Network Working Group
Internet-Draft
Intended status: Experimental
Expires: October 9, 2016

M. Portoles
V. Ashtaputre
V. Moreno
F. Maino
Cisco Systems
D. Farinacci
lispers.net
April 7, 2016

LISP L2/L3 EID Mobility Using a Unified Control Plane
draft-portoles-lisp-eid-mobility-00

Abstract

The LISP control plane offers the flexibility to support multiple overlay flavors simultaneously. This document specifies how LISP can be used to provide control-plane support to deploy a unified L2 and L3 overlay solution, as well as analyzing possible deployment options and models.

Requirements Language

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

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

This document describes the architecture and design options required to offer a unified L2 and L3 overlay solution with the LISP control-plane.

The architecture takes advantage of the flexibility that LISP provides to simultaneously support different overlay flavors. In this sense, while the LISP specification defines both data-plane and control-plane solutions, this document focuses on the use of the control-plane piece, which can be combined with the data-plane of choice (e.g., [VXLAN-GPE], [VXLAN], or [LISP].

The recommended selection of whether a flow is sent over the L2 or the L3 overlay is mapped to bridged (intra-subnet or non-IP) or routed (inter-subnet) traffic, respectively. This allows treating both overlays as separate segments, and enables L2-only and L3-only deployments (and combinations of them) without modifying the architecture.

The unified solution for L2 and L3 overlays offers the possibility to extend subnets and routing domains (as required in state-of-art Datacenter and Enterprise architectures) with traffic optimization.

An important use-case of the unified architecture is that, while most data centers are complete layer-3 routing domains, legacy applications either have not converted to IP or still use auto-discovery at layer-2 and assume all nodes in an application cluster belong to the same subnet. For these applications, the L2-overlay limits the functionality to where the legacy app lives versus having to extend layer-2 into the network.

Broadcast, Unknown and Multicast traffic on the overlay are supported by either replicated unicast, or underlay (RLOC) multicast as specified in [RFC6831], [I-D.ietf-lisp-signal-free-multicast].

2. Definition of Terms

LISP related terms are defined as part of the LISP specification [RFC6830], notably EID, RLOC, Map-Request, Map- Reply, Map-Notify, Ingress Tunnel Router (ITR), Egress Tunnel Router (ETR), Map- Server (MS) and Map-Resolver (MR).

3. Reference System and Packet Flows

This section introduces a reference system to support the description of the solution and it walks the supported packet flows.

3.1. Reference System

The following figure illustrates the reference system used to support the packet flow description. The system presents 4 sites. Site A and Site D provide access to different subnets (non-extended), while Site B and Site C extend a common subnet. The xTR in each one of the sites registers EIDs from the sites with the LISP Mapping System and provides support to encapsulate overlay (EID) traffic through the underlay (RLOC space).

```
                            ,-------------.
                           (Mapping System )
                            -+-------------+
                           +--+---+  +-+----+
                           |MS/MR|   |MS/MR|
                           +-+----+  +------+
              .-._..-._.--._..|._.,.-._.,|-._.-_._.-.._.-.
          .-.'                                            '.-.
         (                    L3 Underlay                     )
          (                   (RLOC Space)                    )
          '.-._.'.-._.'--'._.'.-._.'.-._.'.-._.'.-._.'.-._.-.'
            /              |                 |               \
    RLOC=IP_A          RLOC=IP_B         RLOC=IP_C         RLOC=IP_D
    +-+--+--+          +-+--+--+         +-+--+--+         +-+--+--+
```

```
   .| xTR A |.-.      .| xTR B |.-.     .| xTR C |.-.     .| xTR D |.-.
  ( +-+--+--+   )    ( +-+--+--+   )   ( +-+--+--+   )   ( +-+--+--+   )
 .'   Site A   )   .'   Site B    )   .'   Site C    )   .'   Site D    )
 ( 1.0.0.0/24 .    ( 3.0.0.0/24  .   ( 3.0.0.0/24 .   (  2.0.0.0/24 .
 '--'._.'.     )   '--'._.'.     )   '--'._.'.     )  '--'._.'.      )
      /  '--'           |  '--'           |   '--'              \ '--'
   '--------'         '--------'        '--------'       '--------'
   :  End   :         :  End   :        :  End   :       :  End   :
   :Device 1:         :Device 2:        :Device 3:       :Device 4:
   '--------'         '--------'        '--------'       '--------'
   IP: 1.0.0.1        IP: 3.0.0.2       IP: 3.0.0.3      IP: 2.0.0.4
                    MAC: 0:0:3:0:0:2   MAC: 0:0:3:0:0:3
```

