

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: April 11, 2017

M. Portoles
V. Ashtaputre
V. Moreno
F. Maino
Cisco Systems
D. Farinacci
lispers.net
October 8, 2016

LISP L2/L3 EID Mobility Using a Unified Control Plane
draft-portoles-lisp-eid-mobility-01

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

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

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

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

This Internet-Draft will expire on April 11, 2017.

Copyright Notice

Copyright (c) 2016 IETF Trust and the persons identified as the document authors. All rights reserved.

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

Table of Contents

1. Introduction	3
2. Definition of Terms	4
3. Reference System	4
4. L3 Overlays and Mobility Support	5
4.1. Reference Architecture and packet flows	5
4.1.1. Routed Traffic Flow: L3 Overlay use	6
4.1.2. L3 Mobility Flow	6
4.2. Implementation Considerations	7
4.2.1. L3 Segmentation	7
4.2.2. L3 Database-Mappings	8
4.2.3. LISP Mapping System support	8
4.2.4. Using SMRs to Track Moved-Away Hosts	9
4.2.5. L3 multicast support	9
4.2.6. Time-to-Live Handling in Data-Plane	9
5. L2 Overlays and Mobility Support	9
5.1. Reference Architecture and packet flows	10
5.1.1. Bridged Traffic Flow: L2 Overlay use	10
5.1.2. L2 Mobility Flow	11
5.2. Implementation Considerations	12
5.2.1. L2 Segmentation	12
5.2.2. L2 Database-Mappings	12
5.2.3. Interface to the LISP Mapping System	13
5.2.4. SMR support to track L2 hosts that moved away	13
5.2.5. L2 Broadcast and Multicast traffic	14
5.2.6. L2 Unknown Unicast Support	14
5.2.7. Time-to-Live Handling in Data-Plane	15
5.3. Support to ARP resolution through Mapping System	15
5.3.1. Map-Server support to ARP resolution: Packet Flow	15
5.3.2. ARP registrations: MAC as a locator record	16
5.3.3. Implementation Considerations	18
6. Optional Deployment Models	19

6.1. IP Forwarding of Intra-subnet Traffic	19
6.2. Data-plane Encapsulation Options	20
7. IANA Considerations	21
8. Acknowledgements	21
9. References	21
9.1. Normative References	21
9.2. Informative References	22
Authors' Addresses	22

1. Introduction

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

The architecture takes advantage of the flexibility that LISP provides to simultaneously support different overlay types. While the LISP specification defines both the data-plane and the control-plane, this document focuses on the use of the control-plane to provide L2 and L3 overlay services with mobility. The control plane may be combined with a data-plane of choice e.g., [LISP], [VXLAN-GPE], or [VXLAN].

The recommendation on whether a flow is sent over the L2 or the L3 overlay is based on whether the traffic is bridged (intra-subnet or non-IP) or routed (inter-subnet), 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 mobility support and 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] and [I-D.ietf-lisp-signal-free-multicast].

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. The rest of the document 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 LIISP specification and supports the L3 overlay.

4. L3 Overlays and Mobility Support

4.1. Reference Architecture and packet flows

In order to support the packet flow descriptions in this section we use Figure 1 as reference. This section uses Sites A and D to describe the flows.

