[I-D.hmm-dmm-5g-uplane-analysis]
Homma, S., Miyasaka, T., and S. Matsushima, "User Plane Protocol and Architectural Analysis on 3GPP 5G System", 2018.

[I-D.homma-dmm-5gs-id-loc-coexistence]
Homma, S., Kawakami, K., and A. Akhavain, "Co-existence of 3GPP 5GS and Identifier Locator Separation Architecture", draft-homma-dmm-5gs-id-loc-coexistence-01 (work in progress), May 2018.

[I-D.ietf-6man-segment-routing-header]
Previdi, S., Filsfils, C., Leddy, J., Matsushima, S., and d. daniel.voyer@bell.ca, "IPv6 Segment Routing Header (SRH)", draft-ietf-6man-segment-routing-header-13 (work in progress), May 2018.

[I-D.ietf-dmm-srv6-mobile-uplane]
Matsushima, S., Filsfils, C., Kohno, M., Camarillo, P., daniel.voyer@bell.ca, d., and C. Perkins, "Segment Routing IPv6 for Mobile User Plane", draft-ietf-dmm-srv6-mobile-uplane-01 (work in progress), March 2018.

[I-D.ietf-lisp-eid-mobility]
Portoles-Comeras, M., Ashtaputre, V., Moreno, V., Maino, F., and D. Farinacci, "LISP L2/L3 EID Mobility Using a Unified Control Plane", draft-ietf-lisp-eid-mobility-02 (work in progress), May 2018.

[I-D.ietf-lisp-introduction]
Cabellos-Aparicio, A. and D. Saucez, "An Architectural Introduction to the Locator/ID Separation Protocol (LISP)", draft-ietf-lisp-introduction-13 (work in progress), April 2015.

[I-D.ietf-lisp-mn]
Farinacci, D., Lewis, D., Meyer, D., and C. White, "LISP Mobile Node", draft-ietf-lisp-mn-02 (work in progress), April 2018.

[I-D.ietf-lisp-pubsub]
Rodriguez-Natal, A., Ermagan, V., Leong, J., Maino, F., Cabellos-Aparicio, A., Barkai, S., Farinacci, D., Boucadair, M., Jacquenet, C., and S. Secci, "Publish/Subscribe Functionality for LISP", draft-ietf-lisp-pubsub-00 (work in progress), April 2018.

[I-D.ietf-lisp-rfc6830bis]
Farinacci, D., Fuller, V., Meyer, D., Lewis, D., and A. Cabellos-Aparicio, "The Locator/ID Separation Protocol (LISP)", draft-ietf-lisp-rfc6830bis-12 (work in progress), March 2018.

[I-D.ietf-lisp-rfc6833bis]
Fuller, V., Farinacci, D., and A. Cabellos-Aparicio, "Locator/ID Separation Protocol (LISP) Control-Plane", draft-ietf-lisp-rfc6833bis-10 (work in progress), March 2018.

[I-D.ietf-spring-segment-routing]
Filsfils, C., Previdi, S., Ginsberg, L., Decraene, B., Litkowski, S., and R. Shakir, "Segment Routing Architecture", draft-ietf-spring-segment-routing-15 (work in progress), January 2018.

[I-D.irtf-icnrg-ccnxsemantics]
Mosko, M., Solis, I., and C. Wood, "CCNx Semantics", draft-irtf-icnrg-ccnxsemantics-07 (work in progress), March 2018.

[I-D.irtf-icnrg-mapme]
Auge, J., Carofiglio, G., Muscariello, L., and M. Papalini, "MAP-Me : Managing Anchorless Mobility in Content Centric Networking", draft-irtf-icnrg-mapme-00 (work in progress), March 2018.

[I-D.irtf-icnrg-terminology]
Wissingh, B., Wood, C., Afanasyev, A., Zhang, L., Oran, D., and C. Tschudin, "Information-Centric Networking (ICN): CCN and NDN Terminology", draft-irtf-icnrg-terminology-00 (work in progress), December 2017.

[I-D.kouvelas-lisp-map-server-reliable-transport]
Cassar, C., Leong, J., Lewis, D., Kouvelas, I., and J. Arango, "LISP Map Server Reliable Transport", draft-kouvelas-lisp-map-server-reliable-transport-04 (work in progress), September 2017.

[I-D.lapukhov-bgp-ila-afi]
Lapukhov, P., "Use of BGP for dissemination of ILA mapping information", draft-lapukhov-bgp-ila-afi-02 (work in progress), October 2016.

[I-D.muscariello-intarea-hicn]
Muscariello, L., Carofiglio, G., Auge, J., and M. Papalini, "Hybrid Information-Centric Networking", draft-muscariello-intarea-hicn-00 (work in progress), June 2018.

[I-D.rodrigueznatal-ila-lisp]
Rodriguez-Natal, A., Ermagan, V., Maino, F., and A. Cabellos-Aparicio, "LISP control-plane for Identifier Locator Addressing (ILA)", draft-rodrigueznatal-ila-lisp-01 (work in progress), April 2018.

[I-D.vonhugo-5gangip-ip-issues]
Hugo, D. and B. Sarikaya, "Review on issues in discussion of next generation converged networks (5G) from an IP point of view", draft-vonhugo-5gangip-ip-issues-03 (work in progress), March 2017.

[I-D.xuclad-spring-sr-service-chaining]
Clad, F., Xu, X., Filsfils, C., daniel.bernier@bell.ca, d., Li, C., Decraene, B., Ma, S., Yadlapalli, C., Henderickx, W., and S. Salsano, "Segment Routing for Service Chaining", draft-xuclad-spring-sr-service-chaining-01 (work in progress), March 2018.

[ILAGRPS] "Identifier Group in ILA (To be published)", n.d..

[MAPME] Auge, J., Carofiglio, G., Grassi, G., Muscariello, L., Pau, G., and X. Zeng, "MAP-Me: Managing Anchor-Less Producer Mobility in Content-Centric Networks", IEEE Transactions on Network and Service Management Vol. 15, pp. 596-610, DOI 10.1109/tnsm.2018.2796720, June 2018.

[TR.29.891-3GPP]
3rd Generation Partnership Project (3GPP), "5G System ? Phase 1, CT WG4 Aspects, 3GPP TR 29.891 v15.0.0", December 2017.

[TS.23.203-3GPP]
3rd Generation Partnership Project (3GPP), "Policy and Charging Control Architecture, 3GPP TS 23.203 v2.0.1", December 2017.

[TS.23.501-3GPP]
3rd Generation Partnership Project (3GPP), "System Architecture for 5G System; Stage 2, 3GPP TS 23.501 v2.0.1", December 2017.

[TS.23.502-3GPP]
3rd Generation Partnership Project (3GPP), "Procedures for 5G System; Stage 2, 3GPP TS 23.502, v2.0.0", December 2017.

[TS.23.503-3GPP]
3rd Generation Partnership Project (3GPP), "Policy and Charging Control System for 5G Framework; Stage 2, 3GPP TS 23.503 v1.0.0", December 2017.

[TS.29.244-3GPP]
3rd Generation Partnership Project (3GPP), "Interface between the Control Plane and the User Plane Nodes; Stage 3, 3GPP TS 29.244 v15.0.0", December 2017.

[TS.29.281-3GPP]
3rd Generation Partnership Project (3GPP), "GPRS Tunneling Protocol User Plane (GTPv1-U), 3GPP TS 29.281 v15.1.0", December 2017.

[TS.38.300-3GPP]
3rd Generation Partnership Project (3GPP), "NR and NG-RAN Overall Description: Stage 2, 3GPP TS 38.300 v2.0.0", December 2017.

[TS.38.401-3GPP]
3rd Generation Partnership Project (3GPP), "NG-RAN: Architecture Description, 3GPP TS 38.401 v1.0.0", December 2017.

[WLDR] Carofiglio, G., Muscariello, L., Papalini, M., Rozhnova, N., and X. Zeng, "Leveraging ICN In-network Control for Loss Detection and Recovery in Wireless Mobile networks", Proceedings of the 2016 conference on 3rd ACM Conference on Information-Centric Networking - ACM-ICN ’16, DOI 10.1145/2984356.2984361, 2016.

Appendix A. SRv6 Based Solution

A.1. Overview

Segment Routing (SR), defined in [I-D.ietf-spring-segment-routing] generalises the source routing paradigm with an ordered list of global and/or nodal instructions (segments) prepended in an SR header in order to either steer traffic flows through the network while confining flow states to the ingress nodes in the SR domains and/or to indicate functions that are performed at specific network locations.