Figure 1: Reference System Architecture with unified L2 and L3 overlays

3.2. Packet Flows

The recommended selection between the use of L2 and L3 overlays is to map them to bridged (intra-subnet or non-IP) and routed (inter-subnet) traffic. This section follows this recommendation to describe the packet flows.

However, note that in a different selection approach, intra-subnet traffic MAY also be sent over the L3 overlay. Section 6.1 specifies the changes needed to send all IP traffic using the L3 overlay and restricting the use of the L2 overlay to non-IP traffic.

When required, the control plane makes use of two basic types of EID-to-RLOC mappings associated to end-hosts and in order to support the unified architecture:

- o EID = <IID, MAC> to RLOC=<IP>. This is used to support the L2 overlay.

- o EID = <IID, IP> to RLOC=<IP>. This is the traditional mapping as defined in the original LISP specification and supports the L3 overlay.

3.2.1. Bridged Traffic: Intra-subnet and Non-IP

Bridged traffic is encapsulated using the L2 overlay. This section provides an example of the unicast packet flow and the control plane operations when in the topology shown in Figure 1, the End-Device 2 in site B communicates with the End-Device 3 in site C. In this case we assume that End Device 2, knows the MAC address of End-Device 3 (e.g., learned through ARP).

- End-Device 2 sends an Ethernet/IEEE 802 MAC frame with destination 0:0:3:0:0:3 and source 0:0:3:0:0:2.

- ITR B does a L2 lookup in its local map-cache for the destination MAC 0:0:3:0:0:3. When the lookup of 0:0:3:0:0:3 is a miss, the ITR sends a Map-Request to the mapping database system looking up for MAC 0:0:3:0:0:3.

- The mapping systems forwards the Map-Request to ETR C, that has registered the EID-to-RLOC mapping for MAC 0:0:3:0:0:3. Alternatively, depending on the mapping system configuration, a Map-Server which is part of the mapping database system MAY send a Map-Reply directly to ITR B.

- ETR C sends a Map-Reply to ITR B that includes the EID-to-RLOC mapping: MAC 0:0:3:0:0:3 -> RLOC=IP_C, where IP_C is the locator of ETR C.

- ITR B populates the local map-cache with the EID to RLOC mapping, and encapsulates all subsequent packets with a destination MAC 0:0:3:0:0:3 using destination RLOC=IP_C.

3.2.2. Routed Traffic: Inter-subnet

Inter-subnet traffic is encapsulated using the L3 overlay. The process to encapsulate this traffic is the same as described in the original specification [RFC6830]. We describe the packet flow here for completeness

The following is a sequence example of the unicast packet flow and the control plane operations when in the topology shown in Figure 1 End-Device 1, in LISP site A, wants to communicate with End-Device 4 in LISP site D. Note that both end systems reside in different subnets. We'll assume that End-Device 1 knows the EID IP address of End-Device 4 (e.g. it is learned using a DNS service).

- o End-Device 1 sends an IP packet frame with destination 2.0.0.4 and source 1.0.0.1. As the destination address lies on a different subnet End-Device 1 sends the packet following its routing table to ITR A (e.g., it is its default gateway).

- o ITR A does a L3 lookup in its local map-cache for the destination IP 2.0.0.4. When the lookup of 2.0.0.4 is a miss, the ITR sends a Map-request to the mapping database system looking up for IP 2.0.0.4.