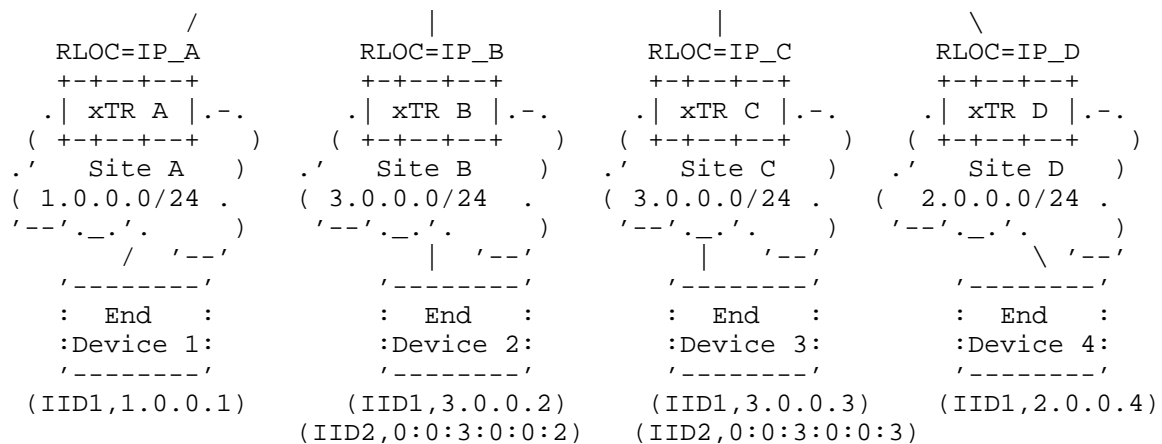


Figure 2: Reference Mobility Architecture

4.1.1.1. Routed Traffic Flow: L3 Overlay use

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 EID=<IID1,2.0.0.4>.
- o The mapping systems forwards the Map-Request to ETR D, that has registered the EID-to-RLOC mapping of EID=<IID1,2.0.0.4>.
- o ETR D sends a Map-Reply to ITR A that includes the EID-to-RLOC mapping: EID=<IID1,2.0.0.4> -> 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.

4.1.1.2. L3 Mobility 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:

- o 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.
- o 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 stops registering EID=<IID1, 1.0.0.1>.
- o 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 to 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.

4.2. Implementation Considerations

4.2.1. 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].

Instance-IDs can be used to support L3 overlay segmentation, such as in extended VRFs or multi-VPN overlays.

this same EID. This Map-Notify is used to track moved-away state of L3 EIDs as described in Section 4.2.4.

4.2.4. Using SMRs to Track Moved-Away Hosts

One of the key elements to support end-host mobility using the LISP architecture is the Solicit-Map-Request (SMR). This is a 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].

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.

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 Map-Notify messages 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 L3 EID records, the 2-tuple (IID, IP), for which a SMR message SHOULD be generated.

4.2.5. L3 multicast support

L3 Multicast traffic on the overlay MAY be supported by either replicated unicast, or underlay (RLOC) multicast. Specific solutions to support L3 multicast over LISP controlled overlays are specified in in [RFC6831], [I-D.ietf-lisp-signal-free-multicast] and [I-D.coras-lisp-re].

4.2.6. 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.

5. L2 Overlays and Mobility Support

5.1. Reference Architecture and packet flows

In order to support L2 packet flow descriptions in this section we use Figure 1 as reference. This section uses Sites B and C to describe the flows.

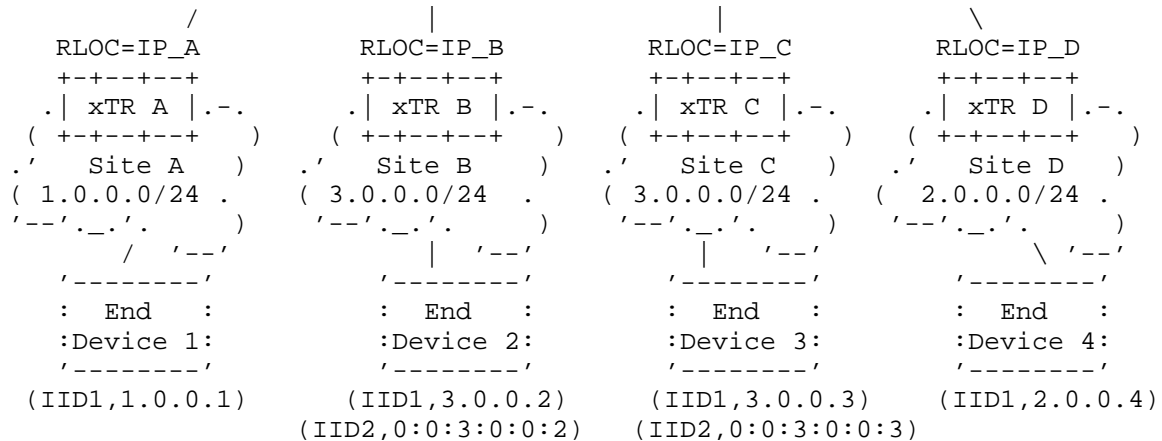


Figure 3: Reference Mobility Architecture

5.1.1. Bridged Traffic Flow: L2 Overlay use

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).