The IPV6 realisation of SR (SRV6) defines a SR Header (SRH), see [I-D.ietf-6man-segment-routing-header]. SRV6 encodes segments as IPV6 addresses in the Segment List (SL) of its header. The packet destination address in SRV6 specifies the active segment while an index field in the SRH points to the next active segment in the list. The index field in SRH is decremented as SRV6 progressively forces packet flows through different segments over the IPV6 data plane.

The versatility and adaptability of SR combined with IPV6's ample and flexible address space positions SRV6 as a viable data path technology for the next generation of mobile user plane, in particular the 3GPP N9 (UPF to UPF).

This section starts by summarising the use of SRV6 as a drop-in alternative for GTP-U over the N9 interface connecting different User Plane Functions (UPF). It then shows how SRV6 as a GTP-U replacement can then provide additional features such as TE, sparing, rate limiting, and service chaining that are not natively available by GTP.

The discussion then focuses on advanced routing with the Identifier/ Locator paradigm and shows how SRV6 can be used to realise this model in the mobile back-haul in either an anchored or anchorless mode of operation.

SRV6 appears well placed as a mechanism to replace GTP-U with initially no control plane changes, but to then offer a progressive path towards many innovations in routing.

A.2. SRV6 as Drop-In Alternative for GTP-U

Existing mobile back-haul employs GTP tunnels to carry user traffic flows in the network. These tunnels are unidirectional, are established via the control plane for a particular QoS level, and run on links between access and the different anchor nodes all the way to DN gateways. 3GPP uses the term UPF to refer to the variety of functions performing different tasks on user traffic along the data path in 5G networks and suggests the use of GTP tunnels to carry user traffic between these UPFs (N9 interface).

The Tunnel Id (TEID) field in the GTP tunnel plays a crucial role in stitching the data path between the above mentioned network nodes for a particular user flow. In other words, TEIDs are used to coordinate traffic hand off between different UPFs.

In its most basic form, SRV6 can be used as a simple drop-in alternative for GTP tunnels. The control plane in this approach remains the same, and still attempts to establish GTP-U tunnels and communicate TEIDs between the tunnel end points. However, at the user plane, SRV6 capable nodes use SIDs to direct user traffic between the UPFs.

The simplest option is to encapsulate the entire GTP frame as a payload within SRV6. However, this scheme still carries the GTP header as the payload and as such doesn’t offer significant advantage.

A much more promising option however is to use SIDs to carry tunnel related information. Here, TEIDs and other relevant data can be encoded into SRV6 SIDs which can be mapped back to TEID’s at the intermediate UPFs thus requiring no changes except at the encapsulation and de-encapsulation points in the UPF chains.

[I-D.ietf-dmm-srv6-mobile-uplane] discusses the details of leveraging the existing control plane for distributing GTP tunnel information between the end nodes and employing SRV6 in data plane for UPF connectivity. The document defines a SID structure for conveying TEID, DA, and SA of GTP tunnels, shows how hybrid IPV4/IPV6 networks are supported by this model and in doing so, it paves a migration path toward a full SRV6 data plane.

Another alternative that can provide for a smooth migration toward SRV6 data plane between UPFs is via the use of "Tag", and optional TLV fields in SRH. Similar to the previously described method, this approach takes advantage of the existing control plane to deliver GTP tunnel information to the UPF endpoints. "Tag" and optional TLV fields in SRH are then used to encode tunnel information in the SRV6 data plane where the UPFs can determine the TEID etc. by inverting the mapping.

In yet another option, GTP tunnel information can be encoded as a separate SID either within the same SRH after the SID that identifies the UPF itself (SRH-UPF)or inside a separate SRH (SRH-G). In this option, SID representing the GTP tunnel information acts as both start and end point of a segment within the UPF. This option resembles the MPLS label stacking mechanism which is widely used in different VPN scenarios.

It must be noted that in any of the above mentioned approaches, the ingress UPF in SRV6 domain can insert a SRH containing the list of SIDs that corresponds to all UPFs along the path. Alternatively, UPFs can stack a new SRH on top of the one inserted by the previous one as packets traverse network paths between different pairs of UPFs in the network.

```
                           +-------+
                    +------+  SMF  +------+
                    |      +-------+      |
                   N4                    N4
                    |                     |
+--------+      +---+---+             +---+---+      +--------+
|  RAN   |--N3--|  UPF  |             |  UPF  |--N6--|  DN    |
+--------+      +---+---+             +---+---+      +--------+
                    |        SRV6 N9      |
                    |   carries GTP info. |
                    +---------------------+
```

Figure 15: SRV6 as Drop-In replacement for GTP-U in 5G

A.3. SRV6 as Drop-In GTP Replacement with TE

The previous section discussed using SRV6 as a drop-in replacement for GTP tunnels in existing mobile networks. No new capabilities were introduced by this simple 1 to 1 replacement. We now explore additional possible features once SRV6 has been introduced.

Traffic engineering is an integral feature of SR. The SRV6 variant of SR of course supports both strict and loose models of source routing. Here, the SID list in SRH can represent a loose or strict path to UPFs. Therefore, traffic engineering can easily be supported regardless of any of the aforementioned approaches.

For loose paths to UPFs, a set of one or more SIDs in SRH's SID list identifies on or more, but not all the intermediate nodes to a particular UPF. Packets then follow the IGP shortest path through the network to each specified intermediate node till they reach the target UPF.

In the case of strict path to UPFs, SRH contains a set of SIDs representing all the intermediate nodes and links that the packet must visit on its route to a particular UPF. The last SID in the set represents the target UPF itself or the last link to this UPF. Here, SRV6 packet processing at each node invokes the function(s) that is associated with SID[SL], the packet then receives the required treatment and gets forwarded over the SRH's specified path toward the target UPF.

It must be noted that the SRH could contain multiple sets of SIDs each representing a TE path between a pair of UPFs. Alternatively, the SRH can contain a fully resolved end to end TE path that covers every intermediate node and UPF along the data plane.

SR considers segments to be instructions. Therefore each SID can represent a function that enforces a specific nodal or global packet treatment. Attributes such as jitter and delay requirement, rate limiting factors, etc. can be easily encoded in to SIDs in order to apply the desired treatment as packets traverse the network from UPF to UPF. [I-D.ietf-dmm-srv6-mobile-uplane] suggests a SID encoding mechanism for rate limiting purposes.

Please refer to the followings for further details about SR and SRV6 traffic engineering capabilities, network programming concept, and a list of some of the main SR functions.

- [I-D.ietf-spring-segment-routing]
- [I-D.ietf-6man-segment-routing-header]
- [I-D.filsfils-spring-srv6-network-programming]
- [I-D.gundavelli-dmm-mfa]

A.4. UPF Chaining with SRV6

Service or function chaining is another intrinsic feature of SR and its SRV6 derivative. Using this capability, operators can direct user traffic through a set of UPFs where each UPF performs a specific task or executes certain functions on the traffic.

UPF chaining is achieved through the use of SIDs in SRV6 in the manner identical to what was described in the previous section regarding SRV6 support for traffic engineering.

Generally speaking, the SRH is populated with a set of SIDs with each SID identifying a specific UPF in the network. Starting from the ingress SRV6 node, packets are then forwarded through the network in either loose or explicit mode toward each UPF.

Please refer to [I-D.xuclad-spring-sr-service-chaining] for further detail.

A.5. SRV6 and Entropy

Ability to provide a good level of entropy is an important aspect of data plane protocols. The TEID field in GTP tunnels if included in network node's hashing algorithms can result in good load balancing. Therefore, any new data plane proposal should be able to deal with entropy in an efficient manner.

SRV6 SIDs can easily accommodate entropy at either hop by hop or global level via reserving a set of bits in the SID construct itself; and hence, eliminate the need for a special entropy Segment ID in SRH. Here, the hashing algorithm at different nodes distribute traffic flows based on the SID which has been copied to IPV6 DA field.

Alternatively, entropy related information can be encoded as optional TLV field in SRV6's SRH.

A.6. SRV6 and 5G Slicing

Slicing is one of the main features in 5G [3GPP 23501]. Several Slices with different requirements can coexist on top of the common network infrastructure. Diverse flows belonging to different 5G slices can be completely disjoint or can share different parts of the network infrastructure. SRV6's native features such as TE, Chaining, one-plus-one protection, etc. either in stand-alone or in conjunction with other alternatives for mobility support such as ID-LOC model lend themselves well to 5G slicing paradigm.

```
                             +--+       +--+
              +--------------+N1| . . . |Nn+----------------+
              |              ++-+       +-++                |
              |               |           |                 |
          +---+---+           |           |             +---+---+
          | UPF-1 |    +-------+          +-------+     | UPF-2 |
          +---+---+    |                          |     +---+---+
                       |                          |
   Slice-A             |                          |
   --------------------|--------------------------|--------------------
   Slice-B             |                          |
              +--------+                          +---------+
              |                                             |
          +---+---+                                     +---+---+
          | UPF-1'|                                     | UPF-2'|
          +---+---+                                     +---+---+
              |              +--+       +--+                |
              +--------------+M1| . . . |Mn+----------------+
                             +--+       +--+
```

Figure 16: SRV6 TE, Service Chaining, Sparing, and Protection for 5G Slices

A.7. SRV6 and Lawful Interception for 5G

TBD

A.8. SRV6 and Alternative Approaches to Advanced Mobility Support

SRV6 flexibility enables it to support different methods of providing mobility in the network. ID-LOC for mobility support is one such option.