- o The mapping systems forwards the Map-Request to ETR D, that has registered the EID-to-RLOC mapping of IP 2.0.0.4.

- o ETR D sends a Map-Reply to ITR A that includes the EID-to-RLOC mapping: EID IP 2.0.0.4 -> RLOC=IP_D, where IP_D is the locator of ETR D.

- o ITR A populates the local map-cache with the EID to RLOC mapping, and encapsulates all subsequent packets with a destination IP 2.0.0.4 using destination RLOC=IP_D.

3.2.3. ARP Resolution

A large majority of applications are IP based and, as a consequence, end systems are typically provisioned with IP addresses as well as MAC addresses.

In this case, to limit the flooding of ARP traffic and reduce the use of multicast in the RLOC network, the LISP mapping system is used to support ARP resolution at the ITR.

In order to provide this support, ETRs handle and register an additional EID-to-RLOC mapping as follows,

- o EID = <IID, IP> to RLOC = <MAC>.

The following packet flow sequence describes the use of the LISP Mapping System to support ARP resolution for hosts residing in a subnet that is extended to multiple sites. Using Figure 1, End-Device 2 tries to find the MAC address of End-Device 3. Note that both have IP addresses within the same subnet:

- End-Device 2 sends a broadcast ARP message to discover the MAC address of End-Device 3. The ARP request targets IP 3.0.0.3.

- ITR B receives the ARP message, but rather than flooding it on the overlay network sends a Map-Request to the mapping database system for IP 3.0.0.3.

- When receiving the Map-Request, the Map-Server sends a Proxy-Map-Reply back to ITR B with the mapping IP 3.0.0.3 -> MAC 0:0:3:0:0:3.

- Using this Map-Reply, ITR B sends an ARP-Reply back to End-Device 2 with the tuple IP 3.0.0.3, MAC 0:0:3:0:0:3.

- End-Device 2 learns MAC 0:0:3:0:0:3 from the ARP message and can now send a L2 traffic to End-Device 3. When this traffic reaches ITR B is sent over the L2-overlay as described above in Section 3.2.1.

This example shows how LISP, by replacing dynamic data plane learning (such as Flood-and-Learn) can reduce the use of multicast in the underlay network.

Note that ARP resolution using the Mapping System is a stateful operation on the ITR. The source IP,MAC tuple coming from the ARP request have to be stored to generate the ARP-reply when the Map-Reply is received.

Note that the ITR SHOULD cache the ARP entry. In that case future ARP-requests can be handled without sending additional Map-Requests.

4. LISP Protocol with L2 and L3 Support

This section describes the specific elements required to support a unified L2 and L3 overlay solution and the packet flows described in the previous section.

4.1. L2 and L3 Segmentation

LISP support of segmentation and multi-tenancy is structured around the propagation and use of Instance-IDs, and handled as part of the EID in control plane operations. The encoding is described in [I-D.ietf-lisp-lcaf] and its use in [I-D.ietf-lisp-ddt].

Depending on the particular deployment, both the L2 and L3 overlays can be segmented. An Instance-ID can be used for L2 overlay segmentation (e.g., VLAN extension) and for L3 overlay segments (e.g., VRF extension or multi-VPN overlays). In all cases, the Instance-ID is a 24-bit value. Instance-IDs are unique to a Mapping System and MAY be used to identify the overlay type.

An important aspect of L2 overlay segmentation is the mapping of VLANs to IIDs. In this case a Bridge Domain (which is the L2 equivalent to a VRF as a forwarding context) maps to an IID, a VLAN-ID may map 1:1 to a bridge domain or different VLAN-IDs on different ports may map to a common Bridge Domain, which in turn maps to an IID in the L2 overlay. When ethernet traffic is double tagged, usually the external 802.1Q tag will be mapped to a bridge domain on a per port basis, and the inner 802.1Q tag will remain part of the payload to be handled by the overlay. The IID should therefore be able to carry ethernet traffic with or without an 802.1Q header. A port may also be configured as a trunk and we may chose to take the encapsulated traffic and map it to a single IID in order to multiplex traffic from multiple VLANs on a single IID. These are all examples of local operations that could be effected on VLANs in order to map them to IIDs, they are provided as examples and are not exhaustive.