- o 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.
- o 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 EID=<IID2,0:0:3:0:0:3>.
- o The mapping systems forwards the Map-Request to ETR C, that has registered the EID-to-RLOC mapping for EID=<IID2,0:0:3:0:0:3>. 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.

- o ETR C sends a Map-Reply to ITR B that includes the EID-to-RLOC mapping: EID=<IID2, 0:0:3:0:0:3> -> RLOC=IP_C, where IP_C is the locator of ETR C.
- o 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.

5.1.2. L2 Mobility 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).
- o 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.

5.2. Implementation Considerations

5.2.1. L2 Segmentation

As with L3 overlays, LISP support of L2 segmentation 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]. Instance-IDs are unique to a Mapping System and MAY be used to identify the overlay type (e.g., L2 or L3 overlay).

An Instance-ID can be used for L2 overlay segmentation. An important aspect of L2 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 outer 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.

5.2.2. L2 Database-Mappings

When an end-host is attached or detected in an ETR that provides L2-overlay services, a database Mapping is registered to the mapping system with the following structure:

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

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

```
+--> +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| |                                     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 = 6          |
d +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| |                     Layer-2 MAC Address ...                      |
| +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
/ | ... Layer-2 MAC Address                |      Priority      |      Weight      |
L +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
o | M Priority   | M Weight   |               Unused Flags               |L|p|R|
c +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| |                 Loc-AFI                  |      Locator....     |
| +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
\ |                 ... Locator
+--> +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
```

The L2 EID record follows the structure as described in [RFC6830].

5.2.3. Interface to the 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].

In order to support system convergence after mobility, when the Map-Server receives a new registration for a specific EID, it MUST send a Map-Notify to the entire RLOC set in the site that last registered this same EID. This Map-Notify is used to track moved-away state of L2 EIDs as described in Section 5.2.4.

5.2.4. SMR support to track L2 hosts that moved away

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.

The particular strategy to maintain a SMR table is implementation specific. In essence the table SHOULD provide support for 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.

5.2.5. L2 Broadcast and Multicast traffic

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

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 register 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 an ITR to know which ETRs are part of the L2 overlay and it can head-end replicate traffic to.

Following the same case, when multicast is not available in the core network, the procedures in [I-D.ietf-lisp-signal-free-multicast] can be used to ensure proper distribution of link-local multicast traffic across xTRs participating in the L2 overlay. In such case, the xTRs SHOULD join a (S,G) entry with S being 0000-0000-0000/0 and where G is 0100-0000-0000/8.

5.2.6. L2 Unknown Unicast Support

An ITR attempts to resolve MAC destination misses through the Mapping System. When the destination host remains undiscovered the destination is considered an Unknown Unicast.

A Map-Server SHOULD respond to a Map-Request for an undiscovered host with a Negative Map-Reply with action "Native Forward". Alternatively the action "Drop" may be used in order to suppress Unknown Unicast forwarding.

An ITR that receives a Negative Map-Reply with Action "Native Forward" will handle traffic for the undiscovered host as L2 Broadcast traffic and will be unicast replicated or flooded using underlay multicast to the rest of ETRs in the Layer-2 overlay.

Upon discovery of a previously unknown unicast MAC EID, a data triggered SMR for the discovered EID should be sent by the discovery ETR back to the ITRs that are flooding the unknown unicast traffic. This would allow ITRs to refresh their caches and stop flooding unknown unicast traffic as necessary.

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

When using a L2 overlay and the encapsulated traffic is IP traffic, the Time-to-Live value of the inner IP header MUST remain unmodified during 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.

5.3. Support to ARP resolution through Mapping System

5.3.1. Map-Server support to ARP resolution: Packet Flow

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 MAY be 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-record contents = <IID, IP>, RLOC-record contents <MAC>.

