

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: 22 July 2022

M. Portoles
V. Ashtaputre
F. Maino
Cisco Systems
V. Moreno
Google LLC
D. Farinacci
lispers.net
18 January 2022

LISP L2/L3 EID Mobility Using a Unified Control Plane
draft-ietf-lisp-eid-mobility-09

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 for End-point Identifier (EID) mobility, 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 <https://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on 22 July 2022.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
2. 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
4.1.2.1. L3 Mobility Flow using Data Driven SMRs	7
4.1.2.2. L3 Mobility Flow using Publish/Subscribe Mechanisms	8
4.2. Implementation Considerations	9
4.2.1. L3 Segmentation	9
4.2.2. L3 Database-Mappings	9
4.2.3. LISP Mapping System support	10
4.2.4. Using SMRs to Track Moved-Away Hosts	10
4.2.5. L3 multicast support	11
4.2.6. Time-to-Live Handling in Data-Plane	11
5. L2 Overlays and Mobility Support	11
5.1. Reference Architecture and packet flows	11
5.1.1. Bridged Traffic Flow: L2 Overlay use	12
5.1.2. L2 Mobility Flow	13
5.1.2.1. L2 Mobility Flow using Data Driven SMRs	13
5.1.2.2. L2 Mobility Flow using Publish/Subscribe mechanisms	14
5.2. Implementation Considerations	14
5.2.1. L2 Segmentation	15
5.2.2. L2 Database-Mappings	15
5.2.3. Interface to the LISP Mapping System	16
5.2.4. SMR support to track L2 hosts that moved away	16
5.2.5. L2 Broadcast and Multicast traffic	17
5.2.6. L2 Unknown Unicast Support	17
5.2.7. Time-to-Live Handling in Data-Plane	18

5.3.	Support to ARP resolution through Mapping System	18
5.3.1.	Map-Server support to ARP resolution: Packet Flow . .	18
5.3.2.	ARP registrations: MAC as a locator record	19
5.3.3.	Implementation Considerations	21
6.	Optional Deployment Models	22
6.1.	IP Forwarding of Intra-subnet Traffic	22
6.2.	Data-plane Encapsulation Options	23
7.	IANA Considerations	24
8.	Acknowledgements	24
9.	References	24
9.1.	Normative References	24
9.2.	Informative References	25
	Authors' Addresses	26

1. Introduction

This document describes the architecture and design options required to offer a unified L2 and L3 overlay solution for End-point Identifier (EID) 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 EID 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 underlay network.

Broadcast, Unknown and Multicast traffic on the overlay are supported by either replicated unicast, or underlay (RLOC) multicast as specified in [RFC6831] and [RFC8378].

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

The following figure illustrates the reference system used to support the packet flow description throughout this document. 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).

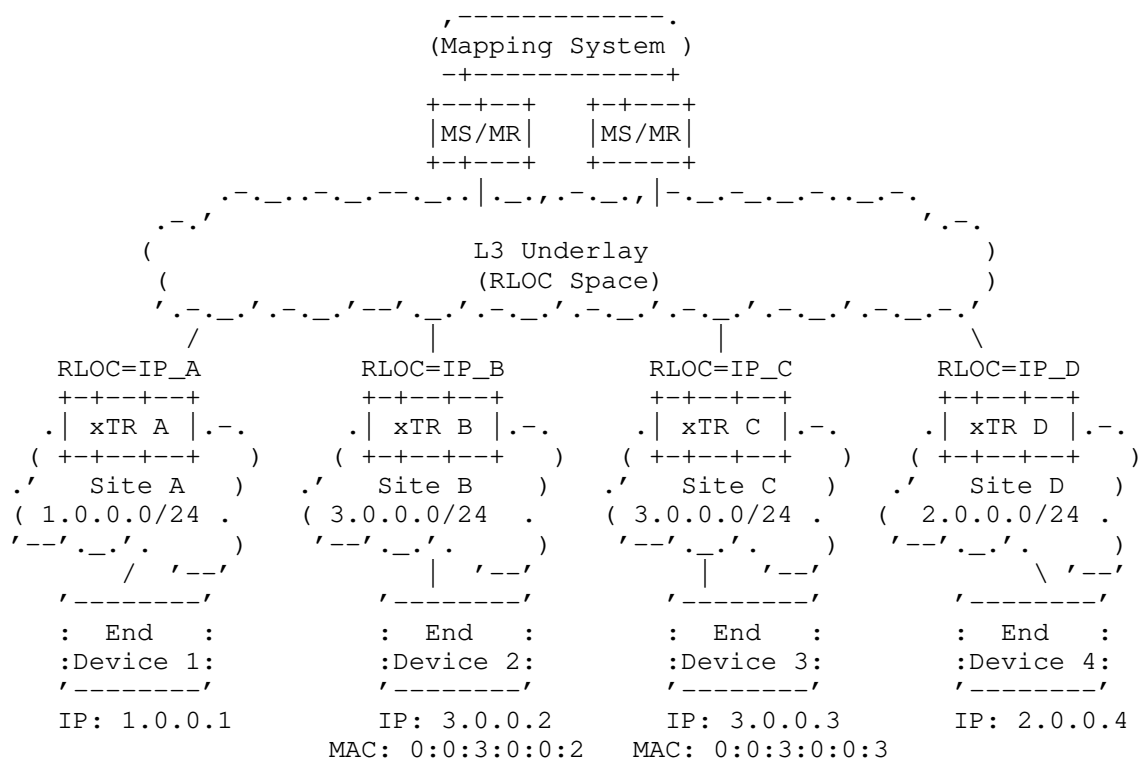


Figure 1: Reference System Architecture with unified L2 and L3 overlays

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:

- * EID = <IID, MAC> to RLOC=<IP>. This is used to support the L2 overlay.
- * EID = <IID, IP> to RLOC=<IP>. This is the traditional mapping as defined in the original LISP 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.

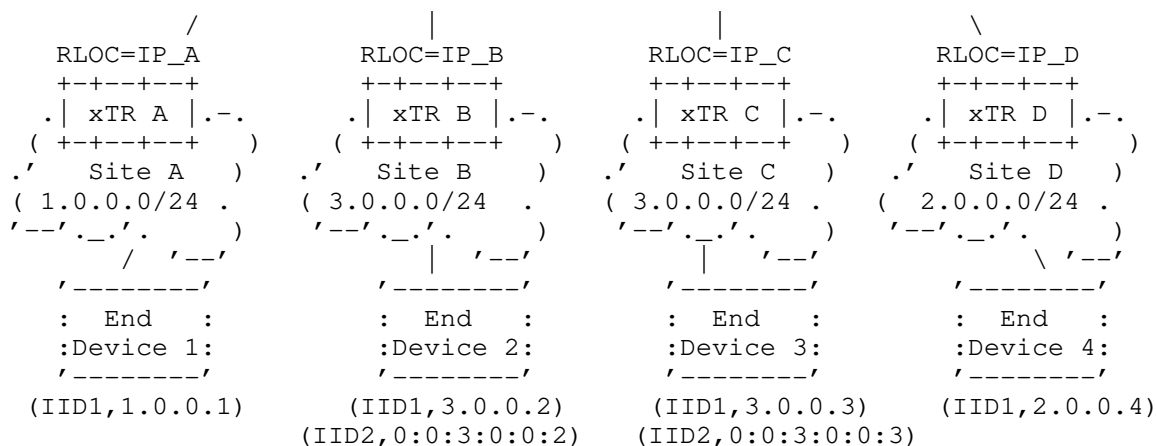


Figure 2: Reference Mobility Architecture

4.1.1. 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 [I-D.ietf-lisp-rfc6830bis] and [I-D.ietf-lisp-rfc6833bis]. 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).

- * 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).
- * 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>.
- * 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>.
- * 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.
- * 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.

4.1.2. 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 update of the ITR map-caches when EIDs move MAY be achieved using Data Driven SMRs or the Publish/Subscribe mechanisms defined in [I-D.ietf-lisp-pubsub]. The following two sections are sequence descriptions of the packet flow when End-Device 1 in the reference figure roams to site D.

4.1.2.1. L3 Mobility Flow using Data Driven SMRs

The following is a sequence description of the packet flow when End-Device 1 in the reference figure roams to site D. This sequence uses Data Driven SMRs to trigger the updates of the ITR map-caches.

- * 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 the departure ETR(s) (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>.
- * 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.
- * 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>.
- * 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.
- * 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.

4.1.2.2. L3 Mobility Flow using Publish/Subscribe Mechanisms

When Publish/Subscribe ([I-D.ietf-lisp-pubsub]) mechanisms are used, the flow of signaling to achieve EID mobility is modified. In this case, when an local end-device connected via an ITR establishes communication with a remote mobile end-device (connected to a remote ETR), the ITR will issue a Map-Request for the mobile end-device. Following the Mapping Request Subscribe Procedures defined in [I-D.ietf-lisp-pubsub], the Map-request will be issued with the N-bit set on the EID-Record so that the ITR is notified of any RLOC-set changes for the mobile EID-prefix.

The following is a sequence description of the packet flow when End-Device 1 in the reference figure roams to site D. This sequence leverages Publish/Subscribe mechanisms to update the ITR map-caches.

- * When an end-Device connected via an ITR establishes communication with a mobile end-device (e.g. end-device 1), the ITR will issue a Map-Request for the mobile end-device. Following the Mapping Request Subscribe Procedures defined in [I-D.ietf-lisp-pubsub], the Map-request will be issued with the N-bit set on the EID-Record so that the ITR is notified of any RLOC-set changes for the mobile EID-prefix.
- * When the mobile end-device (End-Device 1) is attached or detected in a new site (e.g. site D), The ETR at the new site (e.g. ETR D) sets up the database mapping corresponding to the EID of the mobile end-device (e.g. EID=<IID1, 1.0.0.1>). The ETR at the new site (e.g. ETR D) sends a Map-Register to the mapping system registering its RLOCs (e.g. RLOC=IP_D) as locator for the EID of the mobile end-device (e.g. EID=<IID1, 1.0.0.1>). Now the mapping system is updated with the new EID-to-RLOC mapping (location) for the mobile end-device (e.g. End-Device 1).
- * The Mapping System, after receiving the new registration for the EID of the mobile end-device (EID=<IID1, 1.0.0.1>) sends a Map-Notify to the departure site (ETR A) to inform it of the move. Then, the ETR at the departure site (ETR A) removes its local database mapping information and stops registering the EID for the mobile end-device (EID=<IID1, 1.0.0.1>).
- * Any ITR or PiTR participating in the L3 overlay (corresponding to IID1) that were sending traffic to the mobile end-device (1.0.0.1) would have Subscribed to receive notifications of any changes in the RLOC-set for the EID of the mobile end-device (1.0.0.1). The Mapping System publishes the updated RLOC-set to the Subscribed ITRs by sending a Map-Notify to the ITRs or PiTRs per the Mapping Notification Publish Procedures defined in [I-D.ietf-lisp-pubsub].