4.2. Database Mappings in Unified L2 and L3 Overlays

When an end-host is attached or detected in an ETR that provides L2-overlay and L3-overlay services, two Database Mapping entries are registered to the mapping system:

- o The EID 2-tuple (IID, IP) with its binding to a corresponding ETR locator set (IP RLOC)
- o The EID 2-tuple (IID, MAC) with its binding to a locator set (IP RLOC)

These two database mappings are the ones required to support L3 and L2 forwarding.

The registration of these EIDs MUST follow the LCAF format as defined in [I-D.ietf-lisp-lcaf].

4.3. MAC as a Locator Record for ARP Resolution

When an end-host is attached or detected in an ETR that provides L2-overlay services and supports ARP resolution using the LISP control-plane, an additional mapping entry is registered to the mapping system:

- o The EID 2-tuple (IID, IP) and its binding to a corresponding host MAC address.

In this case both the xTRs and the Mapping System MUST support an EID-to-RLOC mapping where the MAC address is set as a locator record.

This mapping is registered with the Mapping System using the following EID record structure,

```
+-> +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   |                          Record TTL                           |
|   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
E   | Locator Count | EID mask-len  | ACT |A|       Reserved        |
I   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
D   | Rsvd  |  Map-Version Number   |          AFI = 16387          |
|   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
r   |    Rsvd1      |     Flags     |   Type = 2    | IID mask-len  |
e   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
c   |             4 + n             |         Instance-ID...        |
o   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
r   |        ...Instance-ID         |         EID-AFI = 1 or 2      |
d   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|   |                    EID-Prefix (IPv4 or IPv6)                  |
|   +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  /|    Priority   |    Weight     |  M Priority   |   M Weight    |
| | +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| M |        Unused Flags     |L|p|R|        AFI  = 16387           |
| A +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| C |    Rsvd1      |     Flags     |   Type = 1    |     Rsvd2     |
| | +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| | |             2 + 6             |             AFI = 6           |
| | +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| | |                    Layer-2 MAC Address  ...                   |
| | +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|  \| ... Layer-2 MAC Address       |
+-> +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+.
```

An EID record with a locator record that carries a MAC address follows the same structure as described in [RFC6830]. However, some fields of the EID record and the locator record require special consideration:

Locator Count: This value SHOULD be set to 1.

Instance-ID: This is the IID used to provide L2-overlay segmentation.

Priority and Weight: IP to MAC bindings are one to one bindings. An

ETR SHOULD not register more than one MAC address in the locator record together with an IP based EID. The Priority of the MAC address record is set to 255. The Weight value SHOULD be ignored and the recommendation is to set it to 0.

L bit: This bit of the locator record SHOULD only be set to 1 when an ETR is registering its own IP to MAC binding.

p bit: This bit of the locator record SHOULD be set to 0.

R bit: This bit of the locator record SHOULD be set to 0.

Note that an IP EID record that carries a MAC address in the locator record, SHOULD be registered with the Proxy Map-Reply bit set.

4.4. LISP Mapping System

The interface between the xTRs and the Mapping System is described in [RFC6833] and this document follows the specification as described there. When available, the registrations MAY be implemented over a reliable transport as described in [I-D.kouvelas-lisp-map-server-reliable-transport].

4.5. Time-to-Live Handling in Data-Plane

The LISP specification ([RFC6830]) describes how to handle Time-to-Live values of the inner and outer headers during encapsulation and decapsulation of packets when using the L3 overlay. The same approach SHOULD be followed for the unified overlay.

When using the L2 overlay and the encapsulated traffic is IP traffic, the Time-to-Live value of the inner IP header MUST remain unmodified at encapsulation and decapsulation. Network hops traversed as part of the L2 overlay SHOULD be hidden to tools like traceroute and applications that require direct L2 connectivity.