A.8.1. SRV6 and Locator/ID Separation Paradigm for N9 Interface

The previous sections discussed how SRV6 could be employed as a replacement for GTP tunnels while leaving the existing control plane intact. This section describes the use of SRV6 as a vehicle to implement Locator/ID Separation model for UPF data plane connectivity.

A.8.2. Brief Overview of Locator-ID Separation

Traditional routing architecture uses IP addresses as both device identity and its location in the network. Locator-ID Separation model establishes a paradigm in which a device identity and its network location are split into two separate namespaces: End-point Identifiers (EID), and Route Locators (RLOC) that are correlated via a control plane, or a dynamic (centralised or distributed) mapping system.

RLOCs are tied to network topology. They represent network devices that are reachable via traditional routing. EIDs, on the other hand, represent mobile or stationary devices, functions, etc. that are reachable via different RLOCs based on the network location where they get instantiated, activated or moved.

Using this model, as long as EID-RLOC relationship remains up to date, EIDs can easily move between the RLOCs. That is the EID namespace can freely move without any impact to the routing paths and connectivity between the Route Locators.

This type of multi encapsulation and routing has been employed in fixed networks (IP, VPN, MPLS, etc.). The use of this paradigm in mobile data plane, therefore, offers an approach that takes advantage of a mature and proven technology to implement the N9 interface for UPF connectivity.

A.8.3. Locator-ID Separation via SRV6 for UPF connectivity

SRV6 can easily implement ID-LOC Separation model for UPF connectivity. The SIDs are once again the main vehicle here. In this model, UPFs are considered to be the IDs while the nodes where the UPFs attach to take on the role of the Locators. Multiple UPFs are allowed to attach to the same Locator. It is also possible for a UPF to connect to multiple Locators. There are several implementation options. The followings highlights a few possibilities.

A.8.3.1. Overlay model with SRV6 Locators

In this approach, UPFs connect to SRV6 capable Locators. UPFs use IPV4/IPV6 transport either in conjunction with GTP or without any GTP tunnel and send the packets to their associated Locator at the near end (Ingress SRV6 Locator).

In either case, the ingress SRV6 Locator uses the DA field in arriving packets to identify the far end Locator (Egress SRV6 Locator) where the target UPF is attached and obtains its associated SID.

For GTP encapsulated traffic from UPFs, the ingress SRV6 Locator must also deliver GTP information to the far end Locator. Please see section 5.2. for more information on different methods of conveying GTP information in SRV6 domains.

The ingress SRV6 Locator then constructs the SRH and sends the traffic through the SRV6 network toward the egress RV6 Locator. Egress Locator marks the end of the segment and ships the traffic to the target UPF.

It must be noted that use of GTP at UPFs allows us to leave the 3GPP control plane intact and hence provides a smooth migration path toward SRV6 with ID-Locator model. For inter UPF traffic that doesn't use GTP, the control plane requires some modifications in order to be able to convey endpoint information to interested parties.

```
                                 +----+----+
         +-------------N4--------+   SMF   +--------N4-----------+
         |                       +----+----+                      |
         |                                                        |
         |                                                        |
         |                       +----+----+                      |
         |                       |  ID-Loc |                      |
         |               +----| Mapping |<----+              |
         |               |       +----+----+     |                |
         |               V                       V                |
     +---+---+      +----+----+            +----+----+      +---+---+
--N3-+ UPF-A +------+  RLOC-A +<----SRV6-+  RLOC-B +------+ UPF-B +-N6--
     +---+---+      +----+----+            +----+----+      +---+---+
         <------------------------N9-GTP-----------------------
```

Figure 17: Overlay Model with SRV6 Locator in 5G

A.8.3.2. SRV6 Capable UPFs and RLOCs

In this model, the head end UPF (Ingress UPF) is the ingress node and the entity that constructs the SRH in the SRV6 domain. Here, both UPFs (IDs) and Locators are represented by SIDs in the SRH. The SID list establishes either a partial or the full path to a target or a set of UPFs that traffic is required to traverse.

The 3GPP control plane is responsible for distributing UPF's endpoint information. But it requires some modifications to be able to convey endpoint information to interested parties.

In its simplest form, the SMF using policy information prepares a set of one or more UPFs along the traffic path and distributes this set in the form a SID list to the ingress UPF. This SID list of UPFs is then gets augmented with a set of SIDs identifying the Locators representing the current point of attachment for each UPF along the data path.

Alternatively, the SMF can provide a fully resolved SID list by communicating with a centralised or distributed ID-LOC mapping system containing all the relevant data regarding the UPF-Locator relationship.

In yet another approach, the SMF can provide a partial SID list representing the segment between each pair of UPFs to individual UPFs along the path.

Regardless of the approach, any changes to UPF's point of attachment must be reflected in the mapping system and communicated to the SMF for distribution to the appropriate set of UPFs. Keeping the mapping system current is essential to proper operation. As long as the mapping database is up-to-date, UPFs can be easily moved in the network. Design of ID-Locator mapping system is beyond the scope of this document. However, experiment with distributed mapping systems offered by today's public clouds has shown very promising results which can be further improved and tailored to mobile network requirements.

The following figure shows the use of SRV6 UPFs and RLOCs in 5G.

```
                                                          +----+----+
                                                          |  ID-Loc |
                                                   +---+ Mapping |
                                 +----+----+       |      |         |
                                 +         +<------+      +----+----+
                                 |   SMF   |
         +-------------N4--------+         +---------N4-----------+
         |                       +----+----+                      |
  SID-UA |            SID-RA                   SID-RB             | SID-UB
     +---+---+      +----+----+             +----+----+      +---+---+
--N3-+ UPF-A +-------+  RLOC-A +-------------+  RLOC-B +-------+ UPF-B +-N6--
     +---+---+      +----+----+             +----+----+      +---+---+
         <------------------------N9-SRV6----------------------
```

Figure 18: SRV6 Capable UPFs and Locators in 5G

A.8.4. Advanced Features in ID-Locator Architecture

SRV6's native features such as Traffic Engineering, QoS support, UPF Chaining, etc. can be easily added to ID-Locator support. As it was noted earlier, these features are not readily available by GTP.

A.9. Areas of Concerns

Support for IPV6 is a precondition for SRV6. Although SRV6 can support hybrid IPV4/IPV6 mobile data plane through an interworking node, support of UPFs with IPV4 address is rather complex.

Due to IPV6 128-bit address space, large SRH size can have a negative impact on MTU. Large SRH size can also exert undesirable header tax especially in the case of small payload size. Furthermore, compound SID processing at each node might affect line rate.

ID-LOC architecture relies on high performance mapping systems. Distributed mapping systems using some form Distributed Hash Table(DHT) exhibit very promising results. But further investigation is required to ensure mobility requirements in mobile data plane.

Appendix B. LISP based Solution

```
         +------------------------------------------------------+
         |                         SMF                          |
         +-+-----------+- - - - - - - - - - - - - +---+-------+-+
           |           |      Mapping System      |   |       |
           |           +--+----+---------------+--+   |       |
           N4             |    |               |      N4      N4
           |              |  LISP-CP           |      |       |
           |              |    |               |      |       |
        +--+---+          |    | +------+      |      |   +---+--+
        |      |          |    | |      +-------------+   |      |
   --N3-+ UPF  +          |    | + UPF  +      |          + UPF  +-N6--
        |      +--LISP-CP-+    +-+      |      +--LISP-CP-+      |
        +--+-+-+                 +-+--+-+                 +-+-+--+
           | |                     |  |                     | |
           | +-------LISP-DP-------+  +--------LISP-DP------+ |
           |                                                  |
           +----------------------LISP-DP---------------------+
```

Figure 19: LISP in the 5G architecture

B.1. Overview

The Locator/Identifier Separation Protocol (LISP), which provides a set of functions for routers to exchange information used to map from Endpoint Identifiers (EIDs) that are not globally routable to routable Routing Locators (RLOCs). It also defines a mechanism for these LISP routers to encapsulate IP packets addressed with EIDs for transmission across a network infrastructure that uses RLOCs for routing and forwarding.