- * Upon receiving the Map-Notify message, the ITR updates the RLOC-set in its local map-cache entry for the EID of the mobile end-device (EID=<IID1, 1.0.0.1>). Once the map-cache is updated, traffic is tunneled by the ITR to the new location.

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 [RFC8060] and its use in [RFC8111].

Instance-IDs can be used to support L3 overlay segmentation, such as in extended VRFs or multi-VPN overlays ([I-D.ietf-lisp-vpn]).

4.2.2. L3 Database-Mappings

When an end-host is attached or detected in an ETR that provides L3-overlay services and mobility, a database Mapping is registered to the mapping system with the following structure:

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

The registration of these EIDs MUST follow the LCAF format as defined in [RFC8060] and the specific EID record to be used is 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 = 1 or 2
d +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
| EID-Prefix (IPv4 or IPv6)
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
/ | Priority | Weight | M Priority | M Weight
L +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
o | Unused Flags | L | p | R | Loc-AFI
c +-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
\ |
+-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+
+--> |-----+-----+-----+-----+-----+-----+-----+-----+-----+-----+

```

The L3 EID record follows the structure as described in [I-D.ietf-lisp-rfc6833bis].

4.2.3. LISP Mapping System support

The interface between the xTRs and the Mapping System is described in [I-D.ietf-lisp-rfc6833bis] 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 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 in [I-D.ietf-lisp-rfc6833bis].

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:

- * Keep track of end-hosts that were once connected to an ETR and have moved away.
- * 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], and [RFC8378].

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

The LISP specification ([I-D.ietf-lisp-rfc6830bis]) 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.

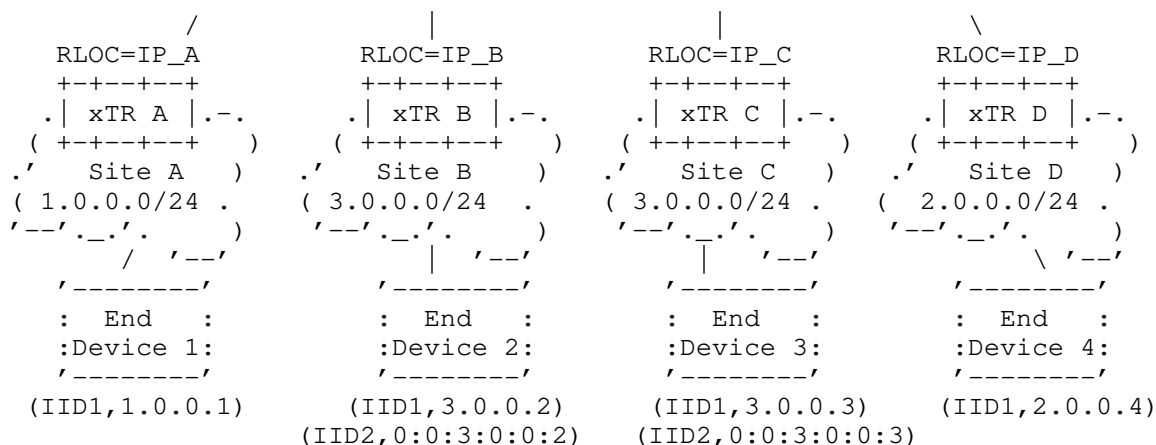


Figure 3: Reference Mobility Architecture

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).

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

- * 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 update of the ITR map-caches when EIDs move MAY be achieved using Data Driven SMRs or the Publish/Subscribe mechanisms defined in [I-D.ietf-lisp-pubsub]. The following two sections are sequence descriptions of the packet flow when End-Device 2 in the reference figure roams to site C, which is extending its own subnet network.

5.1.2.1. L2 Mobility Flow using Data Driven SMRs

The following is a sequence description of the packet flow when End-Device 2 in the reference figure roams to site C. This sequence uses Data Driven SMRs to trigger the updates of the ITR map-caches.

- * 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>.
- * 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>.
- * 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.
- * 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.

5.1.2.2. L2 Mobility Flow using Publish/Subscribe mechanisms

When Publish/Subscribe ([I-D.ietf-lisp-pubsub]) mechanisms are used, the flow of signaling to achieve EID mobility is modified. In this case, when an End-Device connected via an ITR establishes communication with a mobile EID-prefix, the ITR will issue a Map-Request for the mobile End-device. Following the Mapping Request Subscribe Procedures defined in [I-D.ietf-lisp-pubsub], the Map-request will be issued with the N-bit set on the EID-Record so that the ITR is notified of any RLOC-set changes for the mobile EID-prefix.

The following is a sequence description of the packet flow when End-Device 2 in the reference figure roams to site C. This sequence leverages Publish/Subscribe mechanisms to update the ITR map-caches.

- * 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>.
- * The Mapping System, after receiving the new registration for EID=<IID1, 0:0:3:0:0:2> sends a Map-Notify to the departure site (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>.
- * Any ITR or PiTR participating in the same L2-overlay (corresponding to IID2) that was encapsulating traffic to 0:0:3:0:0:2 before the migration would have Subscribed to receive notifications of any changes in the RLOC-set for 0:0:3:0:0:2. The Mapping System publishes the updated RLOC-set to the Subscribed ITRs by sending a Map-Notify to the ITRs or PiTRs per the Mapping Notification Publish Procedures defined in [I-D.ietf-lisp-pubsub].
- * Upon receiving the Map-Notify message, the ITR updates the RLOC-set in its local map-cache entry for EID=<IID2, 0:0:3:0:0:2>. Once the map-cache is updated, traffic is tunneled by the ITR 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 [RFC8060] and its use in [RFC8111]. 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:

- * 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 [RFC8060] 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 [I-D.ietf-lisp-rfc6833bis].

5.2.3. Interface to the LISP Mapping System

The interface between the xTRs and the Mapping System is described in [I-D.ietf-lisp-rfc6833bis] and this document follows the specification as described there. When available, the registrations MAY be implemented over a 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:

- * Keep track of end-hosts that were once connected to an ETR and have moved away.
- * 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 [RFC8378], 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.

Following the same case, when multicast is not available in the core network, the procedures in [RFC8378] 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,

* EID-record contents = <IID, IP>, RLOC-record contents <MAC>.

There is a dedicated IID used for the registration of the ARP/ND related mappings. Thus, a system with L2 and L3 overlays as well as ARP/ND 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:

* EID = <IID1, IP> to RLOC = <IP-RLOC> (L3-overlay)

- * EID = <IID2, MAC> to RLOC = <IP-RLOC> (L2-overlay)
- * 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:

- * 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 EID = <IID2,3.0.0.3>.
- * 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.
- * 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 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:

- * 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 [RFC8060]. 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 [I-D.ietf-lisp-rfc6833bis]. 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:

- * As indicated, when an end-host is attached the ETR maintains and registers a mapping with the binding EID = <IID, IP> -> RLOC = <MAC>.
- * ARP support through the LISP Mapping System is OPTIONAL and the xTRs should allow the possibility to enable or disable the feature.
- * 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.
- * 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).
- * 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.

- * 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):

- * End-Device 1 sends a broadcast ARP message to discover the MAC address of 3.0.0.3.
- * ITR B receives the ARP message and sends a Map-Request to the Mapping System for EID = <IID1,3.0.0.3>.
- * 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 EID = <IID1,3.0.0.3> -> 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 it is 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-lisp-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 LISP specification [I-D.ietf-lisp-rfc6830bis]. 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.

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, <<https://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, <<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>>.
- [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>>.
- [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, <<https://www.rfc-editor.org/info/rfc7348>>.
- [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>>.

- [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>>.
- [RFC8378] Moreno, V. and D. Farinacci, "Signal-Free Locator/ID Separation Protocol (LISP) Multicast", RFC 8378, DOI 10.17487/RFC8378, May 2018, <<https://www.rfc-editor.org/info/rfc8378>>.

9.2. Informative References

- [I-D.ietf-lisp-gpe]
Maino, F., Lemon, J., Agarwal, P., Lewis, D., and M. Smith, "LISP Generic Protocol Extension", Work in Progress, Internet-Draft, draft-ietf-lisp-gpe-19, 26 July 2020, <<https://www.ietf.org/archive/id/draft-ietf-lisp-gpe-19.txt>>.
- [I-D.ietf-lisp-pubsub]
Rodriguez-Natal, A., Ermagan, V., Cabellos, A., Barkai, S., and M. Boucadair, "Publish/Subscribe Functionality for LISP", Work in Progress, Internet-Draft, draft-ietf-lisp-pubsub-09, 28 June 2021, <<https://www.ietf.org/archive/id/draft-ietf-lisp-pubsub-09.txt>>.
- [I-D.ietf-lisp-rfc6830bis]
Farinacci, D., Fuller, V., Meyer, D., Lewis, D., and A. Cabellos, "The Locator/ID Separation Protocol (LISP)", Work in Progress, Internet-Draft, draft-ietf-lisp-rfc6830bis-36, 18 November 2020, <<https://www.ietf.org/archive/id/draft-ietf-lisp-rfc6830bis-36.txt>>.
- [I-D.ietf-lisp-rfc6833bis]
Farinacci, D., Maino, F., Fuller, V., and A. Cabellos, "Locator/ID Separation Protocol (LISP) Control-Plane", Work in Progress, Internet-Draft, draft-ietf-lisp-rfc6833bis-30, 18 November 2020, <<https://www.ietf.org/archive/id/draft-ietf-lisp-rfc6833bis-30.txt>>.
- [I-D.ietf-lisp-vpn]
Moreno, V. and D. Farinacci, "LISP Virtual Private Networks (VPNs)", Work in Progress, Internet-Draft, draft-ietf-lisp-vpn-08, 18 January 2022, <<https://www.ietf.org/archive/id/draft-ietf-lisp-vpn-08.txt>>.

[I-D.kouvelas-lisp-map-server-reliable-transport]

Leong, J., Lewis, D., Pitta, B., Cassar, C., Kouvelas, I.,
and J. Arango, "LISP Map Server Reliable Transport", Work
in Progress, Internet-Draft, draft-kouvelas-lisp-map-
server-reliable-transport-07, 18 January 2022,
<[https://www.ietf.org/archive/id/draft-kouvelas-lisp-map-
server-reliable-transport-07.txt](https://www.ietf.org/archive/id/draft-kouvelas-lisp-map-server-reliable-transport-07.txt)>.

Authors' Addresses

Marc Portoles Comeras
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
United States of America

Email: mportole@cisco.com

Vrushali Ashtaputre
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
United States of America

Email: vrushali@cisco.com

Fabio Maino
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
United States of America

Email: fmaino@cisco.com

Victor Moreno
Google LLC
1600 Amphitheatre Pkwy
Mountain View, CA 94043
United States of America

Email: vimoreno@google.com

Dino Farinacci
lispers.net
San Jose, CA
United States of America

Email: farinacci@gmail.com

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: July 22, 2022

V. Moreno
Google LLC
D. Farinacci
lispers.net
January 18, 2022

LISP Virtual Private Networks (VPNs)
draft-ietf-lisp-vpn-08

Abstract