4.6. Using SMRs to Track Moved-Away State

One of the key elements to support end-host mobility using the LISP architecture is the Solicit-Map-Request (SMR). This is an special message by means of which an ETR can request an ITR to send a new Map-Request for a particular EID record. In essence the SMR message is used as a signal to indicate a change in mapping information and it is described with detail in [RFC6830]. Section 5 provides a packet flow description of the mobility support in a unified overlay.

In order to support mobility, an ETR SHALL maintain a list of EID records for which it has to generate a SMR message whenever it receives traffic with that EID as destination. This is called an Away Table.

The particular strategy to maintain an Away Table is implementation specific and it will be typically based on the strategy to detect the presence of hosts and the use of the Map-Notifies received from the Map-Server. In essence the table SHOULD provide support to the following:

- o Keep track of end-hosts that were once connected to an ETR and have moved away.
- o Support for L2 EID records, the 2-tuple (IID, MAC), for which a SMR message SHOULD be generated.
- o Support for L3 EID records, the 2-tuple (IID, IP), for which a SMR message SHOULD be generated.

4.7. Non-Extended Subnets

The registration and access to non-extended subnets using the L3 overlay follows [RFC6830].

Note also that non-extended subnets can also be treated as non-LISP networks. In that case, providing access to the subnet to participants of LISP overlays requires the use of a PxTR, following the specification in [RFC6830].

4.8. L2 Overlays and Multicast Groups

xTRs that offer L2 overlay services and are part of the same Instance-ID join a common Multicast Group. When required, this group allows ITRs to send traffic that needs to be replicated (flooded) to all ETRs participating in the L2-overlay (e.g., broadcast traffic within a subnet). When the core network (RLOC space) supports native multicast ETRs participating in the L2-overlay join a (*,G) group associated to the Instance-ID.

When multicast is not available in the core network, each xTR that is part of the same instance-ID SHOULD join a (S,G) entry to the mapping system using the procedures described in [I-D.ietf-lisp-signal-free-multicast], where S is 0000-0000-0000/0 and G is ffff-ffff-ffff/48. This strategy allows and ITR to know which ETRs are part of the L2 overlay and it can head-end replicate traffic to.

4.9. L2 Broadcast, Unknown Unicast and Multicast traffic

Broadcast, Unknown Unicast and Multicast (BUM) traffic on the L2-overlay is supported by either replicated unicast, or underlay (RLOC) multicast.

5. EID Mobility Support

Support to end-host mobility is a basic requirement of state-of-art overlay solutions. The unified overlay architecture described here supports mobility both over L2-overlays and L3-overlays. This

section provides a packet flow sequence description of the mobility support, detailing the use of the elements defined in the previous section.

5.1. Reference Architecture

In order to support the packet flow description we use again the same example as in Figure 1. In this particular case hosts may roam and we distinguish the case when we have L2-mobility (e.g., IP hosts move within their own subnet) or L3-mobility.

```
           /                |                 |               \
   RLOC=IP_A          RLOC=IP_B         RLOC=IP_C         RLOC=IP_D
   +-+--+--+          +-+--+--+         +-+--+--+         +-+--+--+
  .| xTR A |.-.      .| xTR B |.-.     .| xTR C |.-.     .| xTR D |.-.
 ( +-+--+--+   )    ( +-+--+--+   )   ( +-+--+--+   )   ( +-+--+--+   )
.'   Site A   )   .'   Site B    )   .'   Site C   )   .'   Site D   )
```

```
( 1.0.0.0/24 .    ( 3.0.0.0/24  .     ( 3.0.0.0/24 .   (  2.0.0.0/24 .
'            )     '            )
'--'._.'.     )    '--'._.'.     )    '--'._.'.    )   '--'._.'.     )
       /  '--'            |  '--'          |   '--'             \ '--'
   '---------'         '---------'       '---------'       '---------'
   :  End   :          :  End   :        :  End   :        :  End   :
   :Device 1:          :Device 2:        :Device 3:        :Device 4:
   '---------'         '---------'       '---------'       '---------'
 (IID1,1.0.0.1)      (IID1,3.0.0.2)     (IID1,3.0.0.3)    (IID1,2.0.0.4)
                  (IID2,0:0:3:0:0:2)   (IID2,0:0:3:0:0:3)
```

Figure 2: Reference Mobility Architecture

5.2. L2 Mobility: Packet Flow

The support to L2 mobility covers the requirements to allow an end-host to move from a given site to another and maintain correspondence with the rest of end-hosts that are connected to the same L2 domain (e.g. extended VLAN). This support MUST ensure convergence of L2 forwarding (MAC based) from the old location to the new one, when the host completes its move.