An introduction to LISP can be found in [I-D.ietf-lisp-introduction].

A complete RFC-set of specifications can be found in [RFC6830], [RFC6831], [RFC6832], [RFC6833], [RFC6836], [RFC7215], [RFC8061], [RFC8111]. They describe support and mechanisms for all combinations of inner and outer IPv4 and IPv6 packet headers for unicast and multicast packet flows that also interwork with non-LISP sites as well as two designs to realize a scalable mapping system.

A standards-track based set of drafts [I-D.ietf-lisp-rfc6830bis] [I-D.ietf-lisp-rfc6833bis] are products and work in progress of the LISP Working Group.

B.2. LISP Data-Plane

LISP uses dynamic tunnel encapsulation as its fundadmental mechanism for the data-plane. Fixed headers are used between the outer and inner IP headers which are 16 bytes in length. Details can be found in [RFC6830].

B.3. LISP Control-Plane

Many years of research dating back to 2007 have gone into LISP scalable mapping systems. They can be found at [LISP-WG] and [IRTF-RRG]. The two that show promise and have deployment experience are LISP-DDT [RFC8111] and LISP-ALT [RFC6836].

The control-plane API which LISP xTRs are the clients of is documented in [RFC6833]. Various mapping system and control-plane tools are available [RFC6835] [RFC8112] and are in operational use.

B.4. LISP Mobility Features

LISP supports multi-homed shortest-path session survivable mobility. An EID can remain fixed for a node that roams while its dynamic binding changes to the RLOCs it uses when it reconnect to the new network location.

When the roaming node supports LISP, its EIDs and RLOCs are local to the node. This form of mobility is call LISP Mobile-Node. Details can be found in [I-D.ietf-lisp-mn].

When the roaming node does not support LISP, but LISP runs in the network the node roams to, the EIDs and RLOCs are not co-located in the same device. In this case, EIDs are assigned to the roaming node and RLOCs are assigned to LISP xTRs. So when the roaming node attaches to the network, its EIDs are mapped to the RLOCs of the LISP xTRs in the network. This form of mobility is called LISP EID-Mobility. Details can be found in [I-D.ietf-lisp-eid-mobility].

For a 3GGP mobile network, the LISP EID-Mobility form of mobility is recommended and is specified in the use-case document [I-D.farinacci-lisp-mobile-network].

B.5. ILSR

ILSR is a specific recommendation for using LISP in the 3GPP 5G mobile network architecture. A detailed whitepaper can be found at [ILSR-WP]. The recommendation is to use the mechanisms in [I-D.farinacci-lisp-mobile-network].

B.6. LISP Control-Plane with ILA Data-Plane

The current LISP control-plane (LISP-CP) specification [I-D.ietf-lisp-rfc6833bis] is data-plane agnostic and can serve as control-plane for different data-plane protocols. In this section we describe how LISP-CP can serve to enable the operation of an ILA data-plane. A similar approach can be followed to use LISP-CP as control-plane for other data-plane protocols (e.g. VXLAN, SRv6, etc).

```
         +------------------------------------------------------+
         |                         SMF                          |
         +-+-----------+- - - - - - - - - - - - - +---+-------+-+
           |           |      Mapping System      |   |       |
           |           +--+----+---------------+--+   |       |
           N4             |    |               |      N4      N4
           |              |  LISP-CP           |      |       |
           |              |    |               |      |       |
        +--+---+          |    | +------+      |      |   +---+--+
        |      |          |    | |      +-------------+   |      |
   --N3-+ UPF  +          |    | + UPF  +      |          + UPF  +-N6--
        |      +--LISP-CP-+    +-+      |      +--LISP-CP-+      |
        +--+-+-+                 +-+--+-+                 +-+-+--+
           | |                     |  |                     | |
           | +---------ILA---------+  +----------ILA--------+ |
           |                                                  |
           +------------------------ILA-----------------------+
```

Figure 20: LISP-CP + ILA in the 5G architecture

Please refer to Section 8 for description of the ILA data-plane. The complete specification of how to use the LISP-CP in conjunction with an ILA data-plane can be found in [I-D.rodrigueznatal-ila-lisp].

Below are summarized the major points to take into account when running LISP-CP as control-plane for ILA.

- o Leveraging on the flexible LISP-CP address encoding defined in [RFC8060], different ILA address types are defined in [I-D.rodrigueznatal-ila-lisp] to carry ILA metadata over the LISP-CP.

- o XTRs can serve as both ILA-Ns (when their map-cache is incomplete) or ILA-Rs (when their map-cache is complete). XTRs serving as ILA-Rs subscribe to the Mapping System to populate their map-cache with all the mappings in the domain (or its shard) using [I-D.ietf-lisp-pubsub].

- o LISP-CP can run over TCP or UDP. The same signaling and logic applies independently of the transport. Additionally, when running over TCP, the optimizations specified in [I-D.kouvelas-lisp-map-server-reliable-transport] can be applied.

- o The ILA control-plane operations "request/response" and "push" are implemented via the LISP mechanisms defined in [I-D.ietf-lisp-rfc6833bis] and [I-D.ietf-lisp-pubsub] respectively. When the Mapping System is co-located with the XTRs serving as ILA-Rs, the ILA "redirect" operation is implemented via the mapping notifications described in [I-D.ietf-lisp-pubsub].

- o XTRs serving as ILA-Ns can use LISP-CP as described in [I-D.ietf-lisp-rfc6833bis] to register and keep updated in the Mapping System the information regarding their local mappings.

- o When using ILA as data-plane, the mobility features and benefits discussed in Section 8 and in [I-D.ietf-lisp-eid-mobility] still apply.

- o As discussed in [I-D.rodrigueznatal-ila-lisp], the LISP-CP can be used not only to resolve ID-Loc mappings but also to obtain the ILA Identifier when it is not possible to locally derivate it from the endpoint address. These two mapping operations can be combined into one to obtain the ILA Identifier and associated locators in a single round of signaling.

B.7. LISP Control-Plane with SRv6 Data-Plane

```
         +------------------------------------------------------+
         |                         SMF                          |
         +-+-----------+- - - - - - - - - - - - - +---+-------+-+
           |           |      Mapping System      |   |       |
           |           +--+----+---------------+--+   |       |
           N4             |    |               |      N4      N4
           |              |  LISP-CP           |      |       |
           |              |    |               |      |       |
        +--+---+          |    | +------+      |      |   +---+--+
        |      |          |    | |      +--------------+  |      |
   --N3-+ UPF  +          |    | + UPF  +      |          + UPF  +-N6--
        |      +--LISP-CP-+    +-+      |      +--LISP-CP-+      |
        +--+-+-+                 +-+--+-+                 +-+-+--+
           | |                     |  |                     | |
           | +---------SRv6--------+  +---------SRv6--------+ |
           |                                                  |
           +-----------------------SRv6-----------------------+
```

Figure 21: LISP-CP + SRv6 in the 5G architecture

Appendix C. ILNP based Solution

TBD

Appendix D. ILA based Solution

Identifier-Locator Addressing [I-D.herbert-intarea-ila] is a protocol to implement transparent network overlays without encapsulation. It addresses the need for network overlays in virtualization and mobility that are efficient, lightweight, performant, scalable, secure, provide seamless mobility, leverage and encourage use of IPv6, provide strong privacy, are interoperable with existing infrastructure, applicable to a variety of use cases, and have simplified control and management.

D.1. Overview of ILA

ILA is a form of identifier/locator split where IPv6 addresses are transformed from application-visible, non-topological "identifier" addresses to topological "locator" addresses. Locator addresses allow packets to be forwarded to the network location where a logical or mobile node currently resides or is attached. Before delivery to the ultimate destination, locator addresses are reverse transformed back to the original application visible addresses. ILA does address "transformation" as opposed to "translation" since address modifications are always undone. ILA is conceptually similar to ILNP and 8+8, however ILA is contained in the network layer. It is not limited to end node deployment, does not require any changes to transport layer protocols, and does not use extension headers.

ILA includes both a data plane and control plane. The data plane defines the address structure and mechanisms for transforming application visible identifier addresses to locator addresses. The control plane's primary focus is a mapping system that includes a database of identifier to locator mappings. This mapping database drives ILA transformations. Control plane protocols disseminate identifier to locator mappings amongst ILA nodes.

The use cases of ILA include mobile networks, datacenter virtualization, and network virtualization. A recent trend in the industry is to build converged networks containing all three of these to provide low latency and high availability. A single network overlay solution that works across multiple use cases is appealing.