This document describes the use of the Locator/ID Separation Protocol (LISP) to create Virtual Private Networks (VPNs). LISP is used to provide segmentation in both the LISP data plane and control plane. These VPNs can be created over the top of the Internet or over private transport networks, and can be implemented by Enterprises or Service Providers. The goal of these VPNs is to leverage the characteristics of LISP - routing scalability, simply expressed Ingress site TE Policy, IP Address Family traversal, and mobility, in ways that provide value to network operators.

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 <https://datatracker.ietf.org/drafts/current/>.

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

This Internet-Draft will expire on July 22, 2022.

Copyright Notice

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

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

Table of Contents

1. Introduction	3
2. Definition of Terms	3
3. LISP Virtual Private Networks (VPNs)	4
3.1. The LISP IID in the Control Plane	6
3.2. The LISP IID in the Data Plane	7
3.3. Locator Network Segmentation	7
3.4. Multicast in LISP VPN environments	8
4. LISP VPN Extranet	8
4.1. LISP Extranet VPN Control Plane	9
4.1.1. LISP Extranet VPN Map Register Procedures	9
4.1.2. LISP Extranet VPN Map Lookup Procedures	10
4.1.3. LISP Extranet Publish/Subscribe Procedures	11
4.1.4. LISP Extranet VPN Home-IID encoding	11
4.2. LISP Extranet VPN Data Plane	12
4.3. LISP Extranet VPN Multicast Considerations	12
4.3.1. LISP Extranet VPN Multicast Control Plane	12
4.3.2. LISP Extranet VPN Multicast Data Plane	13
4.4. LISP Extranet SMR Considerations	13
4.4.1. Home-IID inclusion in SMR messages	14
4.5. LISP Extranet RLOC Probing Considerations	14
5. Security Considerations	14
5.1. LISP VPNs and LISP Crypto	15
6. IANA Considerations	15
7. Acknowledgements	15
8. References	15
8.1. Normative References	15
8.2. Informative References	16
Authors' Addresses	17

1. Introduction

Network virtualization creates multiple, logically separated topologies across one common physical infrastructure. These logically separated topologies are known as Virtual Private Networks (VPNs) and are generally used to create closed groups of end-points. Network reachability within a VPN is restricted to the addresses of the end-points that are members of the VPN. This level of segmentation is useful in providing fault isolation, enforcing access-control restrictions, enabling the use of a single network by multiple tenants and scoping network policy per VPN.

LISP creates two namespaces: The End-point Identifier (EID) namespace and the Routing Locator (RLOC) namespace. The LISP Mapping System maps EIDs to RLOCs. Either the EID space, the RLOC space or both may be segmented. The LISP Mapping System can be used to map a segmented EID address space to the RLOC space. When the EID namespace is segmented, a LISP Instance-ID (IID) is encoded in both the data plane and the control plane to provide segmentation and to disambiguate overlapping EID Prefixes. This allows multiple VRFs to 'share' a common Routing Locator network while maintaining EID prefix segmentation.

LISP VPNs must support Multicast traffic in the EID space and must also support the ability to provide controlled reachability across VPNs which is commonly known as extranet functionality. When data path security is needed, LISP virtualization can be combined with LISP Crypto to provide data path confidentiality, integrity, origin authentication and anti-replay protection.

2. Definition of Terms

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

Terms defining interactions with the LISP Mapping System are defined in [RFC6833].

Terms related to the procedures for signal free multicast are defined in [RFC8378].

The following terms are here defined to facilitate the descriptions and discussions within this particular document.

Forwarding Context - Logical segment of a device's forwarding table and its associated interfaces. This is usually in the form of a VRF

for IP forwarding, may also be in the form of a Bridge Domain or VLAN for MAC forwarding.

Home-IID - In the context of cross VPN connectivity, a particular EID will be registered with multiple Instance-IDs, the Home-IID identifies the Instance-ID associated to the Forwarding Context (VRF) to which an EID is actually connected.

Extranet-VPN - In the context of cross VPN connectivity, a VPN that is reachable by all Extranet-Subscriber-VPNs and can reach all Extranet-Subscriber-VPNs.

Extranet-Subscriber-VPN - The VPNs that can reach the Extranet-VPN, but cannot reach each other.

Extranet Policy - The definition of which VPNs share reachability information with each other in the context of cross VPN connectivity. May be structured as a group of Extranet-Subscriber-VPNs that subscribe to an Extranet-VPN.

3. LISP Virtual Private Networks (VPNs)

A LISP VPN is a collection of LISP Sites building an Overlay Network. These sites share a common control plane, the LISP Mapping System. The members of this VPN also share common RLOC connectivity, whether it be the Internet or a private IP network.

Multiple LISP VPNs may run over a common RLOC space and many LISP VPNs may share one or more locations, requiring XTRs to service multiple VPNs simultaneously.

VPNs must be allowed to have overlapping address space. It is necessary to disambiguate the EID namespace in both the control and data plane as well as maintain forwarding segmentation within the XTRs. The LISP Instance ID (IID) is used to provide a VPN wide unique identifier that can be used both in the control and data planes.

The LISP Instance ID is a 32 bit unstructured namespace that identifies a LISP VPN. The tuple of EID Prefix and IID is referred to as an Extended EID (XEID) [RFC8111]. The LISP IID is used in the data plane of the LISP header [RFC6830], as well as in the LISP control plane [RFC8060].

An implementation should default to an Instance ID value equal to zero when LISP VPNs are not in use.

The operation of a LISP VPN is consistent with the operation of LISP in a non-VPN environment as defined in [RFC6830]. The operation of a LISP VPN is here described at a high level in terms of EID registrations, EID lookups and traffic forwarding:

EID registration: In a LISP VPN, XTRs that are members of the VPN should be configured with a forwarding context (e.g. VRF) and the associated IID for the VPN. Based on this configuration, the ETRs must register the EIDs within the forwarding context as Extended EIDs (IID+EID). The LISP mapping system consolidates the registrations from all the ETRs in the VPN and builds a mapping database for the VPN.

EID Lookup: ITRs that are members of the VPN will do forwarding lookups in the forwarding context where traffic was received. Upon a cache miss within the forwarding context, the ITR must issue a Map-Request for the destination EID and include the VPN's IID. This information must be encoded as an Extended EID (IID+EID) in the Map-Request issued. The IID to associate with the EID in this Map-request is derived from the configuration of the VPN's forwarding context (in which the traffic was received). The Mapping System should reply to the Map Request with a Mapping for the Extended EID (IID+EID), the IID of the Extended EID should be used to identify the forwarding context in which the Mapping received should be cached.

Traffic Forwarding: Once a Mapping has been cached in the VPN's forwarding context, the ITR will encapsulate the traffic towards the RLOC in the mapping. The IID corresponding to the VPN's forwarding context must be included in the Instance-ID field of the data plane header. When the encapsulated traffic is received at the ETR the encapsulation header is removed and the IID received in the header is used to identify the forwarding context to use to do a forwarding lookup for the decapsulated traffic.

A more formal description of the Control and Data Plane procedures for a LISP VPN is documented in the following sections.

In order to create VPNs, the following segmentation functions must be provided:

- o **Device Segmentation.** The forwarding tables of the devices must be segmented so that independent forwarding decisions can be made for each virtual network. Virtual Routing and Forwarding (VRF) contexts may be used to create multiple instances of Layer 3 routing tables virtualization (segmentation) at the device level. If the EID space is in a Layer 2 address family (e.g. MAC addresses), then Layer 2 contexts such as VLANs or bridge domains

may be used to segment the device. We generalize the concept of separate forwarding tables as forwarding contexts.

- o Data Plane Segmentation. Data Plane Forwarding separation is necessary for the devices to maintain virtual network semantics at forwarding time. Data plane separation can be maintained across network paths using either single-hop path segmentation (hop-by-hop) or multi-hop path segmentation. Single-hop path segmentation mechanisms include constructs such as 802.1q VLAN trunks, multi-hop mechanisms include MPLS, LISP, VXLAN and GRE tunnels.
- o Control Plane Segmentation. In order to correctly populate the multiple forwarding tables in the segmented network devices, the control plane needs to be segmented so that the different updates that are conveyed by the control plane contain the necessary virtual network semantics to discriminate between information relevant to one segment vs another. Control plane segmentation is key to allowing sites to use overlapping network prefixes in these logically separate topologies. BGP/MPLS VPNs (ref RFC 4364) are an example of this control plane segmentation.

3.1. The LISP IID in the Control Plane

In a LISP Mapping System supporting VPNs, EID Prefixes should be registered as Extended EID tuples of information that include the EID prefix as well as its corresponding Instance ID (IID) information.

In a segmented LISP network, whenever an EID is present in a LISP message, the EID must be encoded as an extended EID using the Instance ID LCAF type defined in [RFC8060]. This includes all LISP messages pertinent to the EIDs in the segmented space, including, but not limited to, Map-Register, Map-Request, Map-Reply, Map-Notify, SMRs, etc.

On EID registration by an ETR, the Map-Register message sent by the ETR must contain the corresponding IID encoded as part of the EID using the Instance ID LCAF type.

On EID lookup, when an ITR issues a Map-Request, both the Map-Request message and the resulting Map-Reply must contain the IID for the EID encoded using the IID LCAF type. The IID to use for a Map-Request may be derived from the configuration of the ITR Ingress VRF. The mappings received by an ITR in a Map-Reply should be cached in the VRF corresponding (by configuration) to the IID included in the Map-Reply message.

The Mapping System must maintain the IID information that corresponds to any EIDs actively registered with the Mapping System.

3.2. The LISP IID in the Data Plane

A LISP xTR will map, by configuration, a LISP Instance ID to a given forwarding context in its EID namespace. The Instance-ID must be included in the data plane header to allow an xTR to identify which VPN the packet belongs to when encapsulating or decapsulating LISP packets. The LISP header [RFC6830] as well as the VXLAN header [RFC7348] reserve a 24 bit field for the purposes of encoding the Instance-ID (referred to as VNID in the VXLAN specification).