The following is a sequence description of the packet flow when End-Device 2 in the figure moves to Site C, which is extending its own subnet network.

- o When End-Device 2 is attached or detected in site C, ETR C sets up the database mapping corresponding to EID=<IID2, 0:0:3:0:0:2>. ETR C sends a Map-Register to the mapping system registering RLOC=IP_B as locator for EID=<IID2, 0:0:3:0:0:2>.

- o The Mapping System, after receiving the new registration for EID=<IID1, 0:0:3:0:0:2> sends a Map-Notify to ETR B with the new locator set (IP_B). ETR B removes then its local database mapping and stops registering <IID2, 0:0:3:0:0:2>.

- o Any PiTR or ITR participating in the same L2-overlay (corresponding to IID2) that was encapsulating traffic to 0:0:3:0:0:2 before the migration continues encapsulating this traffic to ETR B.

- o Once ETR B is notified by the Mapping system, when it receives traffic from an ITR which is destined to 0:0:3:0:0:2, it will generate a Solicit-Map-Request (SMR) message that is sent to the ITR for (IID2,0:0:3:0:0:2).

- Upon receiving the SMR the ITR sends a new Map-Request for the EID=<IID2,0:0:3:0:0:2>. As a response ETR B sends a Map-Reply that includes the new EID-to-RLOC mapping for <IID2,0:0:3:0:0:2> with RLOC=IP_B. This entry is cached in the L2 table of the ITR, replacing the previous one, and traffic is then forwarded to the new location.

5.3. L3 Mobility: Packet Flow

The support to L3 mobility covers the requirements to allow an end-host to move from a given site to another and maintain correspondence with the rest of end-hosts that are connected to the same L3 routing domain. This support MUST ensure convergence of L3 forwarding (IPv4/IPv6 based) from the old location to the new one when the host completes its move.

The following is a sequence description of the packet flow when End-Device 1 in the reference figure roams to site D:

- When End-Device 1 is attached or detected in site D, ETR D sets up the database mapping corresponding to EID=<IID1, 1.0.0.1>. ETR D sends a Map-Register to the mapping system registering RLOC=IP_D as locator for EID=<IID1, 1.0.0.1>. Now the mapping system is updated with the new EID-to-RLOC mapping (location) for End-Device 1.

- The Mapping System, after receiving the new registration for EID=<IID1, 1.0.0.1> sends a Map-Notify to ETR A to inform it of the move. Then, ETR A removes its local database mapping information and stop registering EID=<IID1, 1.0.0.1>.

- Any ITR or PiTR participating in the L3 overlay (corresponding to IID1) that were sending traffic to 1.0.0.1 before the migration keep sending traffic to ETR A.

o Once ETR A is notified by the Mapping system, when it receives traffic from an ITR with destination 1.0.0.1, it generates a Solicit-Map-Request (SMR) back the ITR (or PiTR) for EID=<IID1, 1.0.0.1>.

o Upon receiving the SMR the ITR invalidates its local map- cache entry for EID=<IID1, 1.0.0.1> and sends a new Map-Request for that EID. The Map-Reply includes the new EID-to-RLOC mapping for End-Device 1 with RLOC=IP_D.

o Similarly, once the local database mapping is removed from ITR A, non-encapsulated packets arriving at ITR A from a local End-Device and destined to End-Device 1 result in a cache miss, which triggers sending a Map-Request for EID=<IID1, 1.0.0.1> to populate the map-cache of ITR A.