Benefits of ILA include:

o Facilitates node mobility and virtualization

o Multiple use cases (mobile, datacenter, cloud)

o Super efficient and performant data plane

o Allows strong privacy in addressing [ADDRPRIV]

o Promotes anchorless mobility

o No typical tunneling issues (e.g. MTU) or management related to encapsulation

o Flexible control plane that splits data and control

o Modern "SDN" control protocols (e.g. RPC/TCP)

o Scale number of nodes to billions for 5G, DC virtualization

o Upstream Linux kernel data path [ILAKERNEL] and open source ctrl plane [ILACONTROL].

The ILA data plane protocol is described in [I-D.herbert-intarea-ila], motivation and problems areas are described in [ILAMOTIVE], ILA in the mobile user-plane is described in detail in [I-D.herbert-ila-mobile].

D.2. ILA in the 5G Architecture

ILA is a proposed alternative to GTP-U and encapsulation. It does not require anchors and simplifies both the data plane and control plane. ILA is a general network overlay protocol can be used to meet the requirements of use cases in a converged network. User Plane Functions (UPF) with ILA are lightweight and stateless such that they can be brought up quickly as needed.

Figure 22, Figure 23 depict two architectural options for the use of ILA in a 5G architecture. ILA is logically a network function and ILA interfaces to the 5G control plane via service based interfaces.

In this architecture, ILA replaces GTP use over the N9 interface. Identifier address to locator address transformations in the downlink from the data network are done by an ILA-R. Transformations for intra domain traffic can be done by an ILA-N close to the gNB or by an ILA-R in the case of a cache miss. Locator address to identifier address transformation happen at ILA-Ns.

ILA could be supported on a gNB. In this case, an ILA-N would be co-resident at a gNB and ILA is used over N3 interface in lieu GTP-U.

Figure 23 and Figure 24 depict two options of how ILA can be used in the 5G architecture. The control plane functions can be implemented as standalone network functions or can be implemented with other network functions. The control plane protocol can be implemented as enhancement to N4, as APIs or as independent protocol.

Use of ILA in roaming scenarios is still TBD.

```
                    Service Based Interfaces
   ----+-----+-----+----+----+----+----+--------+-----+--------
       |     |     |    |    |    |    |        |     |
   +---+---+ |  +--+--+ |  +--+--+ |  +--+--+  +--+--+ |
   | NSSF  | |  | NRF | |  | DSF | |  | UDM |  | NEF | |
   +-------+ |  +-----+ |  +-----+ |  +-----+  +-----+ |
             |          |          |                   |
         +---+----+  +--+--+  +---+---+  +-------------+--+
         |  AMF   |  | PCF |  | AUSF |  |      ILA-M      |
         +---+--+-+  +-----+  +-------+  +-+------------+--+
  +--------+ |   |                          |            |
  | 5G UE |--+   |                          |            |
  +---+---+      |  N2              +-----+-----+  +---+---+      +-----+
      |          |          +------------| ILA-N  |--| ILA-R |------| DN |
      |          |          |   N3      +-+---+--+-+  +-+------+      +-----+
      |          |          |              |   |  |      |
      |      +---+------+-+                +---+  +-------+
      +------|   gNB      |                 N9       N9
             +------------+

          Figure 22: ILA in 5G architecture - Option 1
```

```
                    Service Based Interfaces
   ----+-----+-----+----+----+----+------+----+----+----+--------+--
       |     |     |    |    |    |      |    |    |    |        |
   +---+---+ |  +--+--+ |  +--+--+ |      |    |    | +--+--+  +--+--+
   | NSSF  | |  | NRF | |  | DSF | |      |    |    | | UDM |  | NEF |
   +-------+ |  +-----+ |  +-----+ |      |    |    | +-----+  +-----+
             |          |          |      |    |    |
         +---+----+  +--+--+  +---+---+  |  +--+--+ |
         |  AMF   |  | PCF |  | AUSF |  |  |ILA-M| |
         +---+--+-+  +-----+  +-------+  |  +--+--+ |
  +--------+ |   |                        |          |
  | 5G UE |--+   |                        |          |
  +---+---+      |  N2              +-----+--+  +---+---+      +-----+
      |          |          +------------| ILA-N |--| ILA-R |------| DN |
      |          |          |   N3      ++--+-+-+  +-+------+      +-----+
      |          |          |              |  | |       |
      |      +---+------+-+               +--+ +-------+
      +------|   gNB      |                N9     N9
             +------------+

          Figure 23: ILA in 5G architecture - Option 2
```

```
                     Service Based Interfaces
    ---------+-------+------------------+-----------------------
             |       |                  |
          +--+--+ +--+--+           +---+---+
          | AMF | | SMF |           | ILA-M |
          +--+-++ +--+--+           +---+---+
             N1|                        |
   +-------+ | |          +-------------+----+
   | 5G UE |-+ |          |                  |
   +---+---+   N2     +---+---+    +----+    +-----------+
       |       |   +--| ILAN/R|----| DN |    |           |
       |       |   |  +-------+    +----+    |           |
       |     +-+---+-+                    +--+--+     +--+--+  +----+
       +-----|  gNB  |--------------------|ILAN |--N9-|ILAR |--| DN |
             +-------+                    +-----+     +-----+  +----+
```

Figure 24: ILA in 5G architecture - Option 3

D.3. Protocol Layering

Figure 25 illustrates the protocol layers of packets packets sent over various data plane interfaces in the downlink direction of data network to a mobile node. Note that this assumes the topology shown in Figure 2 where GTP-U is used over N3 and ILA is used on N9.

```
                    -             -            -
       DN to ILA-R      ILA-R to ILA-N   ILA-N to gNB     gNB to UE
      +-------------+  +-------------+  +-------------+  +-------------+
      | Application |  | Application |  | Application |  | Application |
      +-------------+  +-------------+  +-------------+  +-------------+
      |     L4      |  |     L4      |  |     L4      |  |     L4      |
      +-------------+  +-------------+  +-------------+  +-------------+
      |    IPv6     |  | IPv6 (ILA)  |  |    IPv6     |  |  PDU Layer  |
      +-------------+| +-------------+| +-------------+  +-------------+
      |     L2      || |     L2      || |    GTP-U    |  | AN Protocol |
      +-------------+| +-------------+| +-------------+  |   Layers    |
                     |                | |   UDP/IP    |  |             |
                    N6   <--N9    N3 +-------------+  +-------------+
                                        |     L2      |
                                        +-------------+
```

Figure 25: ILA and protocol layer in 5G

D.4. Control Plane

ILA-M provides the interface between the 5G services architecture and the common ILA control plane.

D.4.1. ILA-M services interface

The control interface into ILA is via an ILA-M that interacts with 5G network services. ILA-M uses RESTful APIs to make requests to network services. An ILA-M receives notifications when devices enter the network, leave it, or move within the network. The ILA-M writes the ILA mapping entries accordingly.

ILA is a consumer of several 5G network services. The service operations of interest to ILA are:

- o Nudm (Unified Data Management): Provides subscriber information.

- o Nsmf (Service Managment Function): Provides information about PDU sessions.

- o Namf (Core Access and Mobility Function): Provides notifications of mobility events.

D.4.2. ILA control plane

The ILA control plane is composed of mapping protocols that manage and disseminate information about the mapping database. There are two levels of mapping protocols: one used by ILA routers that require the full set of ILA mappings for a domain, and one used by ILA nodes that maintain a caches of mappings.

The ILA mapping system is effectively a key/value datastore that maps identifiers to locators. The protocol for sharing mapping information amongst ILA routers can thus be implemented by a distributed database [I-D.herbert-ila-ilamp]. ILA separates the control plane from the data plane, so alternative control plane protocols may be used with a common data plane [I-D.lapukhov-bgp-ila-afi], [I-D.rodrigueznatal-ila-lisp].

The ILA Mapping Protocol [I-D.herbert-ila-ilamp] is used between ILA forwarding nodes and ILA mapping routers. The purpose of the protocol is to populate and maintain the ILA mapping cache in forwarding nodes. ILAMP defines redirects, a request/response protocol, and a push mechanism to populate the mapping table. Unlike traditional routing protocols that run over UDP, this protocol is intended to be run over TCP and may be RPC oriented. TCP provides reliability, statefulness implied by established connections, ordering, and security in the form of TLS. Secure redirects are facilitated by the use of TCP. RPC facilities such REST, Thrift, or GRPC leverage widely deployed models that are popular in SDN.

D.5. IP addressing

ILA supports single address assignments as well as prefix assignments. ILA will also support strong privacy in addressing [ADDRPRIV].

D.5.1. Singleton address assignment

Singleton addresses can use a canonical 64/64 locator/identifier split. Singleton addresses can be assigned by DHCPv6.

D.5.2. Network prefix assignment

Prefix assignment can be done via SLAAC or DHCPv6-PD.