LISP ITRs may receive non-encapsulated traffic on an interface that is associated with the forwarding context for a VPN (e.g. VRF). A LISP ITR should do Map-cache lookups for the destination EID within the forwarding context in which it received the traffic. The LISP ITR must encapsulate the traffic to the destination RLOC found in the map-cache and must include, in the header of the encapsulated packet, the IID associated with the forwarding context for the VPN. In the event of a map-cache miss, the LISP ITR must issue a Map-request with the IID associated with the ITR Ingress VRF as described in Section 3.1.

On receipt of an encapsulated LISP packet, a LISP ETR will deliver the decapsulated packets to the VRF associated with the IID received in the LISP header. Standard routing lookups will then take place within the context of the VRF for the forwarding of the decapsulated packet towards its destination.

The use of multiple IIDs on a single site xTR, each mapped to a different EID VRF allows for multiplexing of VPNs over a Locator network.

3.3. Locator Network Segmentation

This document has so far discussed virtualizing the LISP EID namespace, and communication between xTRs and the LISP Mapping System. Implicit in this communication requirement is a network between these devices. LISP VPNs do not require this underlay network connectivity to be in the "default" VRF, just that a given LISP Site and its Mapping System be interconnected via a common VRF.

LISP xTRs may have connectivity to each other via multiple distinct VRFs, as in the case where the LISP VPN is being used to create an Overlay with multiple MPLS-VPN Service Providers being used as the transport. In other words, the RLOC space may also be segmented, the segmentation of the RLOC space is not done by LISP, but the segmentation of the RLOC space is delivered by the routing protocols and data plane used by the RLOC space. When the RLOC space is segmented, different EID segments may use different RLOC segments.

An RLOC segment may service one or many EID segments, allowing a VPN in the RLOC space to service a subset of the VPNs created in the EID space.

3.4. Multicast in LISP VPN environments

Both Signaled and Signal Free Multicast within a VPN will operate without modification in VPN environments provided that all LISP control plane messages include the Instance ID for their VPN as specified in Section 3. Multicast Source (S) state as well as multicast Group (G) state are both scoped within a VPN and therefore the values for S and G may be reused in other VPNs.

4. LISP VPN Extranet

In a multi-tenant network the communication between a shared VPN and a multitude of otherwise isolated VPNs is generally known as extranet communication. Reachability is established between an shared Extranet-VPN and a multitude of Extranet-Subscriber-VPNs without enabling reachability between the different Extranet-Subscriber-VPNs. This section specifies the procedures and protocol encodings necessary to provide extranet functionality in a multi-instance LISP network. The mechanisms described require cross VPN lookups and therefore assume that the EID space across all VPNs involved does not overlap or has been translated to a normalized space that resolves any overlaps.

The operation of a LISP VPN Extranet is consistent with the operation of LISP VPNs as defined in Section 3. The operation of a LISP VPN Extranet is here described at a high level in terms of EID registrations, EID lookups and traffic forwarding:

EID Registration: EIDs in the Extranet-VPN should be registered in their Home-IID as well as in all other IIDs that are part of the Extranet scope. EIDs in the Extranet-Subscriber-VPNs should be registered in their Home-IID and the Extranet-VPN's IID. This makes the EIDs available for lookups in VPNs other than their Home-VPN. When an EID is registered in an IID that it does not belong to, the mapping should include a parameter containing the Home-IID for the EID. As a result any EID that should be reachable based on the Extranet configuration will be registered in every relevant VPN, if the EID is not native to that VPN, the mapping will have a parameter with the Home-IID for the EID.

EID Lookup: Map-requests will be issued within the IID of the requesting VPN as specified in Section 3. If the destination is across VPNs, the mapping for the destination EID should contain the EID's Home-IID as a parameter. The mapping, including the Home-IID

parameter is returned in a Map-Reply and cached by the ITR in the Forwarding Context of the requesting VPN. The cache will include the destination's Home-IID as a parameter of the mapping.

Traffic Forwarding: An ITR will encapsulate traffic to a cross VPN destination using the destination's Home-IID in the data plane header. Upon decapsulation at the ETR, traffic is handed directly to the destination VPN's forwarding context based on the IID used in the header.

A more formal description of the Control and Data Plane procedures for a LISP VPN Extranet is documented in the following sections.

4.1. LISP Extranet VPN Control Plane

In order to achieve reachability across VPNs, EID mapping entries in the Extranet-VPN must be accessible for lookups initiated from an Extranet-Subscriber-VPN and vice-versa.

The definition of which VPNs share reachability information is governed by configurable Extranet Policy. The Extranet Policy will simply state which VPNs are extranet-subscribers to a particular extranet-VPN. There may be multiple Extranet-VPNs in a LISP network and a VPN may subscribe to multiple Extranet-VPNs. An Extranet-subscriber-VPN may act as an Extranet-VPN to provide reachability across Extranet-subscriber-VPNs, this effectively merges the Extranet-subscriber-VPNs together, a scenario that is usually better achieved by creating a single Extranet-subscriber-VPN.

The Instance-ID (IID) for the VPN to which an EID is connected is referred to as the Home-IID of the EID. As cross VPN registrations and lookups take place, the Home-IID for an EID must be preserved and communicated in any pertinent LISP messages.

4.1.1. LISP Extranet VPN Map Register Procedures

An ETR may register EIDs in their Home-IID as well as in the other IIDs within the scope of the Extranet Policy. For example, an EID connected to the Extranet-VPN may be registered by its ETR in its Home-IID and also in all the IIDs corresponding to the Extranet-Subscriber-VPNs defined in the Extranet Policy. When Map-Register messages for an EID are issued in IIDs other than the EID's Home-IID, the Home-IID for the EID must be included in the Map-Register. The Home-IID must be encoded as described in Section 4.1.4.

When registering an EID in multiple IIDs, it is advisable to pack the multiple registrations in a single Map-Register message containing the multiple XEID records.

A Map-Server may be configured with the Extranet Policy. This may suffice for the Map-Server to be able to satisfy cross VPN lookups. In such implementations, ETRs may not be required to register an EID across the entire scope of IIDs defined in the Extranet Policy, but may only require the registration of the EID in its Home-IID.

Which method of cross VPN mapping registration is used (initiated by the ETR or initiated by the Map-Server) should be a configurable option on the XTRs and Map-Server.

4.1.2. LISP Extranet VPN Map Lookup Procedures

Map-Request messages issued by an ITR, their structure and use do not change when a destination EID is outside of the Home-IID for the source EID.

When a Map-Request message is forwarded from the Map-Resolver to an authoritative Map-Server (either directly or by DDT delegation), the IID of the requesting EID must be preserved so that the Map-Reply is sent in the correct context.

Map-Reply messages must use the IID of the requesting EID and must also include the Home-IID of the destination EID. The Home-IID is a parameter of the destination EID, part of the mapping and must be encoded as described in Section 4.1.4. The mapping obtained in the Map-Reply must be cached in the forwarding context of the requesting EID, which is identified by the IID for the requesting EID. The mappings cached will contain the Home-IID of the destination EID whenever this destination EID is cached outside of its Home-IID.

Since each IID at the Map Server has a complete set of EIDs in the scope of the extranet policy, the determination of a covering prefix in the case of a non-LISP or external destination is straightforward and follows the procedures delineated in the procedures for a negative map reply in [RFC6833]. When the Map Server determines that the requested destination EID is either not an EID or not registered it must calculate the covering prefix for the requested EID and reply in one of two ways:

- With a Negative Map Reply per the procedures outlined in [RFC6833]. If using a PeTR, the Home-IID for the PeTR must be configured at the requesting ITR.
- With a Map Reply mapping the calculated EID covering prefix to the RLOCs of a configured or registered PeTR. The Map Reply must contain the Home-IID of the registered PeTR.

4.1.3. LISP Extranet Publish/Subscribe Procedures

When LISP Extranet VPNs are implemented together with LISP Publish/Subscribe functionality [I-D.ietf-lisp-pubsub], the following considerations apply.

Subscriptions initiated across VPNs MUST maintain a record of the IID from which the subscription was requested. An ITR that issues a Map-Request with the N-bit set from an Extranet-Subscriber-VPN will use the IID of the Extranet-Subscriber-VPN as the IID for the XEID it subscribes to. The Map-Request is routed to the Extranet-VPN (Home-VPN) where the EID is registered, as defined in Section 4.1.2. The subscription is maintained in the Home-VPN and will include a record of the IID of the Extranet-Subscriber-VPN from which the subscription was initiated.

Any changes in the RLOC set for the EID will be published using a Map-Notify message. The Map-Notify message will include the Extranet-Subscriber-VPN IID in the XEID and it will also include the IID of the Home-VPN (Home-IID) encoded as specified in Section 4.1.4.

4.1.4. LISP Extranet VPN Home-IID encoding

The Home-IID is an attribute of the EID-RLOC mapping. The Home-IID must be encoded as an additional RLOC within the record carried in Map-Reply messages as defined in [RFC6830]. The Home-IID should also be included in Map-Notify messages when LISP Extranet VPNs are implemented together with LISP Publish/Subscribe functionality [I-D.ietf-lisp-pubsub].

The additional RLOC containing the Home-IID should use AFI = 16387 (LCAF) with a List type as described in Section 4.1.4.1.

4.1.4.1. Home-IID encoded in LCAF List type

The Home-IID may be encoded as LCAF AFI of type Instance ID (Type 2). The IID LCAF AFI entry should be nested within a List Type LCAF (Type 1). The list type is used to include a distinguished name type that would provide the semantical information that identifies this field as a Home-IID to be used for the purposes of Extranet VPNs. Map-Servers and XTRs receiving the encoded messages would leverage the semantical information to parse the control plane message properly. The different LCAF types are documented in [RFC8060]. The logical structure of the nested LCAF structure is depicted below:

