Network Working Group
Internet-Draft
Intended status: Experimental
Expires: November 7, 2016

D. Farinacci
lispers.net
P. Pillay-Esnault
Huawei Technologies

May 6, 2016

LISP EID Anonymity
draft-farinacci-lisp-eid-anonymity-00

Abstract

This specification will describe how ephemeral LISP EIDs can be used to create source anonymity. The idea makes use of frequently changing EIDs much like how a credit-card system uses a different credit-card numbers for each transaction.

Requirements Language

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

Status of This Memo

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

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

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

This Internet-Draft will expire on November 7, 2016.

Copyright Notice

Table of Contents

1. Introduction

The LISP architecture [RFC6830] specifies two namespaces, End-Point IDs (EIDs) and Routing Locators (RLOCs). An EID identifies a node in the network and the RLOC indicates the EID's topological location. Typically EIDs are globally unique so a end-node system can connect to any other end-node system on the Internet. Privately used EIDs are allowed when scoped within a VPN but must always be unique within that scope. Therefore, address allocation is required by network administration to avoid address collisions or duplicate address use. In a multiple namespace architecture like LISP, typically the EID will stay fixed while the RLOC can change. This occurs when the EID is mobile or when the LISP site the EID resides in changes its connection to the Internet.

LISP creates the opportunity where EIDs are fixed and won't change. This can create a privacy problem more so than what we have on the Internet today. This draft will examine a technique to allow a end-node system to use a temporary address. The lifetime of a temporary address can be the same as a lifetime of an address in use today on the Internet or can have traditionally shorter lifetimes, possibly on the order of a day or even change as frequent as new connection attempts.

2. Definition of Terms

Ephemeral-EID - is an IP address that is created randomly for use for a temporary period of time. An Ephemeral-EID has all the properties of an EID as defined in [RFC6830]. Ephemeral-EIDs are not stored in the Domain Name System (DNS) and should not be used in long-term address referrals.

Client End-Node - is a network node that originates and consumes packets. It is a system that originates packets or initiates the establishment of transport-layer connections. It does not offer services as a server system would. It accesses servers and attempts to do it anonymously.

3. Overview

A client end-node can assign its own ephemeral EID and use it to talk to any system on the Internet. The system is acting as a client where it initiates communication and desires to be an inaccessible resource from any other system. The ephemeral EID is used as a destination address solely to return packets to resources the ephemeral EID connects to.

Here is the procedure a client end-node would use:

1. Client end-node desires to talk on the network. It creates and assigns an ephemeral-EID on any interface.

2. If the client end-node is a LISP xTR, it will register the ephemeral-EID with a globally routable RLOC. If the client end-node is not a LISP xTR, it can send packets on the network where a LISP router xTR will register the ephemeral-EID with its RLOC.

3. The client end-node originates packets with a source address equal to the ephemeral-EID and will receive packets addressed to the ephemeral-EID.

4. When the client end-node decides to stop using the ephemeral-EID, it will deregister it from the mapping system and create and assign a new ephemeral-EID, or decide to configure a static global address, or participate in DHCP to get assigned a leased address.

Note that the ephemeral-EID can be mobile just like any other EID so if it is initially registered to the mapping system with one or more RLOCs, later the RLOC-set can change as the ephemeral-EID roams.

4. Design Details

This specification proposes the use of the experimental LISP EID-block 2001:5::/32 when IPv6 is used. See IANA Considerations section for a specific sub-block allocation request. When IPv4 is used, the Class E block 240.0.0.0/4 is being proposed.

The client end-node system will use the rest of the host bits to allocate a random number to be used as the ephemeral-EID. The EID can be created manually or via a programatic interface. When the EID address is going to change frequently, it is suggested to use a programatic interface. The probability of address collision is unlikely for IPv6 EIDs but could occur for IPv4 EIDs. A client end-node can create a ephemeral-EID and then look it up in the mapping system to see if it exists. If the EID exists in the mapping system, the client end-node can attempt creation of a new random number for the ephemeral-EID. See Section 7 where ephemeral-EIDs can be preallocated and registered to the mapping system before use.

When the client end-node system is co-located with the RLOC and acts as an xTR, it should register the binding before sending packets. This eliminates a race condition for returning packets not knowing where to encapsulate packets to the ephemeral-EID's RLOCs. When the client end-node system is not acting as an xTR, it should send some packets so its ephemeral-EID can be discovered by an xTR which supports EID-mobility [I-D.portoles-lisp-eid-mobility] so mapping system registration can occur before the destination returns packets. See Section 7 for alternatives for fixing this race condition problem.

5. Interworking Considerations

If a client end-node is communicating with a system that is not in a LISP site, the procedures from [RFC6832] should be followed. The PITR will be required to originate route advertisements for the ephemeral-EID sub-block [I-D.draft-ietf-lisp-eid-block] so it can attract packets sourced by non-LISP sites destined to ephemeral-EIDs. However, in the general case, the coarse block from [I-D.draft-ietf-lisp-eid-block] will be advertised which would cover the sub-block. For IPv4, the 240.0.0.0/4 must be advertised into the IPv4 routing system.

6. Multicast Considerations

A client end-node system can be a member of a multicast group fairly easily since its address is not used for multicast communication as a receiver. This is due to the design characteristics of IGMP [RFC3376] [RFC2236] [RFC1112] and MLD [RFC2710] [RFC3810].

When a client end-node system is a multicast source, there is ephemeral (S,G) state that is created and maintained in the network via multicast routing protocols such as PIM [RFC4602] and when PIM is used with LISP [RFC6802]. In addition, when [I-D.draft-ietf-lisp-signal-free-multicast] is used, ephemeral-EID state is created in the mapping database. This doesn't present any problems other than the amount of state that may exist in the network if not timed out and removed promptly.

However, there exists a multicast source discovery problem when PIM-SSM [RFC4607] is used. Members that join (S,G) channels via out of band mechanisms. These mechanisms need to support ephemeral-EIDs. Otherwise, PIM-ASM [RFC4602] or PIM-Bidir [RFC5015] will need to be used.

7. Performance Improvements

An optimization to reduce the race condition between registering ephemeral-EIDs and returning packets as well as reducing the probability of ephemeral-EID address collision is to preload the mapping database with a list of ephemeral-EIDs before using them. It comes at a expense of rebinding all of registered ephemeral-EIDs when there is an RLOC change. There is work in progress to consider adding a level of indirection here so a single entry gets the RLOC update and the list of ephemeral-EIDs point to the single entry.

8. Security Considerations

When LISP-crypto [I-D.draft-ietf-lisp-crypto] is used the EID payload is more secure through encryption providing EID obfuscation of the ephemeral-EID as well as the global-EID it is communicating with. But the obfuscation only occurs between xTRs. So the randomness of a ephemeral-EID inside of LISP sites provide a new level of privacy.

9. IANA Considerations

This specification is requesting the sub-block 2001:5:ffff::/48 for ephemeral-EID usage.

10. References

10.1. Normative References

[RFC1112] Deering, S., "Host extensions for IP multicasting", STD 5, RFC 1112, DOI 10.17487/RFC1112, August 1989, <http://www.rfc-editor.org/info/rfc1112>.