To support /64 prefix assignment with ILA, the ILA identifier can be encoded in the upper sixty-four bits of an address. A level of indirection is used so that ILA transforms the upper sixty four bits to contain both a locator and an index into a locator (ILA-N) specific table. The entry in the table provides the original sixty-four bit prefix so that locator to identifier address transformation can be done.

As an example of this scheme, suppose network has a /24 prefix. The identifier address format for /64 assignment might be:

```
+-------------+---------------------|------------------------------+
|  24 bits    |       40 bits       |           64 bits            |
+-------------+---------------------|------------------------------+
| Network     |      Identifier     |             IID              |
+-------------+---------------------+------------------------------+
```

The IID part is arbitrarily assigned by the device, so that is ignored by ILA. All routing, lookups, and transformations (excepting checksum neutral mapping) are based on the upper sixty-four bits.

For identifier to locator address transformation, a lookup is done on the upper sixty-four bits. That returns a value that contains a locator and a locator table index. The resulting packet format may be something like:

```
+-------------+---------------------|------------------------------+
|   24 bits   | 20 bits | 20 bits   |           64 bits            |
+-------------+---------------------|------------------------------+
|   Network   | Locator | Loc index |             IID              |
+-------------+---------+-----------+------------------------------+
```

The packet is forwarded and routed to the ILA-N addressed by locator (/44 route in this case). At the ILA forwarding node, the locator index is used as a key to an ILA-N specific table that returns a 40 bit Identifier. This value is then written in the packet do ILA to identifier address transformation thereby restoring the original destination address.

The locator index is not globally unique, it is specific to each ILA-N. When a node attaches to an ILA-N, an index is chosen so that the table is populated at the ILA-N and the ILA mapping includes the locator and index. When a node detaches from on ILA, it's entry in the table is removed and the index can be reused after a hold-down period to allow stale mappings to be purged.

D.5.3. Strong privacy addresses

Note that when a /64 is assigned to UEs, the assigned prefix may become a persistent identifier for a device. This is a potential privacy issue. [ADDPRIV] describes this problem and suggests some solutions that may be used with ILA.

D.6. Traffic engineering

ILA is primarily a mechanism for mobility and network virtualization. Transport mechanisms for traffic engineering such as MPLS, network slices, encapsulation, routing based on flow hash(flow label) can be applied independently of ILA. This separation allows any discussion related to transport to be left to operator deployment.

D.7. Locator Chaining with ILA

ILA transformations can be performed on a hop-by-hop bases. In this manner a packet can be source routed through a sequence of nodes. At each hop a determination is made as to the next hop the packet should visit. The locator for the target is then written into the destination. Eventually, the packet will be forwarded to an ILA forwarding node that will restore the original address before delivery to the final destination.

D.8. ILA and network slices

Figure 26 illustrates the use of network slices with ILA.

```
       ----+--------------------------------+---------------------
           |                                |
      +----------------------+ +------------------------+
      | +-------+   Slice #1 | | +-----------+ Slice #2 |
      | | SMF   |----+   GTP | | |  ILA-M    |---+  ILA |
      | +--+----+    |       | | +---------+-+   |      |
      | N4 |         | N4    | |     |     |     |      |
      | +--+--+   +--+----+  | | +-------+ | +--+----+ |   +----+
      | | UPF |   | UPF   |  | | | ILA-N | | | ILA-R | |---| DN |
      | +-----+   +-------+  | | +-------+ | +-------+ |   +----+
      +------------ ---------+ +-----------|------------+
                         |                 |
                      +--+-+  +------------|-------------+
                      | DN |  |            |    Slice #3 |
                      +----+  |     +------+----+    ILA |
                              |     |           |        |
                              | +-------+     +-------+  |   +----+
                     +-----+  | | ILA-N |     | ILA-R |  |---| DN |
                     | MEC |--| +-------+     +-------+  |   +----+
                     +-----+  +--------------------------+
```

Figure 26: ILA and network slices in 5G

In this figure, slice #1 illustrates legacy use of UPFs without ILA in a slice. ILA can be deployed incrementally or in parts of the network. As demonstrated, the use of network slices can provide domain isolation for this.

Slice #2 supports ILA. Some number of ILA-Ns and ILA-Rs are deployed. ILA transformations are performed over the N9 interface. ILA-Rs would be deployed at the N6 interface to perform transformations on packets received from a data network. ILA-Ns will be deployed deeper in the network at one side of the N3 interface. ILA-Ns may be supplemented by ILA-Rs that are deployed in the network. ILA-M manages the ILA nodes and mapping database within the slice.

Slice #3 shows another slice that supports ILA. In this scenario, the slice is for Mobile Edge Computing. The slice contains ILA-Rs and ILA-Ns, and as illustrated, it may also contain ILA_Hs that run directly on edge computing servers. Note in this example, one ILA-M, and hence one ILA domain, is shared between slice #2 and slice #3. Alternatively, the two slices could each have their own ILA-M and define separate ILA domains.

D.9. Security considerations

A mobile public infrastructure has many considerations in security as well as privacy. Fundamentally, a system must protect against misdirection for the purposes of hijacking traffic, spoofing, revealing user identities, exposing accurate geo-location, and Denial of Service attacks on the infrastructure.

The ILA mapping system contains personally identifiable information (PII) including user identities and geo-location. The information must be safeguarded. An ILA domain is confined to one administrative domain, only trusted parties entities in the domain participate in ILA. There is no concept of a global, public mapping system and UEs in public networks generally do not participate in ILA protocols since they are untrusted. ILA control protocols, include ILA redirects, use TCP. TLS or other protocols can be applied for strong security.

Privacy in addressing is a consideration. ILA endeavors to provide a mechanism of address assignment that prevents inference of user identity or location. This problem is described in [ADDRPRIV].

Appendix E. hICN-based mobility architecture

A novel mobility management approach is described in [I-D.auge-hicn-mobility], that leverages routable location-independent identifiers (IDs) and an Information-Centric Networking (ICN) communication model integrated in IPv6, (also referred to as Hybrid ICN, or hICN) [I-D.muscariello-intarea-hicn].

Such approach belongs to the category of pure ID-based mobility management schemes whose objective is (i) to overcome the limitations of traditional locator-based solutions like Mobile IP (conf)using locators as identifiers, (ii) to remove the need for a global mapping system as the one required by locator-identifier separation solutions.

E.1. Motivations

The appeal of purely ID-based architectures is that they move Loc/ID split one step further by embedding ID-awareness in the network and transport layer by default and as such completely decoupling data delivery from underlying network connectivity. Forwarding is performed directly based on identifiers stored in routers' FIBs and no mapping of ID into locators is required. In this way, purely ID-based architectures remove the need to maintain a global mapping system at scale, and its intrinsic management complexity.

Additional benefits brought specifically by ICN principles motivate the consideration of ICN solutions for next generation mobility architectures, e.g.

- the flexibility of multi-source/multi-path connectionless pull-based transport. An example is the native support for consumer mobility, i.e. the transparent emission of data requests over multiple and varying available network interfaces during node mobility;

- the opportunity to define fine-grained per-application forwarding and security policies.

An overview of ICN principles and advantages for a simplified mobility management resulting from name-based forwarding can be found in [RFC7476].

E.2. Hybrid Information-Centric Networking (hICN)

Hybrid ICN (hICN) is an ICN architecture that defines integration of ICN semantics within IPv6, instead of over/under/aside. The only difference w.r.t. ICN as defined in [I-D.irtf-icnrg-ccnxsemantics] is that it encodes names inside IP addresses, and thus can use RFC-compliant TCP/IP packets to transport ICN semantics.

The goal of hICN is to ease ICN insertion in existing IP infrastructure by:

1. selective insertion of hICN capabilities in a few network nodes at the edge (no need for pervasive fully hICN network enablement);

2. guaranteed transparent interconnection with hICN-unaware IP nodes, without using overlays;

3. minor modification to existing IP routers/endpoints (e.g. reuse of IP FIBs and of existing buffers, no modifications to L7 applications and user space hICN transport layer introduction in endpoints);

4. re-use of existing IP control plane (e.g. for routing of IP prefixes carrying ID-semantics) along with performing mobility management and caching operations in forwarding plane;

5. fallback capability to traditional IP network/transport layer.

hICN architecture is described in detail in [I-D.muscariello-intarea-hicn]. Hereafter we focus on mobility management in hICN and on the possible deployment options for insertion in 5G SBA.

E.3. hICN-based mobility

In ICN and hICN endpoints can act as consumers and/or producers. Consumers when they emit requests for named data packets (so called Interests), producers when they send data packets in response to consumers request (pull-based transport model).

Clearly a node can be a consumer and a producer at the same time (e.g. in a voice conversation).