```
AFI = LCAF(16387)
Type = LIST(1)
  ITEM1
    AFI = Distinguished Name
    Value = "Home-IID"
  ITEM2
    AFI = LCAF(16387)
    Type = IID(2)
    Value = <Home-IID.value>
```

4.1.4.2. Home-IID encoded in dedicated LCAF Type

Alternatively, a new dedicated LCAF type could be used in order to include application semantics to the encoding of the IID in a purposely structured type. In the future, this document may be updated to provide details of the definition of structure and semantics for a dedicated LCAF type to be used in this application.

4.2. LISP Extranet VPN Data Plane

Traffic will be forwarded according to the procedures outlined in [RFC6830]. The map-cache will include the Home-IID for the destination EID as part of the mapping for the destination EID. In an ITR, unicast traffic will be encapsulated using the Home-IID for the destination EID as the Instance-ID in the encapsulation header. On de-capsulation, the Instance-ID in the header points to the destination VPN already so no further procedures are required.

4.3. LISP Extranet VPN Multicast Considerations

When Multicast traffic needs to be forwarded across VPNs, there are special considerations that are closely tied to the definition of the Extranet functionality. This specification will focus on the use of Signal Free Multicast [RFC8378] for the delivery of a cross VPN multicast service.

4.3.1. LISP Extranet VPN Multicast Control Plane

The Receiver-site Registration procedures described in [RFC8378] are expanded to allow the formation of a replication-list inclusive of Receivers detected in the different VPNs within the scope of the Extranet Policy.

Once the Receiver-ETRs detect the presence of Receivers at the Receiver-site, the Receiver-ETRs will issue Map-Register messages to include the Receiver-ETR RLOCs in the replication-list for the multicast-entry the Receivers joined.

The encodings for Map-Register messages and the EIDs and RLOCs within follow the guidelines defined in [RFC8378].

For VPNs within the scope of the Extranet Policy the multicast receiver registrations will be used to build a common replication list across all VPNs in the Extranet Policy scope. This replication list is maintained within the scope of the VPN where the multicast source resides. When Receivers are in the Extranet-Subscriber-VPN, Multicast sources are assumed to be in the Extranet-VPN and viceversa.

The Instance-ID used to Register the Receiver-ETR RLOCs in the replication-list is the Instance-ID of the Extranet-VPN, i.e. the VPN where the Multicast Source resides. When listeners are detected in the Extranet-VPN, then multiple Registrations must be sent with the Instance-IDs of the Extranet-Subscriber-VPNs under the assumption that the Multicast sources could be in one or more of the Extranet-Subscriber-VPNs.

Source-ITRs will complete lookups for the replication-list of a particular multicast group destination as well as the forwarding of traffic to this multicast group following the procedures defined in [RFC8378] without any change.

4.3.2. LISP Extranet VPN Multicast Data Plane

It is desirable to send a single copy of the Multicast traffic over the transit network and have the Receiver-ETRs locally replicate the traffic to all Receiver-VPNs necessary. This replication is governed by the Extranet Policy configured at the ETR. Thus, ITRs will encapsulate the traffic with the Instance-ID for the VPN where the Multicast Source resides. ETRs will receive traffic in the source IID and replicate it to the Receiver VPNs per the Extranet Policy.

4.4. LISP Extranet SMR Considerations

Data driven SMRs MUST carry the IID for the VPNs of the receivers of traffic. Data driven SMRs MAY carry the IID for the VPNs of senders of traffic if the sender VPN IIDs are known by the ETR generating the SMR. If the sender VPN's Instance-ID is not known, the ETR SHOULD send the SMR to the RLOC of the sending ITR without the sender VPN's IID.

The SMR MUST be replicated to all extranet VPNs that are defined in the Extranet Policy and instantiated at the sending ITR.

When the IID of the sender VPN is known at the ETR, the ETR MAY include the sender VPN's IID in the SMR and issue a separate SMR for

each sender VPN IID known to the ETR. Multicast optimizations could be used to minimize the amount of traffic replicated when sending these SMRs and potentially replicate only at the ITR.

When the IID of the sender VPN is not known at the ETR, the ETR SHOULD issue a single SMR to each of the sending ITRs. The SMR will then be replicated at the ITR to all extranet VPNs that are defined in the Extranet Policy and instantiated at the sending ITR.

4.4.1. Home-IID inclusion in SMR messages

The Instance IDs relevant to the SMR signaling will be encoded in the SMR Map-request message fields as follows:

Source-EID field: If known by the ETR, this field SHOULD carry the instance-ID of the traffic source VPN at the ITR with the obsolete map-cache. This is the IID of the senders of the traffic. Otherwise, the Instance-ID in this field MUST be the same as the Instance-ID of the destination VPN at the ETR generating the SMR.

EID-Prefix field: This field carries the Instance-ID of the destination VPN at the ETR sending the SMR. This is the IID of the receivers of the traffic. This field must always be set with the IID of the receivers.

4.5. LISP Extranet RLOC Probing Considerations

RLOC Probes must be sent with the IID of the VPN originating the probe. The XTR receiving the probe must identify the VPN for the target EID. The XTR receiving the probe should run all verifications as specified in [RFC6830] within the forwarding context corresponding to the VPN where the target EID is connected. Once verifications are completed, the reply to the probe should be sent in the IID of the VPN that originated the probe.

5. Security Considerations

LISP [RFC 6830] incorporates many security mechanisms as part of the mapping database service when using control-plane procedures for obtaining EID-to-RLOC mappings. In general, data plane mechanisms are not of primary concern for general Internet use-case. However, when LISP VPNs are deployed, several additional security mechanisms and considerations must be addressed.

Data plane traffic uses the LISP instance-id (IID) header field for segmentation. in-flight modifications of this IID value could result in violations to the tenant segmentation provided by the IID. Protection against this attack can be achieved by using the integrity

protection mechanisms afforded by LISP Crypto, with or without encryption depending on users' confidentiality requirements (see below).

5.1. LISP VPNs and LISP Crypto

The procedures for data plane confidentiality in LISP are documented in [RFC8061] and are primarily aimed at negotiating secret shared keys between ITR and ETR in map-request and map-reply messages. These secret shared keys are negotiated on a per RLOC basis and without regard for any VPN segmentation done in the EID space. Thus, multiple VPNs using a shared RLOC may also share a common secret key to encrypt communications of the multiple VPNs.

It is possible to negotiate secret shared keys on a per EID basis by applying the procedures described in [RFC8061] to RLOC probes. In a VPN environment, RLOC probes would be aimed at Extended EIDs that contain Instance-ID semantics, therefore resulting in the calculation of different secret shared keys for different XEID. Since the keys are calculated per XEID prefix rather than per VPN, there are scale considerations when implementing this level of key negotiation granularity.

6. IANA Considerations

This document has no IANA implications

7. Acknowledgements

The authors want to thank Marc Portoles, Vrushali Ashtaputre, Johnson Leong, Jesus Arango, Prakash Jain, Sanjay Hooda, Alberto Rodriguez-Natal, Darrel Lewis and Greg Schudel for their insightful contribution to shaping the ideas in this document.

8. References

8.1. 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, <<https://www.rfc-editor.org/info/rfc2119>>.
- [RFC3618] Fenner, B., Ed. and D. Meyer, Ed., "Multicast Source Discovery Protocol (MSDP)", RFC 3618, DOI 10.17487/RFC3618, October 2003, <<https://www.rfc-editor.org/info/rfc3618>>.

- [RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, DOI 10.17487/RFC4601, August 2006, <<https://www.rfc-editor.org/info/rfc4601>>.
- [RFC4607] Holbrook, H. and B. Cain, "Source-Specific Multicast for IP", RFC 4607, DOI 10.17487/RFC4607, August 2006, <<https://www.rfc-editor.org/info/rfc4607>>.

8.2. Informative References

- [I-D.ietf-lisp-pubsub]
Rodriguez-Natal, A., Ermagan, V., Cabellos, A., Barkai, S., and M. Boucadair, "Publish/Subscribe Functionality for LISP", draft-ietf-lisp-pubsub-09 (work in progress), June 2021.
- [RFC6407] Weis, B., Rowles, S., and T. Hardjono, "The Group Domain of Interpretation", RFC 6407, DOI 10.17487/RFC6407, October 2011, <<https://www.rfc-editor.org/info/rfc6407>>.
- [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>>.