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

[RFC2236] Fenner, W., "Internet Group Management Protocol, Version 2", RFC 2236, DOI 10.17487/RFC2236, November 1997, <http://www.rfc-editor.org/info/rfc2236>.

[RFC2710] Deering, S., Fenner, W., and B. Haberman, "Multicast Listener Discovery (MLD) for IPv6", RFC 2710, DOI 10.17487/RFC2710, October 1999, <http://www.rfc-editor.org/info/rfc2710>.

[RFC3376] Cain, B., Deering, S., Kouvelas, I., Fenner, B., and A. Thyagarajan, "Internet Group Management Protocol, Version 3", RFC 3376, DOI 10.17487/RFC3376, October 2002, <http://www.rfc-editor.org/info/rfc3376>.

[RFC3810] Vida, R., Ed. and L. Costa, Ed., "Multicast Listener Discovery Version 2 (MLDv2) for IPv6", RFC 3810, DOI 10.17487/RFC3810, June 2004, <http://www.rfc-editor.org/info/rfc3810>.

[RFC4602] Pusateri, T., "Protocol Independent Multicast - Sparse Mode (PIM-SM) IETF Proposed Standard Requirements Analysis", RFC 4602, DOI 10.17487/RFC4602, August 2006, <http://www.rfc-editor.org/info/rfc4602>.

[RFC4607] Holbrook, H. and B. Cain, "Source-Specific Multicast for IP", RFC 4607, DOI 10.17487/RFC4607, August 2006, <http://www.rfc-editor.org/info/rfc4607>.

[RFC5015] Handley, M., Kouvelas, I., Speakman, T., and L. Vicisano, "Bidirectional Protocol Independent Multicast (BIDIR-PIM)", RFC 5015, DOI 10.17487/RFC5015, October 2007, <http://www.rfc-editor.org/info/rfc5015>.

[RFC6802] Baillargeon, S., Flinta, C., and A. Johnsson, "Ericsson Two-Way Active Measurement Protocol (TWAMP) Value-Added Octets", RFC 6802, DOI 10.17487/RFC6802, November 2012, <http://www.rfc-editor.org/info/rfc6802>.

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

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

10.2. Informative References

[I-D.draft-ietf-lisp-crypto]
Farinacci, D. and B. Weis, "LISP Data-Plane Confidentiality", draft-ietf-lisp-crypto-03 (work in progress).

[I-D.draft-ietf-lisp-eid-block]
Iannone, L., Lewis, D., Meyer, D., and V. Fuller, "LISP EID Block", draft-ietf-lisp-eid-block-13.txt (work in progress).

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

[I-D.portoles-lisp-eid-mobility]
Portoles-Comeras, M., Ashtaputre, V., Moreno, V., Maino, F., and D. Farinacci, "LISP L2/L3 EID Mobility Using a Unified Control Plane", draft-portoles-lisp-eid-mobility-00 (work in progress), April 2016.

Appendix A. Acknowledgments

The author would like to thank the LISP WG for their review and acceptance of this draft.

Authors' Addresses

Dino Farinacci
lispers.net
San Jose, CA
USA

Email: farinacci@gmail.com

Padma Pillay-Esnault
Huawei Technologies
San Clara, CA
USA

Email: padma@huawei.com

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: May 17, 2017

D. Farinacci
lispers.net
P. Pillay-Esnault
Huawei Technologies
November 13, 2016

LISP Predictive RLOCs
draft-farinacci-lisp-predictive-rlocs-01

Abstract

This specification will describe a method to achieve near-zero packet loss when an EID is roaming quickly across RLOCs.

Requirements Language

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

Status of This Memo

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

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

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

This Internet-Draft will expire on May 17, 2017.

Copyright Notice

Table of Contents

1. Introduction

The LISP architecture [RFC6830] specifies two namespaces, End-Point IDs (EIDs) and Routing Locators (RLOCs). An EID identifies a node in the network and the RLOC indicates the EID's topological location. When an node roams in the network, its EID remains fixed and unchanged but the RLOCs associated with it change to reflect its new topological attachment point. This specification will focus EIDs and RLOCs residing in separate nodes. An EID is assigned to a host node that roams while the RLOCs are assigned to network nodes that stay stationary and are part of the network topology. For example, a set of devices on an aircraft are assigned EIDs, and base stations on the ground attached to the Internet infrastructure are configured as LISP xTRs where their RLOCs are used for the bindings of the EIDs on the aircraft up in the air.

The scope of this specification will not emphasize general physical roaming as an aircraft would do in the sky but in a direction that is

more predictable such as a train traveling on a track or vehicle that travels along a road.

2. Definition of Terms

Roaming-EID - is a network node that moves from one topological location in the network to another. The network node uses the same EID when it is roaming. That is, the EID address does not change for reasons of mobility. A roaming-EID can also be a roaming EID-prefix where a set of EIDs covered by the prefix are all roaming and fate-sharing the same set of RLOCs at the same time.

Predictive RLOCs - is a set of ordered RLOCs in a list each assigned to LISP xTRs where the next RLOC in the list has high probability it will be the next LISP xTR in a physical path going in a single predictable direction.

Road-Side-Units (RSUs) - is a network node that acts as a router, more specifically as a LISP xTR. The xTR automatically discovers roaming-EIDs that come into network connectivity range and relays packets to and from the roaming-EID. RSUs are typically deployed along a directional path like a train track or road and are in connectivity range of devices that travel along the directional path.

3. Overview

The goal of this specification is to describe a make-before-break EID-mobility mechanism that offers near-zero packet loss. Offering minimal packet loss, not only allows transport layers to operate more efficiently, but because an EID does not change while moving, transport layer session continuity is maintained. To achieve these requirements, a mechanism that reacts to the mobility event is necessary but not sufficient. So the question is not that there isn’t a reaction but when it happens. By using some predictive algorithms, we can guess with high probability where the EID will roam to next. We can achieve this to a point where packet data will be at the new location when the EID arrives.

First we should examine both the send and receive directions with respect to the roaming-EID. Refer to Figure 1 for discussion. We show a network node with a fixed EID address assigned to a roaming-EID moving along a train track. And there are LISP xTRs deployed as Road-Side-Units to support the connectivity between the roaming-EID and the infrastructure or to another roaming-EID.