There is a dedicated IID used for the registration of the ARP related mappings. Thus, a system with L2 and L3 overlays as well as ARP mappings would have three IIDs at play. In the spirit of providing clarity, we will refer to those IIDs as L2-IID, L3-IID and ARP-IID respectively. By using these definitions, we do not intend to coin new terminology, nor is there anything special about those IIDs that would make them differ from the generic definition of an IID. The types of mappings expected in such a system are summarized below:

- o EID = <IID1, IP> to RLOC = <IP-RLOC> (L3-overlay)
- o EID = <IID2, MAC> to RLOC = <IP-RLOC> (L2-overlay)

- o EID = <IID3, IP> to RLOC = <MAC-RLOC> (ARP/ND mapping)

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:

- o 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.
- o 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 EID = <IID2,3.0.0.3>.
- o When receiving the Map-Request, the Map-Server sends a Proxy-Map-Reply back to ITR B with the mapping EID = <IID2,3.0.0.3> -> MAC 0:0:3:0:0:3.
- o 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.
- o 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 5.1.1.

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.

5.3.2. ARP registrations: MAC as a locator record

When an end-host is attached or detected in an ETR that provides L2-overlay services and also 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.

In order to guarantee compatibility with current implementations of xTRs, the MAC locator record SHALL be encoded following the AFI-List LCAF Type defined in [I-D.ietf-lisp-lcaf]. This option would also allow adding additional attributes to the locator record, while maintaining compatibility with legacy devices.

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 segmentation of the L2-overlays, L3 overlays and ARP tables.

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, SHALL be registered with the Proxy Map-Reply bit set.

5.3.3. Implementation Considerations

While ARP support through the LISP Mapping System fits the LISP Control-Plane there are a series of considerations to take into account when providing this feature:

- o As indicated, when an end-host is attached the ETR maintains and registers a mapping with the binding EID = <IID, IP> -> RLOC = <MAC>.
- o ARP support through the LISP Mapping System is OPTIONAL and the xTRs should allow the possibility to enable or disable the feature.
- o When the ARP entry has not been registered, a Map Server SHOULD send a Negative Map-Reply with action "No Action" as a response to an ARP based Map Request.
- o In case the ITR receives a Negative Map-Reply for an ARP request it should fallback to flooding the ARP packet as any other L2 Broadcast packet (as described in Section 5.2.5).
- o When receiving a positive Map-Reply for an ARP based Map-Request, the ETR MUST recreate the actual ARP Reply, impersonating the real host. As a consequence, ARP support is a stateful operation where the ITR needs to store enough information about the host that generates an ARP request in order to recreate the ARP Reply.

- o ARP replies learned from the Mapping System SHOULD be cached and the information used to reply to subsequent ARP requests to the same hosts.

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 5.3.1 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 EID = <IID1,3.0.0.3>.
- o 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 EID = <IID1,3.0.0.3> -> RLOC=IP_C.
- o 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.
- o 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.
- o As the destination MAC address is the one from xTR B, when xTR B receives this traffic it is forwarded using the L3-overlay.

- o 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]:

- o VXLAN-GPE encaps: 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).
- o LISP encaps: 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.
- o VXLAN encaps: 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.

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.coras-lisp-re]
Coras, F., Cabellos-Aparicio, A., Domingo-Pascual, J., Maino, F., and D. Farinacci, "LISP Replication Engineering", draft-coras-lisp-re-08 (work in progress), November 2015.
- [I-D.ietf-lisp-ddt]
Fuller, V., Lewis, D., Ermagan, V., Jain, A., and A. Smirnov, "LISP Delegated Database Tree", draft-ietf-lisp-ddt-08 (work in progress), September 2016.
- [I-D.ietf-lisp-lcaf]
Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", draft-ietf-lisp-lcaf-16 (work in progress), October 2016.
- [I-D.ietf-lisp-signal-free-multicast]
Moreno, V. and D. Farinacci, "Signal-Free LISP Multicast", draft-ietf-lisp-signal-free-multicast-01 (work in progress), April 2016.
- [I-D.ietf-nvo3-vxlan-gpe]
Kreeger, L. and U. Elzur, "Generic Protocol Extension for VXLAN", draft-ietf-nvo3-vxlan-gpe-02 (work in progress), April 2016.
- [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-02 (work in progress), August 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