- [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>>.
- [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, <<https://www.rfc-editor.org/info/rfc7348>>.
- [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>>.
- [RFC8378] Moreno, V. and D. Farinacci, "Signal-Free Locator/ID Separation Protocol (LISP) Multicast", RFC 8378, DOI 10.17487/RFC8378, May 2018, <<https://www.rfc-editor.org/info/rfc8378>>.

Authors' Addresses

Victor Moreno
Google LLC
1600 Amphitheater Parkway
Mountain View, CA 94043
USA

Email: vimoreno@googleo.com

Dino Farinacci
lispers.net
San Jose, CA 95120
USA

Email: farinacci@gmail.com

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: 22 July 2022

J. Leong
D. Lewis
B. Pitta
Cisco Systems
C. Cassar
Tesla
I. Kouvelas
Arista Networks Inc.
J. Arango
Microsoft
18 January 2022

LISP Map Server Reliable Transport
draft-kouvelas-lisp-map-server-reliable-transport-07

Abstract

The communication between LISP ETRs and Map-Servers is based on unreliable UDP message exchange coupled with periodic message transmission in order to maintain soft state. The drawback of periodic messaging is the constant load imposed on both the ETR and the Map-Server. New use cases for LISP have increased the amount of state that needs to be communicated with requirements that are not satisfied by the current mechanism. This document introduces the use of a reliable transport for ETR to Map-Server communication in order to eliminate the periodic messaging overhead, while providing reliability, flow-control and endpoint liveness detection.

Status of This Memo

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

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

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

This Internet-Draft will expire on 22 July 2022.

Copyright Notice

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

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

Table of Contents

1. Introduction	2
2. Requirements Notation	3
3. Message Format	4
4. Session Establishment	5
5. Error Notifications	5
6. EID Prefix Registration	7
6.1. Reliable Mapping Registration Messages	7
6.1.1. Registration Message	8
6.1.2. Registration Acknowledgement Message	8
6.1.3. Registration Rejection Message	9
6.1.4. Registration Refresh Message	10
6.1.5. Mapping Notification Message	12
6.2. ETR Behavior	13
6.3. Map-Server Behavior	17
7. Security Considerations	18
8. IANA Considerations	18
8.1. LISP Reliable Transport Message Types	18
8.2. Transport Protocol Port Numbers	18
9. Acknowledgments	19
10. Normative References	19
Authors' Addresses	19

1. Introduction

The communication channel between LISP ETRs and Map-Servers is based on unreliable UDP message exchange [I-D.ietf-lisp-rfc6833bis]. Where required, reliability is pursued through periodic retransmissions that maintain soft state on the peer. Map-Register messages are retransmitted every minute by an ETR and the Map-Server times out its state if the state is not refreshed for three successive periods. When registering multiple EID-Prefixes, the ETR includes multiple mapping records in the Map-Register message. Packet size limitations

provide an upper bound to the number of mapping records that can be placed in each Map-Register message. When the ETR has more EID-Prefixes to register than can be packed in a single Map-Register message, the mapping records for the EID-Prefixes are split across multiple Map-Register messages.

The drawback of the periodic registration is the constant load that it introduces on both the ETR and the Map-Server. The ETR uses resources to periodically build and transmit the Map-Register messages, and to process the resulting Map-Notify messages issued by the Map-Server. The Map-Server uses resources to process the received Map-Register messages, update the corresponding registration state, and build and transmit the matching Map-Notify messages. When the number of EID-Prefixes to be registered by an ETR is small, the resulting load imposed by periodic registrations may not be significant. The ETR will only transmit a single Map-Register message each period that contains a small number of mapping records.

In some LISP deployments, a large set of EID-Prefixes must be registered by each ETR (e.g. mobility, database redistribution). Use cases with a large set of EID-Prefixes behind an ETR will result in a much higher load. An example is LISP mobility deployments where EID-Prefixes are limited to host entries. ETRs may have thousands of hosts to register resulting in hundreds of Map-Register and Map-Notify messages per registration period.

A transport is required for the ETR to Map-Server communication that provides reliability, flow-control and endpoint liveness notifications. This document describes the use of TCP or SCTP as a LISP reliable transport. The initial application for the LISP reliable transport session is the support of scalable EID prefix registration. The reliable session mechanism is defined to be extensible so that it can support additional LISP communication requirements as they arise using a single reliable transport session between an ETR and a Map-Server. The use of the reliable transport session for EID prefix registration is an alternative and does not replace the existing UDP based mechanism.

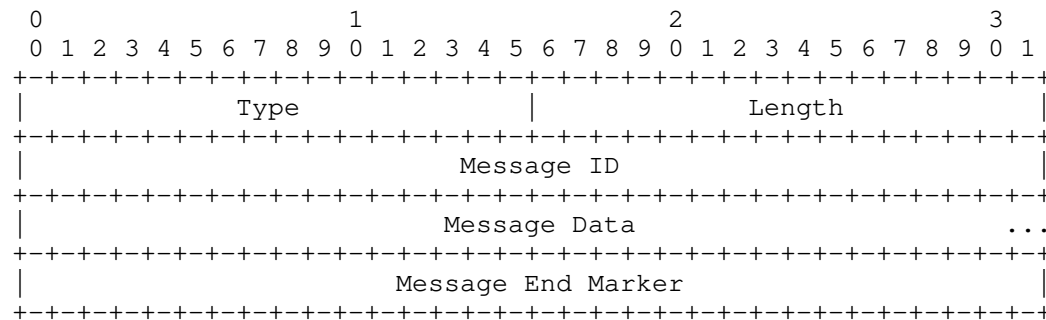
2. Requirements Notation

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].

3. Message Format

A single LISP reliable transport session may carry information for multiple LISP applications. One such application is the registration of EID to RLOC mappings that operates over a session between an ETR and a Map-Server. Communication over a session is based on the exchange of messages. This document defines a base set of messages to support session establishment and management. It also defines the messages for the EID to RLOC mapping registration application.

To support protocol extensibility when new applications, or extensions to existing applications are introduced, the messages are based on a TLV format.



Reliable transport message format

- * Type: 16 bit type field identifying the message type.
- * Length: 16 bit field that provides the total size of the message in octets including the length, type and end marker fields. The length allows the receiver to locate the next message in the TCP stream. The minimum value of the length field is 8.
- * ID: A 32-bit value that identifies the message. May be used by the receiver to identify the message in replies or notification messages.
- * Data: Type specific message contents.
- * End Marker: A 32-bit message end marker that must be set to 0x9FACADE9. The End Marker is used by the receiver to validate that it has correctly parsed or skipped a message and provides a method to detect formatting errors. Note that message data may also contain this marker, and that the marker itself is not sufficient for parsing the message.

The base message format does not indicate how the peer should deal with the message in cases where the message type is not supported/understood. This is best dealt with by the application. For example, in case an error notification is returned, or an expected acknowledgement message is not received, the application might choose various courses of action; from simply logging that the feature is not supported, all the way to tearing the relationship with the peer down for the feature, or for all LISP features.

4. Session Establishment

To ensure backwards compatibility, the map server and ETR MUST communicate via unreliable UDP messages until a TCP session between the two is successfully established.

The map server authenticates the ETR with the authentication data contained in the first UDP map-register message it receives from the ETR. Once the ETR is authenticated, the map server performs a passive open by listening on TCP port 4342, and does not qualify the remote port. As a security measure, the map server accepts TCP connections only from those ETRs that have been authenticated via UDP map-register messages.

The ETR assumes the active role of the TCP session establishment by connecting to the map server once it has received a UDP map-notify message.

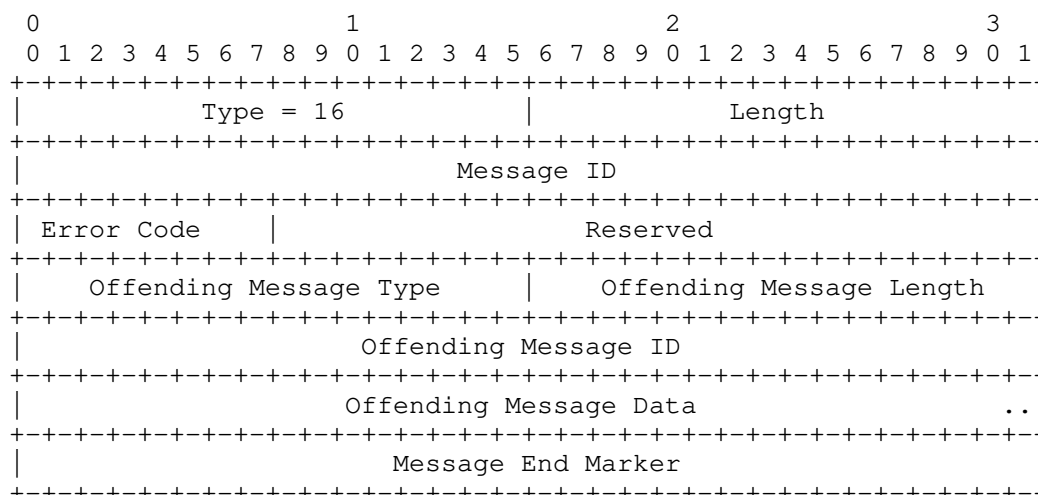
When a TCP session goes down, UDP authentication must take place before a new TCP session is established. The map-server will not accept a connection from the ETR until a UDP map-register has been received. Similarly, the ETR will not attempt to establish a session with the map server until an UDP map-notify message has been received.

A single reliable transport session is established between the map server and the ETR to cover all communication needs. For example, an ETR that has EID prefix registrations for multiple EID instances and EID address families will only establish a single session with the map server.

5. Error Notifications