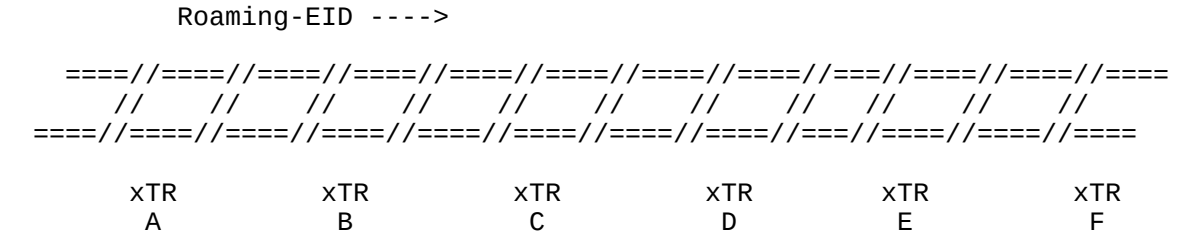


Figure 1: Directional Mobility

For the send direction from roaming-EID to any destination can be accomplish as a local decision. As long as the roaming-EID is in signal range to any xTR along the path, it can use it to forward packets. The LISP xTR, acting as an ITR, can forward packets to destinations in non-LISP sites as well as to stationary and roaming EIDs in LISP sites. This is accomplished by using the LISP overlay via dynamic packet encapsulation. When the roaming-EID sends packets, the LISP xTR must discover the EID and MAY register the EID with a set of RLOCs to the mapping system [I-D.portoles-lisp-eid-mobility]. The discovery process is important because the LISP xTR, acting as an ETR for decapsulating packets that arrive, needs to know what local ports or radios to send packets to the roaming-EID.

Much of the focus of this design is on the packet direction to the roaming-EID. And how remote LISP ITRs find the current location (RLOCs) quickly when the roaming-EID is moving at high speed. This specification solves the fast roaming with the introduction of the Predictive-RLOCs algorithm.

Since a safe assumption is that the roaming-EID is going in one direction and cannot deviate from it allows us to know a priori the next set of RLOCs the roaming-EID will pass by. Referring to Figure 1, if the roaming-EID is in range near xTR-A, then as it moves, it will at some point pass by xTR-B and xTR-C, and so on. As the roaming-EID moves, one could time when the EID is mapped to RLOC A, and when it should change to RLOC B and so on. However, the speed of movement of the roaming-EID won’t be constant and the variables involved in consistent timing cannot be relied on. Furthermore, timing the move is not a make-before-break algorithm, meaning the reaction of the binding happens at the time the roaming-EID is discovered by an xTR. One cannot achieve fast hand-offs when message signaling will be required to inform remote ITRs of the new binding.

The Predictive RLOCs algorithm allows a set of RLOCs, in an ordered list, to be provided to remote ITRs so they have the information

available and local for when they need to use it. Therefore, no control-plane message signaling occurs when the roaming-EID is discovered by LISP xTRs.

4. Design Details

Predictive RLOCs accommodates for encapsulated packets to be delivered to Road-Side-Unit LISP xTRs regardless where the roaming-EID is currently positioned.

Referring to Figure 1, the following sequence is performed:

1. The Predictive RLOCs are registered to the mapping system as a LCAF encoded Replication List Entry (RLE) Type [I-D.ietf-lisp-lcaf]. The registration can happen by one or more RSUs or by a third-party. When registered by an RSU, and when no coordination is desired, they each register their own RLOC with merge-semantics so the list can be created and maintained in the LISP Map-Server. When registered by a third-party, the complete list of RLOCs can be included in the RLE.

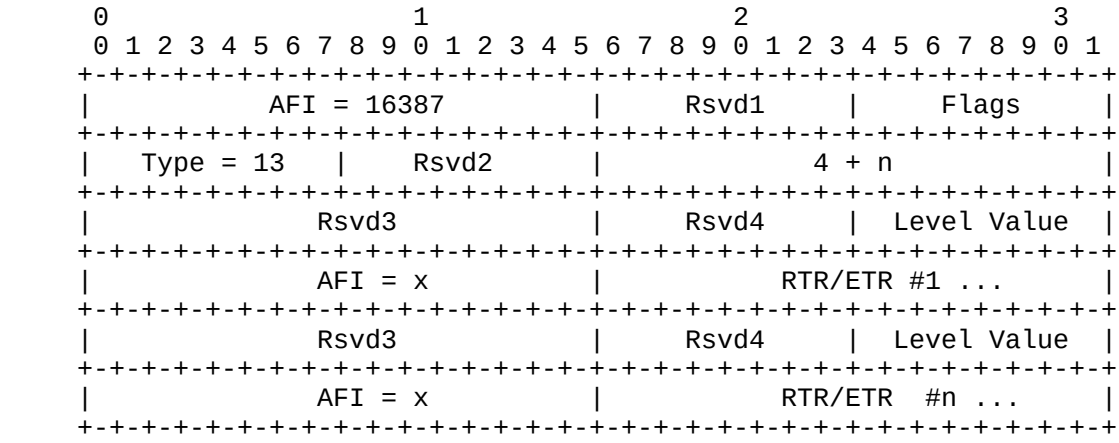
2. There can be multiple RLEs present each as different RLOC-records so a remote ITR can select one RLOC-record versus the other based in priority and weight policy [RFC6830].

3. When a remote ITR receives a packet destined for a roaming-EID, it encapsulates and replicates to each RLOC in the RLE thereby delivering the packet to the locations the roaming-EID is about to appear. There are some cases where packets will go to locations where the roaming-EID has already been, but see Section 4.2 for packet delivery optimizations.

4. When the ETR resident RSU receives an encapsulated packet, it decapsulates the packet and then determines if the roaming-EID had been previously discovered. If the EID has not been discovered, the ETR drops the packet. Otherwise, the ETR delivers the decapsulated packet on the port interface the roaming-EID was discovered on.

4.1. RLE Encoding

The LCAF [I-D.ietf-lisp-lcaf] Replication List Entry (RLE) will be used to encode the Predictive RLOCs in an RLOC-record for Map-Registers, Map-Reply, and Map-Notify messages [RFC6830].



When the RLOC-record contains an RLE with RLOC entries all with the same level value, it means the physical order listed is the directional path of the RSUs. This will typically be the result of a third-party doing the registration where it knows ahead of time the RSU deployment.

When each RSU is registering with merge-semantics on their own, the level number is used to place them in an ordered list. Since the registrations come at different times and therefore arrive in different order than the physical RSU path, the level number creates the necessary sequencing. Each RSU needs to know its position in the path relative to other RSUs. For example, in xTR-B, it would register with level 1 since it is after xTR-A (and before xTR-C). So if the registration order was xTR-B with level 1, xTR-C with level 2, and xTR-A with level 0, the RLE list stored in the mapping system would be (xTR-A, xTR-B, xTR-C). It is recommended that level numbers be assigned in increments of 10 so latter insertion is possible.

The use of Geo-Prefixes and Geo-Points can be used to compare the physical presence of each RSU with respect to each other, so they can choose level numbers to sequence themselves. Also if the xTRs register with a Geo-Point in an RLOC-record, then perhaps the Map-Server could sequence the RLE list.

4.2. Packet Delivery Optimizations

Since the remote ITR will replicate to all RLOCs in the RLE, a situation is created where packets go to RLOCs that don’t need to. For instance, if the roaming-EID is along side of xTR-B and the RLE is (xTR-A, xTR-B, xTR-C), there is no reason to replicate to xTR-A since the roaming-EID has passed it and the the signal range is weak or lost. However, replicating to xTR-B and xTR-C is important to

deliver packets to where the roaming-EID resides and where it is about to go to.