6. Optional Deployment Models

The support of an integrated L2 and L3 overlay solution takes multiple architectural design options, that depend on the specific requirements of the deployment environment. While some of the previous describe specific packet flows and solutions based on the recommended solution, this section documents OPTIONAL design considerations that differ from the recommended one but that MAY be required on alternative deployment environments.

6.1. IP Forwarding of Intra-subnet Traffic

As pointed out at the beginning the recommended selection of the L2 and L3 overlays is not the only one possible. In fact, providing L2 extension to some cloud platforms is not always possible and subnets need to be extended using the L3 overlay.

In order to send all IP traffic (intra- and inter-subnet) through the L3 overlay the solution MUST change the ARP resolution process described in Section 3.2.3 to the following one (we follow again Figure 1 to drive the example. End-Device 2 queries about End-Device 3):

o End-Device 1 sends a broadcast ARP message to discover the MAC address of 3.0.0.3.

o ITR B receives the ARP message and sends a Map-Request to the Mapping System for 3.0.0.3.

- In this case, the Map-Request is routed by the Mapping system infrastructure to ETR C, that will send a Map-Reply back to ITR B containing the mapping 3.0.0.3 -> RLOC=IP_C.

- ITR B populates its local cache with the received entry on the L3 forwarding table. Then, using the cache information it sends a Proxy ARP-reply with its own MAC (MAC_xTR_B) address to end End-Device 1.

- End-Device 1 learns MAC_ITR_B from the proxy ARP-reply and sends traffic with destination address 3.0.0.3 and destination MAC, MAC_xTR_B.

- As the destination MAC address is the one from xTR B, when xTR B receives this traffic is it forwarded using the L3-overlay.

- Note that when implementing this solution, when a host that is local to an ETR moves away, the ETR SHOULD locally send a Gratuitous ARP with its own MAC address and the IP of the moved host, to refresh the ARP tables of local hosts and guarantee the use of the L3 overlay when connecting to the remote host.

It is also important to note that using this strategy to extend subnets through the L3 overlay but still keeping the L2 overlay for the rest of traffic MAY lead to flow asymmetries. This MAY be the case in deployments that filter Gratuitous ARPs, or when moved hosts continue using actual L2 information collected before a migration.

6.2. Data-plane Encapsulation Options

The LISP control-plane offers independence from the data-plane encapsulation. Any encapsulation format that can carry a 24-bit instance-ID can be used to provide the unified overlay.

Common encapsulation formats that can be used are [VXLAN-GPE], [LISP] and [VXLAN]:

- VXLAN-GPE encap: This encapsulation format is defined in [I-D.ietf-nvo3-vxlan-gpe]. It allows encapsulation both L2 and L3 packets and the VNI field directly maps to the Instance-ID used in the control plane. Note that when using this encapsulation for a unified solution the P-bit is set and the Next-Protocol field is used (typically with values 0x1 (IPv4) or 0x2 (IPv6) in L3-overlays, and value 0x3 in L2-overlays).

- LISP encap: This is the encapsulation format defined in the original LISP specification [RFC6830]. The encapsulation allows encapsulating both L2 and L3 packets. The Instance-ID used in the

EIDs directly maps to the Instance-ID that the LISP header carries. At the ETR, after decapsulation, the IID MAY be used to decide between L2 processing or L3 processing.

- VXLAN encap: This is a L2 encapsulation format defined in [RFC7348]. While being a L2 encapsulation it can be used both for L2 and L3 overlays. The Instance-ID used in LISP signaling maps to the VNI field of the VXLAN header. Providing L3 overlays using VXLAN generally requires using the ETR MAC address as destination MAC address of the inner Ethernet header. The process to learn or derive this ETR MAC address is not included as part of this document.

6.3. L2-only Deployments

The Unified architecture that this document specifies allows the deployment of L3-only or L2-only overlays. By having a single control plane where the mapping database can hold both MAC EIDs and IP EIDs, the deployment of L2-only or L3-only architectures consists in using only the relevant database mappings.