The error notification message is used to communicate base reliable transport session communication errors. LISP applications making use of the reliable transport session and having to communicate application specific errors must define their own messages to do so. An error notification is issued when the receiver of a message does not recognize the message type or cannot parse the message contents.

The notification includes the offending message type and ID and as much of the offending message data as the notification sender wishes to.



Error Notification message format

- * Error Code: An 8 bit field identifying the type of error that occurred. Defined errors are:
 - Unrecognized message type.
 - Message format error.
- * Reserved: Set to zero by the sender and ignored by the receiver.
- * Offending Message Type: 16 bit type field identifying the message type of the offending message that triggered this error notification. This is copied from the Type field of the offending message.
- * Offending Message Length: 16 bit field that provides the total size of the offending message in octets. This is copied from the Length field of the offending message.
- * Offending Message ID: A 32-bit field that is set to the Message ID field of the offending message.

- * **Offending Message Data:** The Data from the offending message that triggered this error notification. The sender of the notification may include as much of the original data as is deemed necessary. The length of the Offending Message Data field is not provided by the Offending Message Length field and is determined by subtracting the size of the other fields in the message from the Length field. It is valid to not include any of the offending message data when sending an error notification.
- * **End Marker:** A 32-bit message end marker that must be set to 0x9FACADE9. The End Marker is used by the receiver to validate that it has correctly parsed or skipped a message and provides a method to detect formatting errors. Note that message data may also contain this marker, and that the marker itself is not sufficient for parsing the message.

An error notification cannot be the offending message in another error notification and MUST NOT trigger such a message.

6. EID Prefix Registration

EID prefix registration uses the reliable transport session between an ETR and a Map-Server to communicate the ETR local EID database EID to RLOC mappings to the Map-Server. In contrast to the UDP based periodic registration, mapping information over the reliable transport session is only sent when there is new information available for the Map-Server. The Map-Server does not maintain a timer to expire registrations communicated over the reliable transport session. Instead an explicit de-registration (a registration carrying a zero TTL) is needed to delete the state maintained by the Map-Server.

The key used to identify registration mapping records in the ETR to Map-Server communication is the EID prefix. The prefix may be specified using an LCAF encoding that includes an EID instance ID.

When the reliable transport session goes down, registration mappings learned by the Map-Server are treated as periodic UDP registrations and a timer is used to expire them after 3 minutes. During this period UDP based registrations or the re-establishment of the reliable transport session and subsequent communication of a new mapping can update the EID prefix mapping state.

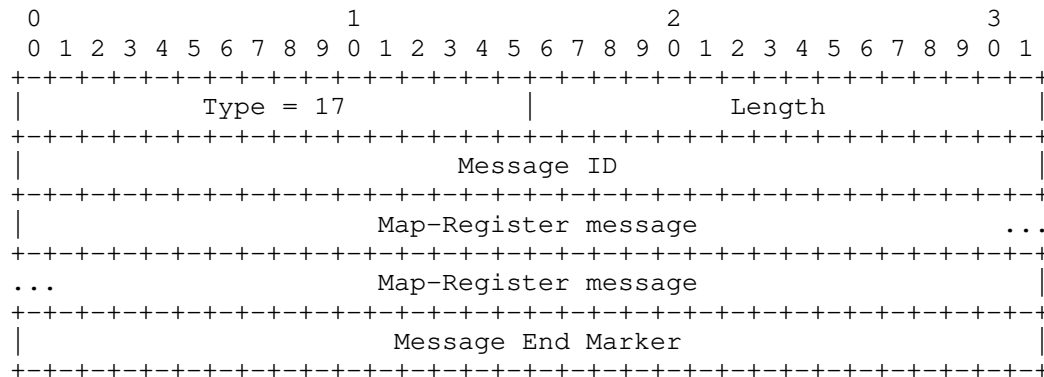
6.1. Reliable Mapping Registration Messages

This section defines the LISP reliable transport session messages used to communicate local EID database registrations between the ETR and the Map-Server.

6.1.1. Registration Message

The reliable transport registration message is used to communicate EID to RLOC mapping registrations from the ETR to the Map-Server. To increase code reuse, the "Message Data" field uses the same format as the UDP Map-Registers but without the IP and UDP headers. A reliable registration message MUST contain a single mapping-record. The map server MUST discard any reliable registration message that contains more than one mapping record.

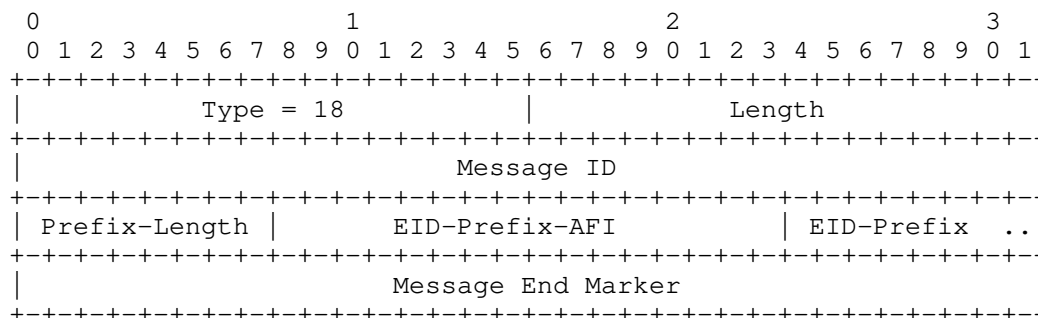
The reliable transport session is authenticated by means of the session establishment procedure. Thus, although the Map-Register MUST carry the authentication data, it is up to the map server to determine if each individual reliable registration message should be authenticated.



Registration message format

6.1.2. Registration Acknowledgement Message

The Acknowledgement message is sent from the Map-Server to the ETR to confirm successful registration of an EID prefix previously communicated by a reliable transport session Registration message. The Registration Acknowledgement message does not carry a mapping record (the map servers view of the mapping). This is accomplished by the LISP reliable transport Map Notification message.

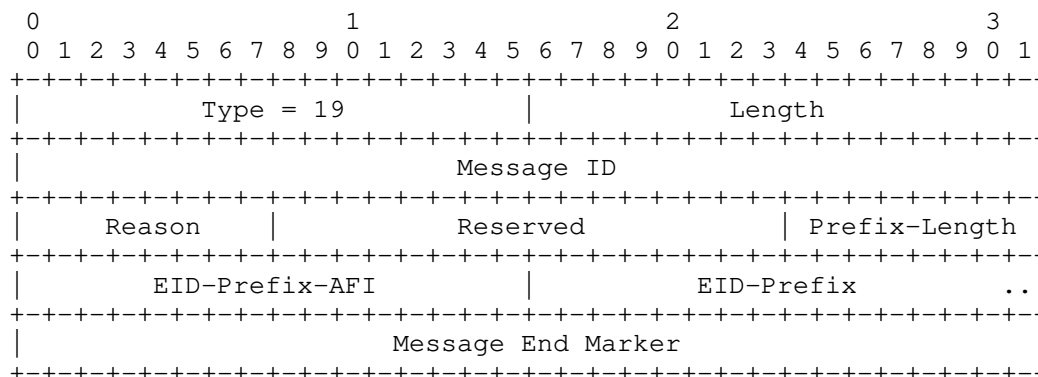


Registration Acknowledgement message format

- * Prefix-Length: Mask length for the EID prefix.
- * EID-Prefix AFI: Address family identifier for the EID prefix in the following field.
- * EID-Prefix: The EID prefix from the received Registration.

6.1.3. Registration Rejection Message

The Registration Rejection Message is sent by the map server to the ETR to indicate that the registration of a specific EID prefix is being rejected or withdrawn. A rejection refers to a recently-sent registration that is being immediately rejected. A withdrawal refers to a previously accepted registration that is no longer acceptable, perhaps due to a configuration change in the map-server. The ETR must keep track of rejected EID prefixes and be prepared to re-register their mappings when requested through a registration refresh message.

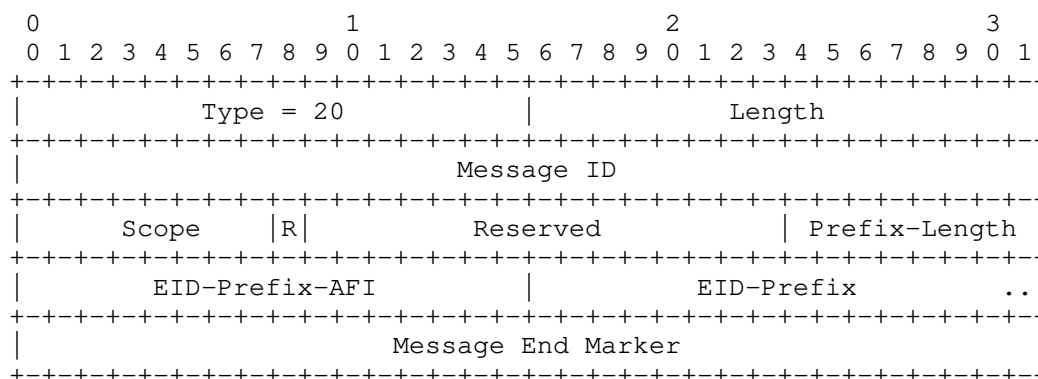


Registration Rejection message format

- * Reason: Code identifying the reason for which the Map-Server rejected or withdrew the registration.
 - 1 - Not a valid site EID prefix.
 - 2 - Authentication failure.
 - 3 - Locator set not allowed.
 - 4 - Used to cover reason that's not defined.
- * Reserved: This field is reserved for future use. Set to zero by the sender and ignored by the receiver.
- * Prefix-Length: Mask length for the EID prefix.
- * EID-Prefix-AFI: Address family identifier for the EID prefix in the following field.
- * EID-Prefix: The EID prefix being rejected or withdrawn.

6.1.4. Registration Refresh Message

Sent by the Map-Server to the ETR to request the (re-)transmission of EID prefix database mapping Registration messages.



Registration Refresh message format

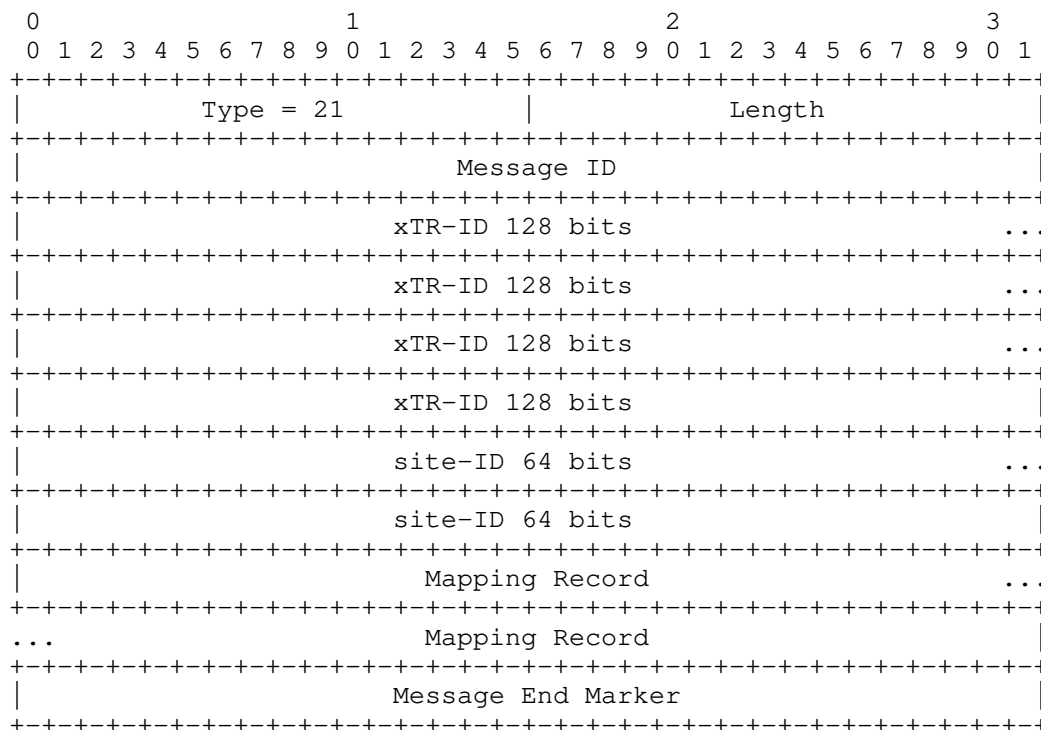
- * Scope: Determines the set of registrations being refreshed.
 - 0 - All prefixes under all address families under all EID instances are being refreshed. When using this scope the Prefix-Length, EID-Prefix-AFI, and EID-Prefix fields MUST be omitted. That is, the Message End Marker follows immediately after the Reserved field. The total length of the message MUST be 15 bytes.
 - 1 - All prefixes under all address families under a single EID instance are being refreshed. The Prefix-Length MUST be set to zero, EID-Prefix-AFI MUST be set to LCAF type, the EID-Prefix encodes the LCAF Instance ID, the LCAF address AFI MUST be set to UNSPECIFIED. The total length of the message MUST be 30 bytes.
 - 2 - All prefixes under a single address family under a single EID instance are being refreshed. The Prefix-Length MUST be set to zero, the EID-Prefix-AFI MUST be set to LCAF type and the EID-Prefix MUST encode the Instance ID. The LCAF address AFI MUST specify the address family to refresh, the actual address SHOULD be set to zero.
 - 3 - All prefixes covered by a specific EID prefix in a single EID instance is being refreshed. The Prefix-Length, EID-Prefix-AFI and EID prefix MUST be encoded accordingly.
 - 4 - A specific EID prefix in a single EID instance is being refreshed. The Prefix-Length, EID-Prefix-AFI and EID prefix MUST be encoded accordingly.

The map-server has the flexibility to control the granularity of the refresh by issuing refresh with different scopes. It can send a single refresh with a coarse scope or send individual refreshes with narrower scope. The ETR MUST be able to process all scopes to ensure the map-server registration states are synchronized with the ETR.

- * R: Request from the ETR to only refresh registrations that have been previously rejected by the Map-Server. If the R bit is set then the scope cannot have a value of 3 and the EID-Prefix and Prefix-Length fields must be omitted.
- * Reserved: This field is reserved for future use. Set to zero by the sender and ignored by the receiver.
- * Prefix-Length: Mask length for the EID prefix. Refer to scope for more details.
- * EID-Prefix-AFI: Address family identifier for the EID prefix in the following field. Refer to scope for more details.
- * EID-Prefix: The EID prefix being refreshed. Refer to scope for more details.

6.1.5. Mapping Notification Message