A simple data-plane option, which converges fairly quickly is to have the remote xTR, acting as an ETR, when packets are sent from the roaming-EID, examine the source RLOC in the outer header of the encapsulated packet. If the source RLOC is xTR-B, the remote xTR can determine that the roaming-EID has moved past xTR-A and no longer needs to encapsulate packets to xTR-A's RLOC.

In addition, the remote ITR can use RLOC-probing to determine if each RLOC in the RLE is reachable. And if not reachable, exclude from the list of RLOCs to replicate to.

This solution also handles the case where xTR-A and xTR-B may overlap in radio signal range, but the signal is weak from the roaming-EID to xTR-A but stronger to xTR-B. In this case, the roaming-EID selects xTR-B to send packets that inform the remote xTR that return packets should not be encapsulated to xTR-A.

There are also situations where the RSUs are in signal range of each other in which case they could report reachability status of each other. The use of the Locator-Status-Bits of the LISP encapsulation header could be used to convey this information to the remote xTR. This would only occur when the roaming-EID was discovered by both xTR-A and xTR-B so it was possible for either xTR to reach the roaming-EID. Either an IGP like routing protocol would be required to allow each xTR to know the other could reach the roaming-EID or a path trace tool (i.e. traceroute) could be originated by one xTR targeted for the roaming-EID but MAC-forwarded through the other xTR. These and other roaming-EID reachability mechanisms are work in progress and for further study.

4.3. Trading Off Replication Cost

If RLE lists are large, packet replication can occur to locations well before the roaming-EID arrives. Making RLE lists small is useful without sacrificing hand-off issues or incurring packet loss to the application. By having overlapping RLEs in separate RLOC-records we a simple mechanism to solve this problem. Here is an example mapping entry to illustrate the point:

```
EID = <roaming-EID>, RLOC-records:
  RLOC = (RLE: xTR-A, xTR-B)
  RLOC = (RLE: xTR-B, xTR-C, xTR-D, xTR-E)
  RLOC = (RLE: xTR-E, xTR-F)
```

When the remote ITR is encapsulating to xTR-B as a decision to use the first RLOC-record, it can decide to move to use the second RLOC-record because xTR-B is the last entry in the first RLOC-record and the first entry in the second RLOC-record. When there are overlapping RLEs, the remote ITR can decide when it is more efficient to switch over. For example, when the roaming-EID is in range of xTR-A, the remote ITR uses the first RLOC-record so the wasted replication cost is to xTR-B only versus a worse cost when using the second RLOC-record. But when the roaming-EID is in range of xTR-B, then replicating to the other xTRs in the second RLOC-record may be crucial if the roaming-EID has increased speed. And when the roaming-EID may be at rest in a parked mode, then the remote ITR encapsulates to only xTR-F using the third RLOC-record since the roaming-EID has moved past xTR-E.

In addition, to eliminate unnecessary replication to xTRs further down a directional path, GEO-prefixes [I-D.farinacci-lisp-geo] can be used so only nearby xTRs that the roaming-EID is about to come in contact with are the only ones to receive encapsulated packets.

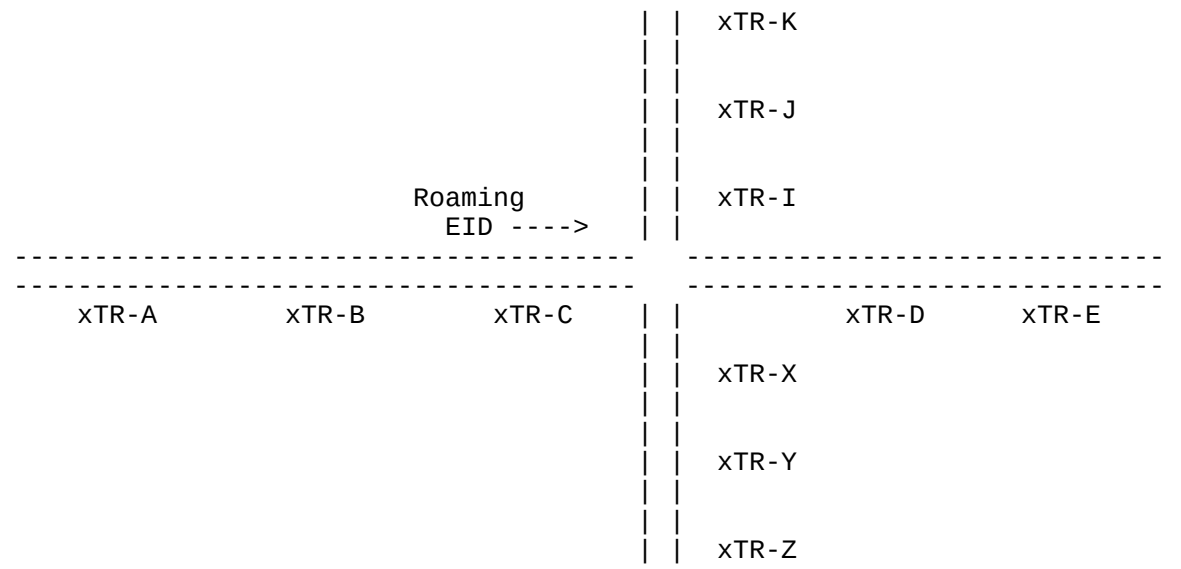
Even when replication lists are not large, we can reduce the cost of replication that the entire network bears by moving the replicator away from the the source (i.e. the ITR) and closer to the RSUs (i.e. the ETRs). See the use of RTRs for Replication Engineering techniques in [I-D.ietf-lisp-signal-free-multicast].

5. Directional Paths with Intersections

A roaming-EID could be registered to the mapping system with the following nested RLE mapping:

```
EID = <roaming-EID>, RLOC-records:
  RLOC = (RLE: xTR-A, xTR-B, xTR-C, (RLE: xTR-X, xTR-Y, xTR-Z),
                (RLE: xTR-I, xTR-J, xTR-K), xTR-D, xTR-E)
```

The mapping entry above describes 3 directional paths where the ordered list has encoded one-level of two nested RLEs to denote intersections in a horizontal path. Which is drawn as:



When the roaming-EID is on the horizontal path, the remote-ITRs typically replicate to the rest the of the xTRs in the ordered list. When a list has nested RLEs, the replication should occur to at least the first RLOC in a nested RLE list. So if the remote-ITR is replicating to xTR-C, xTR-D, and xTR-E, it should also replicate to xTR-X and xTR-I anticipating a possible turn at the intersection. But when the roaming-EID is known to be at xTR-D (a left or right hand turn was not taken), replication should only occur to xTR-D and xTR-E. Once either xTR-I or xTR-X is determined to be where the roaming-EID resides, then the replication occurs on the respective directional path only.

When nested RLEs are used it may be difficult to get merge-semantics to work when each xTR registers itself. So it is suggested a third-party registers nested RLEs. It is left to further study to understand better how to automate this.

6. Multicast Considerations

In this design, the remote ITR is receiving a unicast packet from an EID and replicating and encapsulating to each RLOC in an RLE list. This form of replication is no different than a traditional multicast replication function. So replicating multicast packets in the same fashion is a fallout from this design.