Consumer and producer mobility are handled in a different way due to the pull-based request model. More specifically, consumer mobility is natively supported: consumers pull traffic by sending Interest packets towards named content (wherever produced/stored, the source is a priori unknown by the consumer). Interests are named-based forwarded using the information found in traversed routers' FIBs.

In case of consumer mobility, i.e. mobility of the endpoint issuing the requests, selection of a new available output interface and retransmission of not-yet-satisfied Interests is sufficient for data delivery to continue, independently from the underlying change of locators. Consumer mobility is fully anchorless with hICN, and does not incur any signalization nor tunneling overhead.

Producer mobility is not natively supported by ICN architecture, rather handled in different ways according to the selected producer mobility management scheme.

In [I-D.irtf-icnrg-mapme] and [MAPME] producer mobility schemes are classified into four classes:

o _Resolution based_ solutions rely on dedicated rendez-vous nodes (similar to DNS) which map content names into routable location identifiers. To maintain this mapping updated, the producer signals every movement to the mapping system. Once the resolution is performed, packets can be correctly routed directly to the producer.

o _Anchor-based_ proposals are inspired by Mobile IP, and maintain a mapping at network-layer by using a stable home address advertised by a rendez-vous node, or anchor. This acts as a relay, forwarding through tunneling both interests to the producer, and data packets coming back.

o _Tracing-based_ solutions allow the mobile node to create a hop-by-hop forwarding reverse path from its RV back to itself by propagating and keeping alive traces stored by all involved routers. Forwarding to the new location is enabled without tunneling.

o _Anchorless_ approaches allow the mobile nodes to advertise their mobility to the network without requiring any specific node to act as a rendez-vous point.

The selected mobility management scheme for hICN is MAP-Me, an anchorless producer mobility management solution originally proposed for ICN [I-D.irtf-icnrg-mapme] [MAPME] and further extended to hICN in [I-D.auge-hicn-mobility].

MAP-Me belongs to the class of anchorless approaches that relies on scope-limited forwarding updates triggered by producer mobility events to keep locally up-to-date FIB information for a low-latency guaranteed reroute of consumer Interests towards changing location of the producer.

The difference w.r.t. to other classes of approaches is that it does not require an anchor neither in forwarding plane (traffic does not need to pass through a specific network node), nor in the control plane (no rendez-vous point, no mapping system).

As detailed in [I-D.auge-hicn-mobility] using MAP-Me, hICN provides anchorless consumer and producer mobility removing the need for tunnels and for ID mapping.

Signaling of mobility is only required upon producer movements and limited in scope to current-to-previous network hops. Unlike routing updates, it is not necessary to update all routers' FIBs after a node has moved, but only those located on the path between the new and a former position of the producer. MAP-Me realizes this process in a distributed fashion through in-band signaling packets, hence its anchorless property with respect to the control plane too. Scalability of producer mobility is guaranteed by an efficient and secure FIB update process with minimal and bounded path stretch.

Forwarding and mobility management operations in hICN are based only location-independent identifiers, preserving coexistence with IP locators whose existence may be required by non-hICN services and by control/management plane operations specific to the considered network architecture. As an example, in the case of 3GPP architectures, MAP-Me mobility management does not require an additional control plane anchor.

E.4. hICN insertion in the 3GPP 5G architecture

[I-D.auge-hicn-mobility-deployment-options] reviews various insertion strategies for hICN, including overlay deployments using local breakout to hICN instances situated in MEC, or hICN forwarders deployed within an UPF. While those approaches have the merit of allowing an easy or early integration of hICN and exploiting some of its benefits, they do not fully exploit purely ID-based capabilities nor the dynamic hICN forwarding.

Thus, in this section, we focus our attention on more integrated approaches leveraging hICN-enriched mobile backhaul network to offer an alternative to GTP-U tunnels over the N9 (and possibly N3) Interfaces, as shown in Figure 2 and Figure 8.

E.4.1. Control plane considerations

By operating directly on routers' FIBs for mobility updates, dynamic hop-by-hop forwarding strategies etc., hICN inherits the simplicity of IP forwarding and reuses IP routing protocols for ID prefixes advertisement and routing. In this way it removes the challenges of managing a distributed mapping service at scale (cache update/ refresh, etc.). In addition it remains compatible with the exiting control plane architecture as proposed in the 3GPP standard, with no change required to N1, N2 or N4.

MAP-Me anchorless producer mobility management does not imply SMF interaction, but does not exclude neither to use SMF signaling to trigger MAP-Me updates or to handle FIB updates, at the condition to follow the same procedure described for MAP-Me. However, the absence of SMF interaction might be beneficial in case of dense deployments or failure of the central control entities (infrastructure-less communication scenarios) to empower distributed control of local mobility within an area.

E.4.2. Replacement of N9 interface only

Replacing only the N9 interface (which represents the interface between UPFs, and as such most of the backhaul network) is the initial target of our study. This has the advantage of not touching the gNB as illustrated in Figure 27. The corresponding protocol layering is shown in Figure 10 where we assume hICN-enablement of the end-points (the suboptimal case of hICN enablement via proxies is not considered in this document). We remark that in the protocol layer, hICN is associated to IPv6 PDU layer, transported over N9 directly over L2.

```
     +------------------+          +------------------+
     |       AMF        |          |       SMF        |
     +-+--------------+-+          +-+--------------+-+
       |              |              |              |
       | N1           | N2           | N4           | N4
       |              |              |              |
   +---+---+      +---+---+  N3  +---+---+  N9  +---+---+  N6  +-------+
   |  UE   +------+  gNB  +------+  UPF  +------+  UPF  +------+  DN   |
   +-------+      +-------+      +-------+      +-------+      +-------+
                                            ^
                                            |
                                      hICN insertion
```

Figure 27: Replacement of N9 interface

```
       UE            5G-AN        N3         UPF        N9   UPF    N6
                                  |                     |            |
   +--------+                     |                     |            |
   |  App.  |--------------------------------------------------------|
   +--------+                     |                     | +--------+ |
   | IP PDU |                     |                     | | IP PDU | |
   | (hICN) |---------------------------------------------| (hICN) | |
   +--------+ +-----------------+ | +-----------------+ | |        | |
   |        | |\     relay     /| | |\     decap     /  | |        | |
   |        | | \_____________/ |-|-| \_____________/   | |        | |
   |        | |        | GTP-U  | | | GTP-U  |          | |        | |
   |        | |        +--------+ | +--------+          | |        | |
   |   5G   | |   5G   |  UDP   |-|-|  UDP   |          | |        | |
   |   AN   |-|   AN   +--------+ | +--------+          | |        | |
   |protocol| |protocol|   IP   |-|-|   IP   |          | |        | |
   | layers | | layers +--------+ | +--------+--------+ | +--------+ |
   |        | |        |   L2   |-|-|   L2   |   L2   |-|-|   L2   | |
   |        | |        +--------+ | +--------+--------+ | +--------+ |
   |        | |        |   L1   |-|-|   L1   |   L1   |-|-|   L1   | |
   +--------+ +-----------------+ | +-----------------+ | +--------+ |
                                  |                     |            |
```

Figure 28: Replacement of N9 interface - Protocol layers

E.4.3. Replacement of both N3 and N9 interfaces

This option additionally removes the GTP tunnels between the RAN and the first UPF. It is illustrated in Figure 29 and Figure 30.

```
     +------------------+          +------------------+
     |        AMF       |          |        SMF       |
     +-+--------------+-+          +-+--------------+-+
       |              |              |              |
       | N1           | N2           | N4           | N4
       |              |              |              |
   +---+---+      +---+---+  N3  +---+---+  N9  +---+---+  N6  +-------+
   |  UE   +------+  gNB  +------+  UPF  +------+  UPF  +------+  DN   |
   +-------+      +-------+  ^   +-------+  ^   +-------+      +-------+
                             |              |
                             |              |
                              hICN insertion
```

Figure 29: Replacement of N3 and N9 interfaces

```
       UE            5G-AN        N3         UPF        N9   UPF    N6
                                  |                     |            |
   +--------+                     |                     |            |
   |  App.  |--------------------------------------------------------|
   +--------+                     | +--------+--------+ | +--------+ |
   | IP PDU |                     | | IP PDU | IP PDU | | | IP PDU | |
   | (hICN) |-----------------------| (hICN) | (hICN) |-|-| (hICN) | |
   +--------+ +-----------------+ | |        |        | | |        | |
   |        | |\     decap     /  | |        |        | | |        | |
   |        | | \_____________/   | |        |        | | |        | |
   |        | |        |          | |        |        | | |        | |
   |        | |        |          | |        |        | | |        | |
   |   5G   | |   5G   |          | |        |        | | |        | |
   |   AN   |-|   AN   |          | |        |        | | |        | |
   |protocol| |protocol|          | |        |        | | |        | |
   | layers | | layers +--------+ | +--------+--------+ | +--------+ |
   |        | |        |   L2   |-|-|   L2   |   L2   |-|-|   L2   | |
   |        | |        +--------+ | +--------+--------+ | +--------+ |
   |        | |        |   L1   |-|-|   L1   |   L1   |-|-|   L1   | |
   +--------+ +-----------------+ | +-----------------+ | +--------+ |
                                  |                     |            |
```

Figure 30: Replacement of N3 and N9 interfaces : Protocol layers

E.4.4. Enhanced data plane: hICN/SRv6 combination

hICN is designed to operate inside an IPv6 network by means of an enriched communication layer supporting ICN primitives. The targeted deployment of a few hICN-empowered nodes leads to the tradeoff between incremental deployment and benefits which are proportionally related to the degree of hICN penetration. The association of hICN with other data planes technologies is investigated as a possibility to overcome the above-mentioned tradeoff yielding to a selective, yet fully beneficial insertion of hICN in IP networks.