Mapping Notification messages communicate the Map-Server view of the mapping for an EID prefix and no longer serve as a registration acknowledgement. Mapping Notifications do not need message level authentication as they are received over a reliable transport session to a known Map-Server. Note that reliable transport Mapping Notification messages do not reuse the UDP Map-Notify message format.



Mapping Notification message format

- * xTR-ID: xTR-ID taken from the last valid registration the map-server received for the EID-prefix conveyed in the mapping record.
- * site-ID: site-ID taken from the last valid registration the map-server received for the EID-prefix conveyed in the mapping record.
- * Mapping Record: Mapping record of the EID-prefix the map-server is conveying to the ETR.

6.2. ETR Behavior

The ETR operates the following per EID prefix, per MS state machine that defines the reliable transport EID prefix registration behavior.

There are five states:

- * No state: The local EID database prefix does not exist.
- * Periodic: The local EID database prefix is being periodically registered through UDP Map-Register messages as specified in [1].

- * **Stable:** From the ETR's perspective, no registrations are due to be sent to the peer. The session to the peer is up, and the peer has either acknowledged the registration, or is expected to request a refresh in the future.
- * **AckWait:** A Registration message for the prefix has been transmitted to the Map-Server and the ETR is waiting for either a Registration Acknowledge or Registration Rejected reply from the Map-Server.
- * **Reject:** The reliable transport registration for the local EID database prefix was rejected by the Map-Server. From the ETR's perspective, no registration is due to the peer AND the peer is known to have rejected the registration.

The following events drive the state transitions:

- * **DB creation:** The local EID database entry for the EID prefix is created.
- * **DB deletion:** The local EID database entry for the EID prefix is deleted.
- * **DB change:** The mapping contents or authentication information for the local EID database entry changes.
- * **Session up:** The reliable transport session to the Map-Server is established.
- * **Session down:** The reliable transport session the Map-Server goes down.
- * **Recv Refresh:** A Registration refresh message is received from the Map-Server.
- * **Recv ACK:** A Registration Acknowledge message is received from the Map-Server.
- * **Recv Rejected:** A Registration Rejected message is received from the Map-Server.
- * **Periodic timer:** The timer that drives generation of periodic UDP Map-Register messages fires.

The state machine is:

Event	Prev State	
	No state	Periodic
DB creation [session down]	-> Periodic A1	N/A
DB creation [session up]	-> AckWait A2	N/A
DB deletion	N/A	-> No state A3
DB change	N/A	- A1
Session up	-	-> Stable A4
Session down	-	N/A
Recv Refresh	-	N/A
Recv Refresh [rejected]	-	N/A
Recv ACK	-	N/A
Recv Rejection	-	N/A
Timer	N/A	- A5

xTR per EID prefix per MS state machine

Event	Prev State		
	Stable	AckWait	Rejected
DB creation	N/A	N/A	N/A
DB deletion	-> No state A6	-> No state A6	-> No state
DB change	-> AckWait A2	- A2	-> AckWait A2
Session up	N/A	N/A	N/A
Session down	-> Periodic A7	-> Periodic A7	-> Periodic A7
Recv Refresh	-> AckWait A2	- A2	-> AckWait A2
Recv Refresh [rejected]	-	- A2	-> AckWait A2
Recv ACK	-	-> Stable	-> AckWait A2
Recv Rejection	-> Rejected	-> Rejected	-
Timer	N/A	N/A	N/A

xTR per EID prefix per MS state machine

Action descriptions:

- * A1: Start periodic registration timer with zero delay.
- * A2: Send Registration over reliable transport session.
- * A3: Send UDP registration with zero TTL.
- * A4: Stop periodic registration timer.

- * A7: Send UDP registration and start periodic registration timer with registration period.
- * A6: Send Registration with TTL zero over reliable transport session.
- * A7: Start periodic registration timer with registration period.

All timer start actions must be jittered.

When the reliable transport session is established the ETR moves the state machine into the Stable state without first registering the EID prefix over the reliable transport session. The map server will send a refresh message with a scope of 0 that will trigger the registration message to be sent. Because other applications may be using the reliable session, the refresh message signals the ETR that the map server supports reliable map registration messages. This model will also allow future optimizations where the Map-Server may retain registration state from a previous instantiation of the reliable transport session with the ETR and only request the refresh of EID prefix state beyond some negotiated session progress marker.

Aa Map-Server authentication key change is treated as a DB change event and will result in triggering a new Registration message to be transmitted.

6.3. Map-Server Behavior

Received registrations create/update or delete mapping state.

A refresh with global scope is sent when a session between the ETR and map-server is first established so the map-server can obtain the complete database contents from the ETR. This refresh is also serving as a capability signaling from the map-server to the ETR that it can support reliable registration.

Refresh for rejected registrations sent (R bit set) when a new EID prefix is configured on the Map-Server.

Refresh is sent whenever authentication key is changed or EID prefix is deconfigured. Upon reception of the registration map-server can decide whether to acknowledge the registration or issue rejection.

Mapping Notification message sent whenever the mapping for a registered or more specific prefix for which notifications are requested changes. ETR acknowledgement or rejection messaging for Mapping Notification is not required because the ETR decides how to process the message based on the registered mapping information. If the mapping information changes the resulting registration will trigger a new Mapping Notification message from the Map-Server.

7. Security Considerations

The LISP reliable transport session SHOULD be authenticated. On controlled RLOC networks that can guarantee that the source RLOC address of data packets cannot be spoofed, the authentication check can be a source address validation on the reliable transport packets. When the RLOC network does not provide such guarantees, reliable transport authentication SHOULD be used. Implementations SHOULD support the TCP Authentication Option (TCP-AO) [RFC5925] and SCTP Authenticated Chunks [RFC4895].

8. IANA Considerations

8.1. LISP Reliable Transport Message Types

Assignment of new LISP reliable transport message types is done according to the "IETF Review" model defined in [RFC5266].

The initial content of the registry should be as follows.

Type	Name	Reference
0-15	Reserved	This document
16	Error Notification	This document
17	Registration Message	This document
18	Registration Acknowledgement Message	This document
19	Registration Rejected Message	This document
20	Registration Refresh Message	This document
21	Mapping Notification Message	This document
22-30	Reserved for EID membership distribution	TBD
31-64999	Unassigned	
65000-65535	Reserved for Experimental Use	

8.2. Transport Protocol Port Numbers

TCP port 4342 already reserved for LISP CONS that is now obsolete. Repurpose for reliable transport over TCP. Reserve an SCTP port.

9. Acknowledgments

The authors would like to thank Noel Chiappa, Dino Farinacci, Jesper Skriver, Andre Pelletier and Les Ginsberg for their contributions to this document.

10. Normative References

- [I-D.ietf-lisp-rfc6833bis]
Farinacci, D., Maino, F., Fuller, V., and A. Cabellos,
"Locator/ID Separation Protocol (LISP) Control-Plane",
Work in Progress, Internet-Draft, draft-ietf-lisp-
rfc6833bis-30, 18 November 2020,
<[https://www.ietf.org/archive/id/draft-ietf-lisp-
rfc6833bis-30.txt](https://www.ietf.org/archive/id/draft-ietf-lisp-rfc6833bis-30.txt)>.
- [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>>.
- [RFC5266] Devarapalli, V. and P. Eronen, "Secure Connectivity and
Mobility Using Mobile IPv4 and IKEv2 Mobility and
Multihoming (MOBIKE)", BCP 136, RFC 5266,
DOI 10.17487/RFC5266, June 2008,
<<https://www.rfc-editor.org/info/rfc5266>>.
- [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>>.

Authors' Addresses

Johnson Leong
Cisco Systems
Tasman Drive
San Jose, CA 95134
United States of America

Email: joleong@cisco.com

Darrel Lewis
Cisco Systems
Tasman Drive
San Jose, CA 95134
United States of America

Email: darlewis@cisco.com

Balaji Pitta
Cisco Systems
Tasman Drive
San Jose, CA 95134
United States of America

Email: bvenkata@cisco.com

Chris Cassar
Tesla
10 New Square Park
Bedfont Lakes
TW14 8HA
United Kingdom

Email: christiancassar@acm.org

Isidor Kouvelas
Arista Networks Inc.
5453 Great America Parkway
Santa Clara, CA 95054
United States of America

Email: isidor@kouvelas.net

Jesus Arango
Microsoft

Email: jearango@microsoft.com