If there are multiple roaming-EIDs joined to the same multicast group but reside at different RSUs, a merge has to be done of any pruned RLEs used for forwarding. So if roaming-EID-1 resides at xTR-A and

roaming-EID-2 resides at xTR-B and the RLE list is (xTR-A, xTR-B, xTR-C), and they are joined to the same multicast group, then replication occurs to all of xTR-A, xTR-B, and xTR-C. Even since roaming-EID-2 is past xTR-A, packets need to be delivered to xTR-A for roaming-EID-1. In addition, packets need to be delivered to xTR-C because roaming-EID-1 and roaming-EID-2 will get to xTR-C (and roaming-EID-1 may get there sooner if it is traveling faster than roaming-EID-2).

When a roaming-EID is a multicast source, procedures from [I-D.ietf-lisp-signal-free-multicast] are used to deliver packets to multicast group members anywhere in the network. The solution requires no signaling to the RSUs. When RSUs receive multicast packets from a roaming-EID, they do a (roaming-EID,G) mapping database lookup to find the replication list of ETRs to encapsulate to.

7. Multiple Address-Family Considerations

Note that roaming-EIDs can be assigned IPv6 EID addresses while the RSU xTRs could be using IPv4 RLOC addresses. Any combination of address-families can be supported as well as for multicast packet forwarding, where (S,G) are IPv6 addresses entries and replication is done with IPv4 RLOCs in the outer header.

8. Scaling Considerations

One can imagine there will be a large number of roaming-EIDs. So there is a strong desire to efficiently store state in the mapping database and the in remote ITRs map-caches. It is likely, that roaming-EIDs may share the same path and move at the same speed (EID devices on a train) and therefore share the same Predictive RLOCs. And since EIDs are not reassigned for mobility purposes or may be temporal , they will not be topologically aggregatable, so they cannot compress into a single EID-prefix mapping entry that share the same RLOC-set.

By using a level of indirection with the mapping system this problem can be solved. The following mapping entries could exist in the mapping database:

```
EID = <eid1>, RLOC-records:
  RLOC = (afi=<dist-name>: "am-train-to-paris")
EID = <eid2>, RLOC-records:
  RLOC = (afi=<dist-name>: "am-train-to-paris")
EID = <eid3>, RLOC-records:
  RLOC = (afi=<dist-name>: "am-train-to-paris")

EID = "am-train-to-paris", RLOC-records:
  RLOC = (afi=lcaf/RLE-type: xTR-A, xTR-B, xTR-C)

EID = "am-train-to-paris-passengers", RLOC-records:
  RLOC = (afi=lcaf/afi-list-type: <eid1>, <eid2>, <eid3>)
```

Each passenger that boards a train has their EID registered to point to the name of the train "am-train-to-paris". And then the train with EID "am-train-to-paris" stores the Predictive RLOC-set. When a remote-ITR wants to encapsulate packets for an EID, it looks up the EID in the mapping database gets the name "am-train-to-paris" returned. Then the remote-ITR does another lookup for the name "am-train-to-paris" to get the RLE list returned.

When new EIDs board the train, the RLE mapping entry does not need to be modified. Only an EID-to-name mapping is registered for the specific new EID. Optionally, another name "am-train-to-paris-passengers" can be registered as an EID to allow mapping to all specific EIDs which are on the train. This can be used for inventory, billing, or security purposes.

This optimization comes at a cost of a 2-stage lookup. However, if both sets of mapping entries are registered to the same Map-Server, a combined RLOC-set could be returned. This idea is for further study.

9. Security Considerations

LISP has procedures for supporting both control-plane security [I-D.ietf-lisp-sec] and data-plane security [I-D.ietf-lisp-crypto].

10. IANA Considerations

At this time there are no requests for IANA.

11. References

11.1. Normative References

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

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

11.2. Informative References

[I-D.farinacci-lisp-geo] Farinacci, D., "LISP Geo-Coordinate Use-Cases", draft-farinacci-lisp-geo-02 (work in progress), October 2016.

[I-D.ietf-lisp-crypto] Farinacci, D. and B. Weis, "LISP Data-Plane Confidentiality", draft-ietf-lisp-crypto-10 (work in progress), October 2016.

[I-D.ietf-lisp-lcaf] Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", draft-ietf-lisp-lcaf-20 (work in progress), October 2016.

[I-D.ietf-lisp-sec] Maino, F., Ermagan, V., Cabellos-Aparicio, A., and D. Saucez, "LISP-Security (LISP-SEC)", draft-ietf-lisp-sec-11 (work in progress), October 2016.

[I-D.ietf-lisp-signal-free-multicast] Moreno, V. and D. Farinacci, "Signal-Free LISP Multicast", draft-ietf-lisp-signal-free-multicast-02 (work in progress), October 2016.

[I-D.portoles-lisp-eid-mobility] Portoles-Comeras, M., Ashtaputre, V., Moreno, V., Maino, F., and D. Farinacci, "LISP L2/L3 EID Mobility Using a Unified Control Plane", draft-portoles-lisp-eid-mobility-01 (work in progress), October 2016.

Appendix A. Acknowledgments

The author would like to thank the LISP WG for their review and acceptance of this draft.

Appendix B. Document Change Log

[RFC Editor: Please delete this section on publication as RFC.]

B.1. Changes to draft-farinacci-lisp-predictive-rlocs-01.txt

- Posted November 2016 to update document timer.

B.2. Changes to draft-farinacci-lisp-predictive-rlocs-00.txt

- Initial post April 2016.

Authors' Addresses

Dino Farinacci
lispers.net
San Jose, CA
USA

Email: farinacci@gmail.com

Padma Pillay-Esnault
Huawei Technologies
San Clara, CA
USA

Email: padma@huawei.com

Internet Engineering Task Force
Internet-Draft
Intended status: Experimental
Expires: September 1, 2017

D. Farinacci
lispers.net
M. Kowal
cisco Systems
P. Lahiri
February 28, 2017

LISP Traffic Engineering Use-Cases
draft-farinacci-lisp-te-12

Abstract

This document describes how LISP reencapsulating tunnels can be used for Traffic Engineering purposes. The mechanisms described in this document require no LISP protocol changes but do introduce a new locator (RLOC) encoding. The Traffic Engineering features provided by these LISP mechanisms can span intra-domain, inter-domain, or combination of both.

Table of Contents

1. Requirements Language

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

2. Introduction

This document describes the Locator/Identifier Separation Protocol (LISP), which provides a set of functions for routers to exchange information used to map from non globally routeable Endpoint Identifiers (EIDs) to routeable Routing Locators (RLOCs). It also defines a mechanism for these LISP routers to encapsulate IP packets addressed with EIDs for transmission across the Internet that uses RLOCs for routing and forwarding.

When LISP routers encapsulate packets to other LISP routers, the path stretch is typically 1, meaning the packet travels on a direct path from the encapsulating ITR to the decapsulating ETR at the destination site. The direct path is determined by the underlying routing protocol and metrics it uses to find the shortest path.