To this aim, we focus on hICN insertion in a Segment Routing (SR) enhanced data plane, specifically considering SRv6 instantiation of SR. More details are provided in [I-D.auge-hicn-mobility-deployment-options].

hICN/SRv6 combination inherits all SRv6 advantages presented in SR-dedicated section of this document, namely "underlay" management (fast reroute, etc.), service chaining or fine-grained TE for instance.

In addition, it allows extending the reach of hICN on regular IP routers with SRv6 functionality. One realization being to create SRv6 domains in between hICN nodes. The hICN router (through forwarding strategies) would then act as a control plane for SRv6 by specifying the list of SIDs to insert in the packet.

SRv6 forwarding of packets between hICN hops would allow to enforce dynamic per-application hICN forwarding strategies and their objectives (path steering, QoS, etc.), which would be otherwise not possible over not hICN-enabled IP network segments. It would also allow dynamic multi-path and load balancing in hICN-unaware IP network segments and it could guarantee request/reply IP path symmetry (instrumental for efficient round trip delay measurements and rate/congestion control).

E.5. Benefits

Benefits of the deployed solution result both from the purely identifier-based approach, as well as from specific hICN properties. We provide an overview of expected benefits, described in more detail with examples in [I-D.auge-hicn-mobility-deployment-options].

E.5.1. hICN benefits

We review benefits resulting from the deployment of hICN nodes, not specific to the replacement of N9.

- o Low-latency and multicast capabilities by means of in-path edge caching

The ability to satisfy UE requests close to the access is important for latency-sensitive applications such as AR/VR and for cost-efficient transport (that minimizes high-throughput traffic carrying to the core when it can be satisfied locally).

In addition, hICN ability to serve requests from cache or to aggregate them implicitly realizes opportunistic multicast distribution of popular content improving users' QoE for service such as VoD or live streaming, while greatly offloading the mobile core.

o Consumer mobility improvements

The combined use of identifiers and a pull-model allows for seamless mobility across heterogeneous wired and wireless networks, as well as their simultaneous use for multi-homing and bandwidth aggregation. A specific feature of hICN is to allow the use of multiple path/sources of content at the same time, for instance edge caches located at the edge of different fixed/mobile network accesses.

o Network-assisted transport

On-path insertion of hICN enables the use of network buffers to optimize traffic engineering including multipath and load balancing support. One interesting deployment consists in enabling hICN in the PDU Session Anchor, so as to allow two downstream paths back to the mobile edge (due to hICN replacing the datagram source address by the locator of the output UPF), thereby exploiting additional upstream path diversity / resiliency. The ability to perform in-network assistance (for rate/loss/congestion control) is an additional advantage, e.g. to support rate adaptation in the case of dynamic adaptive streaming, or to improve reliability of WiFi connection through transparent wireless detection and recovery [WLDR].

E.5.2. Additional benefits resulting from N9 replacement

o Anchorless consumer and producer mobility

Removing N9 tunnels is key to allow a fully anchorless solution, by not forcing all traffic to transit through the anchoring point in the core. Local forwarding decisions will allow the offloading of device-to-device communications when the different UEs are topologically close, as well as network operations when some areas are disconnected from the core (disaster recovery for instance).

o Dynamic forwarding strategies / UPF selection

Dynamic selection of next hop or exit point is simplified as it can be performed locally based on identifiers and/or locally available information (e.g. interface measurements) in virtue of service-specific forwarding strategies.

E.5.3. Additional benefits resulting from N3 replacement

- Removal of tunnel management state and signaling

The clear advantage is the complete removal of tunnels. Forwarding beyond the radio access is directly managed through hICN. As a consequence, no additional state nor signaling is required for static and mobile consumers, nor for static producers. The impact of producer mobility is low because of the small number of impacted routers.

Deploying hICN on both N3 and N9 further presents the advantage of removing the need for inter-networking with remaining GTP tunnels between those two interfaces. The result is a simpler and lighter architecture, allowing convergence with other non-3GPP accesses.

- Dynamic first UPF selection

Dynamic forwarding capabilities are extended in this configuration to the selection of the first UPF, with the potential of additional performance improvement and higher traffic offload because of the deployment of hICN functionalities closer to the UE. A significant advantage arises in dense deployments scenarios where it becomes possible to isolate the core network from the locally-management mobility (a design objective of the mobile architecture), while allowing distributed selection of ingress UPFs, and dynamic per-packet load balancing of traffic across them.

E.5.4. Deployment considerations

The benefits previously described can be obtained by an upgrade of only a few selected routers at the network edge. The design of hICN allows the rest of the infrastructure to remain unmodified, and to leverage existing management and monitoring tools.

E.5.4.1. hICN in a slice

The use of hICN does not impose any specific slicing of the network. Rather, it can assist a transition of services towards hICN, and/or the coexistence of different hICN deployment options.

As an example of use of hICN in a slice, a service provider might for instance decide to use an hICN-enabled slice dedicated to video delivery, with appropriate mobility management, and dedicated hICN nodes with appropriate caching/forwarding strategies at places aggregating considerable number of user requests.

E.5.4.2. End-to-end deployment

The deployment of an hICN stack in endpoints is the preferred option and offers the full range of benefits. The hICN network stack and forwarder are available through two reference implementations based on the CICN project [CICN]. They share an objective of smooth deployment in existing devices, and are fully userland based. The first is built on top of existing IP primitives and proposed as an application/library for all major OS vendors including iOS, Android, Linux and Windows. The second targets high-performance routers and servers, and leverages the VPP kernel-bypass technology.

E.5.4.3. Network-contained deployment

It is not always possible nor desirable to affect endpoints, and a deployment fully contained in the network, or within the N9 interface is possible through the deployment of proxies. An overview of different options implemented at the network, transport or application level are considered. An example would be the deployment of HTTP proxies at the ingress and egress (resp. first and last UPFs), in order to benefit from content awareness in the network. Such configuration however reduces the flexibility and dynamic forwarding capabilities in endpoints. In particular, existing transport protocols have limited support for dynamically changing paths or network conditions.

Traffic that is not handled though hICN mechanisms can still benefit from the lower overhead and anchorless mobility capabilities coming from the removal of GTP tunnels, as well as dynamic forwarding capabilities that are inherent to the forwarding pipeline. This results from the ability to assign location-independent identifiers to endpoints. It preserves the advantage of removing the mapping system, and of a lightweight FIB update process. No encapsulation is required and packet headers are not modified, which allows the network to have visibility in the source and/or destination identifiers.

E.6. Summary

hICN proposes a general purpose architecture that combines the benefits of a pure-ID architecture with those of ICN. While a full deployment is recommended to make efficient use of available network resources, it is still possible to opt for a partial or phased deployment, with the associated tradeoffs that we have reviewed here.

An hICN enabled network offers native offloading capabilities thanks to the anchorless properties resulting from the pure-ID communication scheme. It does so without the need for a third party mapping system, and further requires no change in the 5G architecture nor in its control plane. The architecture will further leverage the incremental insertion of information centric functionalities through proxies or direct insertion in user devices as the technology gets adopted and deployed.

Authors' Addresses

Kalyani Bogigeni
Verizon

Email: kalyani.bogineni@verizon.com

Arashmid Akhavain
Huawei Canada Research Centre

Email: arashmid.akhavain@huawei.com

Tom Herbert
Quantonium

Email: tom@quantonium.net

Dino Farinacci
lispers.net

Email: farinacci@gmail.com

Alberto Rodriguez-Natal
Cisco

Email: natal@cisco.com

Giovanna Carofiglio
Cisco

Email: gcarofig@cisco.com

Jordan Auge
Cisco

Email: jordan.auge@cisco.com

Shunsuke Homma
NTT

Email: homma.shunsuke@lab.ntt.co.jp