The requirements and use of LISP to support a L3-only overlay are extensively documented in the original LISP specification and related documents.

The provision of a L2-only overlay MUST provide support for intra-subnet connectivity of end-hosts belonging to the same tenant, including them in a unique L2 broadcast domain extended across the network.

Provision such L2-only overlay SHALL take the following aspects into account, as described before in Section 4:

- When an end-host is attached the ETR maintains and registers the mappings EID = <IID, MAC> -> RLOC = <IP> and EID = <IID, IP> -> RLOC = <MAC>. The second mapping is optional and is meant to be used for ARP resolution.

- An ITR and Mapping-System provides support for ARP lookup and MAC lookup using the lisp control-plane as described before in this document.

o xTRs MUST provide support for Broadcast, Unknown and Multicast (BUM) traffic through either replicated unicast or underlay (RLOC) multicast.

o An ITR MUST treat a destination MAC for which it receives a Negative Map-Reply with Native Forward action as BUM traffic and replicate it to the ETRs in the Layer-2 overlay.

o To support end-host mobility, an ETR MUST be able to support an Away Table (as described above) to keep track of end-hosts and generate SMR messages when receiving traffic for end-hosts not locally attached.

o TTL value of the inner-IP header SHOULD not be modified when traversing the L2 overlay.

7. IANA Considerations

This memo includes no request to IANA.

8. Acknowledgements

This draft builds on top of two expired drafts that introduced the concept of LISP L2/L3 overlays (draft-maino-nvo3-lisp-cp and draft-hertoghs-nvo3-lisp-controlplane-unified). Many thanks to the combined authors of those drafts, that SHOULD be considered main contributors of this draft as well: Vina Ermagan, Dino Farinacci, Yves Hertoghs, Luigi Iannone, Fabio Maino, Victor Moreno, and Michael Smith.

9. References

9.1. Normative References

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

<http://www.rfc-editor.org/info/rfc2119>.

[RFC6830] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "The Locator/ID Separation Protocol (LISP)", RFC 6830, DOI 10.17487/RFC6830, January 2013, <http://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, <http://www.rfc-editor.org/info/rfc6831>.

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

[RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "Virtual eXtensible Local Area Network (VXLAN): A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014, <http://www.rfc-editor.org/info/rfc7348>.

9.2. Informative References

[I-D.ietf-lisp-ddt] Fuller, V., Lewis, D., Ermagan, V., and A. Jain, "LISP Delegated Database Tree", draft-ietf-lisp-ddt-04 (work in progress), March 2016.

[I-D.ietf-lisp-lcaf] Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", draft-ietf-lisp-lcaf-11 (work in progress), September 2015.

[I-D.ietf-lisp-signal-free-multicast] Moreno, V. and D. Farinacci, "Signal-Free LISP Multicast", draft-ietf-lisp-signal-free-multicast-00 (work in progress), December 2015.

[I-D.ietf-nvo3-vxlan-gpe]
Quinn, P., Manur, R., Kreeger, L., Lewis, D., Maino, F., Smith, M., Agarwal, P., Yong, L., Xu, X., Elzur, U., Garg, P., and D. Melman, "Generic Protocol Extension for VXLAN", draft-ietf-nvo3-vxlan-gpe-01 (work in progress), November 2015.

[I-D.kouvelas-lisp-map-server-reliable-transport]
Cassar, C., Kouvelas, I., Lewis, D., Arango, J., and J. Leong, "LISP Map Server Reliable Transport", draft-kouvelas-lisp-map-server-reliable-transport-01 (work in progress), February 2016.

Authors' Addresses

Marc Portoles Comeras
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
USA

Email: mportole@cisco.com

Vrushali Ashtaputre
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
USA

Email: vrushali@cisco.com

Victor Moreno

Cisco Systems
170 Tasman Drive
San Jose, CA 95134
USA

Email: vimoreno@cisco.com

Fabio Maino
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
USA

Email: fmaino@cisco.com

Dino Farinacci
lispers.net
San Jose, CA
USA

Email: farinacci@gmail.com