This specification will examine how reencapsulating tunnels [RFC6830] can be used so a packet can take an adminstratively specified path, a congestion avoidance path, a failure recovery path, or multiple load-shared paths, as it travels from ITR to ETR. By introducing an Explicit Locator Path (ELP) locator encoding [RFC8060], an ITR can encapsulate a packet to a Reencapsulating Tunnel Router (RTR) which decapsulates the packet, then encapsulates it to the next locator in the ELP.

3. Definition of Terms

Endpoint ID (EID): An EID is a 32-bit (for IPv4) or 128-bit (for IPv6) value used in the source and destination address fields of the first (most inner) LISP header of a packet. The host obtains a destination EID the same way it obtains an destination address today, for example through a Domain Name System (DNS) [RFC1034] lookup or Session Invitation Protocol (SIP) [RFC3261] exchange. The source EID is obtained via existing mechanisms used to set a host’s "local" IP address. An EID used on the public Internet must have the same properties as any other IP address used in that manner; this means, among other things, that it must be globally unique. An EID is allocated to a host from an EID-prefix block associated with the site where the host is located. An EID can be used by a host to refer to other hosts. EIDs MUST NOT be used as LISP RLOCs. Note that EID blocks MAY be assigned in a hierarchical manner, independent of the network topology, to facilitate scaling of the mapping database. In addition, an EID block assigned to a site may have site-local structure (subnetting) for routing within the site; this structure is not visible to the global routing system. In theory, the bit string that represents an EID for one device can represent an RLOC for a different device. As the architecture is realized, if a given bit

string is both an RLOC and an EID, it must refer to the same entity in both cases. When used in discussions with other Locator/ID separation proposals, a LISP EID will be called a "LEID". Throughout this document, any references to "EID" refers to an LEID.

Routing Locator (RLOC): A RLOC is an IPv4 [RFC0791] or IPv6 [RFC2460] address of an egress tunnel router (ETR). A RLOC is the output of an EID-to-RLOC mapping lookup. An EID maps to one or more RLOCs. Typically, RLOCs are numbered from topologically-aggregatable blocks that are assigned to a site at each point to which it attaches to the global Internet; where the topology is defined by the connectivity of provider networks, RLOCs can be thought of as PA addresses. Multiple RLOCs can be assigned to the same ETR device or to multiple ETR devices at a site.

Reencapsulating Tunnel Router (RTR): An RTR is a router that acts as an ETR (or PETR) by decapsulating packets where the destination address in the "outer" IP header is one of its own RLOCs. Then acts as an ITR (or PITR) by making a decision where to encapsulate the packet based on the next locator in the ELP towards the final destination ETR.

Explicit Locator Path (ELP): The ELP is an explicit list of RLOCs for each RTR a packet must travel to along its path toward a final destination ETR (or PETR). The list is a strict ordering where each RLOC in the list is visited. However, the path from one RTR to another is determined by the underlying routing protocol and how the infrastructure assigns metrics and policies for the path.

Recursive Tunneling: Recursive tunneling occurs when a packet has more than one LISP IP header. Additional layers of tunneling MAY be employed to implement traffic engineering or other re-routing as needed. When this is done, an additional "outer" LISP header is added and the original RLOCs are preserved in the "inner" header. Any references to tunnels in this specification refers to dynamic encapsulating tunnels and they are never statically configured.

Reencapsulating Tunnels: Reencapsulating tunneling occurs when an ETR removes a LISP header, then acts as an ITR to prepend another LISP header. Doing this allows a packet to be re-routed by the reencapsulating router without adding the overhead of additional tunnel headers. Any references to tunnels in this specification refers to dynamic encapsulating tunnels and they are never statically configured. When using multiple mapping database systems, care must be taken to not create reencapsulation loops through misconfiguration.

4. Overview

Typically, a packet's path from source EID to destination EID travels through the locator core via the encapsulating ITR directly to the decapsulating ETR as the following diagram illustrates:

Legend:

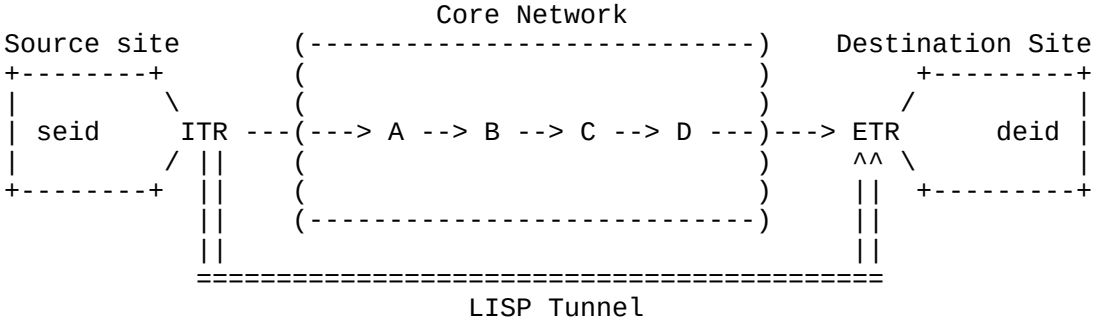
seid: Packet is originated by source EID 'seid'.

deid: Packet is consumed by destination EID 'deid'.

A,B,C,D : Core routers in different ASes.

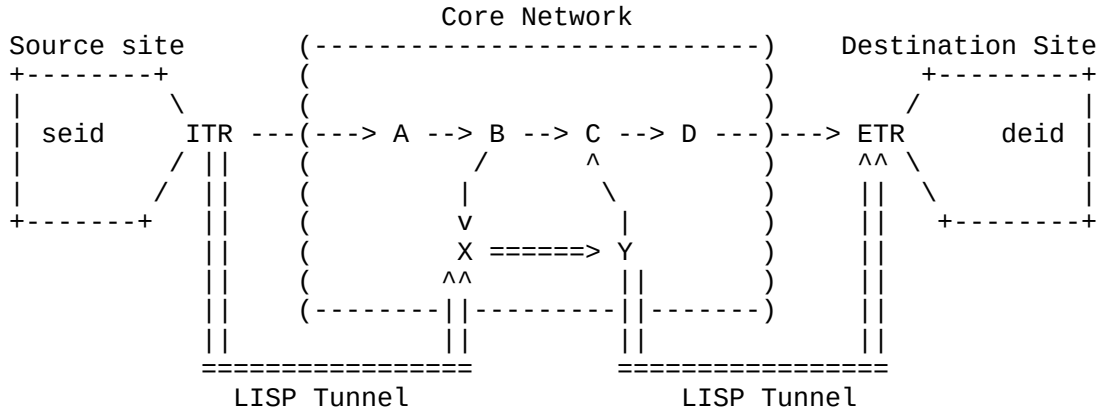
---> : The physical topological path between two routers.

===> : A multi-hop LISP dynamic tunnel between LISP routers.



Typical Data Path from ITR to ETR

Let's introduce RTRs 'X' and 'Y' so that, for example, if it is desirable to route around the path from B to C, one could provide an ELP of (X,Y,etr):



ELP tunnel path ITR ==> X, then X ==> Y, and then Y ==> ETR

There are various reasons why the path from 'seid' to 'deid' may want to avoid the path from B to C. To list a few:

- There may not be sufficient capacity provided by the networks that connect B and C together.
- There may be a policy reason to avoid the ASes that make up the path between B and C.
- There may be a failure on the path between B and C which makes the path unreliable.
- There may be monitoring or traffic inspection resources close to RTRs X and Y that do network accounting or measurement.
- There may be a chain of services performed at RTRs X and Y regardless if the path from ITR to ETR is through B and C.

5. Explicit Locator Paths

The notation for a general formatted ELP is (x, y, etr) which represents the list of RTRs a packet SHOULD travel through to reach the final tunnel hop to the ETR.

The procedure for using an ELP at each tunnel hop is as follows:

1. The ITR will retrieve the ELP from the mapping database.
2. The ITR will encapsulate the packet to RLOC 'x'.

3. The RTR with RLOC 'x' will decapsulate the packet. It will use the decapsulated packet's destination address as a lookup into the mapping database to retrieve the ELP.

4. RTR 'x' will encapsulate the packet to RTR with RLOC 'y'.

5. The RTR with RLOC 'y' will decapsulate the packet. It will use the decapsulated packet's destination address as a lookup into the mapping database to retrieve the ELP.

6. RTR 'y' will encapsulate the packet on the final tunnel hop to ETR with RLOC 'etr'.

7. The ETR will decapsulate the packet and deliver the packet to the EID inside of its site.

The specific format for the ELP can be found in [RFC8060]. It is defined that an ELP will appear as a single encoded locator in a locator-set. Say for instance, we have a mapping entry for EID-prefix 10.0.0.0/8 that is reachable via 4 locators. Two locators are being used as active/active and the other two are used as active/active if the first two go unreachable (as noted by the priority assignments below). This is what the mapping entry would look like:

```
EID-prefix:   10.0.0.0/8
Locator-set:  ETR-A: priority 1, weight 50
              ETR-B: priority 1, weight 50
              ETR-C: priority 2, weight 50
              ETR-D: priority 2, weight 50
```

If an ELP is going to be used to have a policy path to ETR-A and possibly another policy path to ETR-B, the locator-set would be encoded as follows:

```
EID-prefix:   10.0.0.0/8
Locator-set:  (x, y, ETR-A): priority 1, weight 50
              (q, r, ETR-B): priority 1, weight 50
              ETR-C:         priority 2, weight 50
              ETR-D:         priority 2, weight 50
```

The mapping entry with ELP locators is registered to the mapping database system just like any other mapping entry would. The registration is typically performed by the ETR(s) that are assigned and own the EID-prefix. That is, the destination site makes the choice of the RTRs in the ELP. However, it may be common practice for a provisioning system to program the mapping database with ELPs.

Another case where a locator-set can be used for flow-based load-sharing across multiple paths to the same destination site:

```
EID-prefix:   10.0.0.0/8
Locator-set:  (x, y, ETR-A): priority 1, weight 75
              (q, r, ETR-A): priority 1, weight 25
```

Using this mapping entry, an ITR would load split 75% of the EID flows on the (x, y, ETR-A) ELP path and 25% of the EID flows on the (q, r, ETR-A) ELP path. If any of the ELPs go down, then the other can take 100% of the load.

5.1. ELP Re-optimization

ELP re-optimization is a process of changing the RLOCs of an ELP due to underlying network change conditions. Just like when there is any locator change for a locator-set, the procedures from the main LISP specification [RFC6830] are followed.

When a RLOC from an ELP is changed, Map-Notify messages [RFC6833] can be used to inform the existing RTRs in the ELP so they can do a lookup to obtain the latest version of the ELP. Map-Notify messages can also be sent to new RTRs in an ELP so they can get the ELP in advance to receiving packets that will use the ELP. This can minimize packet loss during mapping database lookups in RTRs.

5.2. Using Recursion

In the previous examples, we showed how an ITR encapsulates using an ELP of (x, y, etr). When a packet is encapsulated by the ITR to RTR 'x', the RTR may want a policy path to RTR 'y' and run another level of reencapsulating tunnels for packets destined to RTR 'y'. In this case, RTR 'x' does not encapsulate packets to 'y' but rather performs a mapping database lookup on the address 'y', requests the ELP for RTR 'y', and encapsulates packets to the first-hop of the returned ELP. This can be done when using a public or private mapping database. The decision to use address 'y' as an encapsulation address versus a lookup address is based on the L-bit setting for 'y' in the ELP entry. The decision and policy of ELP encodings are local to the entity which registers the EID-prefix associated with the ELP.

Another example of recursion is when the ITR uses the ELP (x, y, etr) to first prepend a header with a destination RLOC of the ETR and then prepend another header and encapsulate the packet to RTR 'x'. When RTR 'x' decapsulates the packet, rather than doing a mapping database lookup on RTR 'y' the last example showed, instead RTR 'x' does a mapping database lookup on ETR 'etr'. In this scenario, RTR 'x' can

choose an ELP from the locator-set by considering the source RLOC address of the ITR versus considering the source EID.

This additional level of recursion also brings advantages for the provider of RTR 'x' to store less state. Since RTR 'x' does not need to look at the inner most header, it does not need to store EID state. It only stores an entry for RTR 'y' which many EID flows could share for scaling benefits. The locator-set for entry 'y' could either be a list of typical locators, a list of ELPs, or combination of both. Another advantage is that packet load-splitting can be accomplished by examining the source of a packet. If the source is an ITR versus the source being the last-hop of an ELP the last-hop selected, different forwarding paths can be used.

5.3. ELP Selection based on Class of Service

Paths to an ETR may want to be selected based on different classes of service. Packets from a set of sources that have premium service can use ELP paths that are less congested where normal sources use ELP paths that compete for less resources or use longer paths for best effort service.

Using source/destination lookups into the mapping database can yield different ELPs. So for example, a premium service flow with (source=1.1.1.1, dest=10.1.1.1) can be described by using the following mapping entry:

```
EID-prefix:   (1.0.0.0/8, 10.0.0.0/8)
Locator-set:  (x, y, ETR-A): priority 1, weight 50
              (q, r, ETR-A): priority 1, weight 50
```

And all other best-effort sources would use different mapping entry described by:

```
EID-prefix:   (0.0.0.0/0, 10.0.0.0/8)
Locator-set:  (x, x', y, y', ETR-A): priority 1, weight 50
              (q, q', r, r', ETR-A): priority 1, weight 50
```

If the source/destination lookup is coupled with recursive lookups, then an ITR can encapsulate to the ETR, prepending a header that selects source address ITR-1 based on the premium class of service source, or selects source address ITR-2 for best-effort sources with normal class of service. The ITR then does another lookup in the mapping database on the prepended header using lookup key (source=ITR-1, dest=10.1.1.1) that returns the following mapping entry:

```
EID-prefix:   (ITR-1, 10.0.0.0/8)
Locator-set:  (x, y, ETR-A): priority 1, weight 50
              (q, r, ETR-A): priority 1, weight 50
```

And all other sources would use different mapping entry with a lookup key of (source=ITR-2, dest=10.1.1.1):

```
EID-prefix:   (ITR-2, 10.0.0.0/8)
Locator-set:  (x, x', y, y', ETR-A): priority 1, weight 50
              (q, q', r, r', ETR-A): priority 1, weight 50
```

This will scale the mapping system better by having fewer source/destination combinations. Refer to the Source/Dest LCAF type described in [RFC8060] for encoding EIDs in Map-Request and Map-Register messages.

5.4. Packet Loop Avoidance

An ELP that is first used by an ITR must be inspected for encoding loops. If any RLOC appears twice in the ELP, it MUST not be used.

Since it is expected that multiple mapping systems will be used, there can be a loop across ELPs when registered in different mapping systems. The TTL copying procedures for reencapsulating tunnels and recursive tunnels in [RFC6830] MUST be followed.

6. Service Chaining

An ELP can be used to deploy services at each reencapsulation point in the network. One example is to implement a scrubber service when a destination EID is being DoS attacked. That is, when a DoS attack is recognized when the encapsulation path is between ITR and ETR, an ELP can be registered for a destination EID to the mapping database system. The ELP can include an RTR so the ITR can encapsulate packets to the RTR which will decapsulate and deliver packets to a scrubber service device. The scrubber could decide if the offending packets are dropped or allowed to be sent to the destination EID. In which case, the scurbber delivers packets back to the RTR which encapsulates to the ETR.

7. RLOC Probing by RTRs

Since an RTR knows the next tunnel hop to encapsulate to, it can monitor the reachability of the next-hop RTR RLOC by doing RLOC-probing according to the procedures in [RFC6830]. When the RLOC is determined unreachable by the RLOC-probing mechanisms, the RTR can use another locator in the locator-set. That could be the final ETR,

a RLOC of another RTR, or an ELP where it must search for itself and use the next RLOC in the ELP list to encapsulate to.

RLOC-probing can also be used to measure delay on the path between RTRs and when it is desirable switch to another lower delay ELP.

8. Interworking Considerations

[RFC6832] defines procedures for how non-LISP sites talk to LISP sites. The network elements defined in the Interworking specification, the proxy ITR (PITR) and proxy ETR (PETR) (as well as their multicast counterparts defined in [RFC6831]) can participate in LISP-TE. That is, a PITR and a PETR can appear in an ELP list and act as an RTR.

Note when an RLOC appears in an ELP, it can be of any address-family. There can be a mix of IPv4 and IPv6 locators present in the same ELP. This can provide benefits where islands of one address-family or the other are supported and connectivity across them is necessary. For instance, an ELP can look like:

(x4, a46, b64, y4, etr)

Where an IPv4 ITR will encapsulate using an IPv4 RLOC ’x4’ and ’x4’ could reach an IPv4 RLOC ’a46’, but RTR ’a46’ encapsulates to an IPv6 RLOC ’b64’ when the network between them is IPv6-only. Then RTR ’b64’ encapsulates to IPv4 RLOC ’y4’ if the network between them is dual-stack.

Note that RTRs can be used for NAT-traversal scenarios [I-D.ermagan-lisp-nat-traversal] as well to reduce the state in both an xTR that resides behind a NAT and the state the NAT needs to maintain. In this case, the xTR only needs a default map-cache entry pointing to the RTR for outbound traffic and all remote ITRs can reach EIDs through the xTR behind a NAT via a single RTR (or a small set RTRs for redundancy).

RTRs have some scaling features to reduce the number of locator-set changes, the amount of state, and control packet overhead:

- When ITRs and PITRs are using a small set of RTRs for encapsulating to "orders of magnitude" more EID-prefixes, the probability of locator-set changes are limited to the RTR RLOC changes versus the RLOC changes for the ETRs associated with the EID-prefixes if the ITRs and PITRs were directly encapsulating to the ETRs. This comes at an expense in packet stretch, but depending on RTR placement, this expense can be mitigated.

- When RTRs are on-path between many pairwise EID flows, ITRs and PITRs can store a small number of coarse EID-prefixes.

- RTRs can be used to help scale RLOC-probing. Instead of ITRs RLOC-probing all ETRs for each destination site it has cached, the ITRs can probe a smaller set of RTRs which in turn, probe the destination sites.

9. Multicast Considerations

ELPs have application in multicast environments. Just like RTRs can be used to provide connectivity across different address family islands, RTRs can help concatenate a multicast region of the network to one that does not support native multicast.

Note there are various combinations of connectivity that can be accomplished with the deployment of RTRs and ELPs:

- Providing multicast forwarding between IPv4-only-unicast regions and IPv4-multicast regions.

- Providing multicast forwarding between IPv6-only-unicast regions and IPv6-multicast regions.

- Providing multicast forwarding between IPv4-only-unicast regions and IPv6-multicast regions.

- Providing multicast forwarding between IPv6-only-unicast regions and IPv4-multicast regions.

- Providing multicast forwarding between IPv4-multicast regions and IPv6-multicast regions.

An ITR or PITR can do a (S-EID,G) lookup into the mapping database. What can be returned is a typical locator-set that could be made up of the various RLOC addresses:

```
Multicast EID key:  (seid, G)
Locator-set:        ETR-A: priority 1, weight 25
                    ETR-B: priority 1, weight 25
                    g1:    priority 1, weight 25
                    g2:    priority 1, weight 25
```

An entry for host 'seid' sending to application group 'G'

The locator-set above can be used as a replication list. That is some RLOCs listed can be unicast RLOCs and some can be delivery group RLOCs. A unicast RLOC in this case is used to encapsulate a

multicast packet originated by a multicast source EID into a unicast packet for unicast delivery on the underlying network. ETR-A could be a IPv4 unicast RLOC address and ETR-B could be a IPv6 unicast RLOC address.

A delivery group address is used when a multicast packet originated by a multicast source EID is encapsulated in a multicast packet for multicast delivery on the underlying network. Group address 'g1' could be a IPv4 delivery group RLOC and group address 'g2' could be an IPv6 delivery group RLOC.

Flexibility for these various types of connectivity combinations can be achieved and provided by the mapping database system. And the RTR placement allows the connectivity to occur where the differences in network functionality are located.

Extending this concept by allowing ELPs in locator-sets, one could have this locator-set registered in the mapping database for (seid, G). For example:

```
Multicast EID key:  (seid, G)
Locator-set:        (x, y, ETR-A):    priority 1, weight 50
                    (a, g, b, ETR-B): priority 1, weight 50
```

Using ELPs for multicast flows

In the above situation, an ITR would encapsulate a multicast packet originated by a multicast source EID to the RTR with unicast RLOC 'x'. Then RTR 'x' would decapsulate and unicast encapsulate to RTR 'y' ('x' or 'y' could be either IPv4 or IPv6 unicast RLOCs), which would decapsulate and unicast encapsulate to the final RLOC 'ETR-A'. The ETR 'ETR-A' would decapsulate and deliver the multicast packet natively to all the receivers joined to application group 'G' inside the LISP site.

Let's look at the ITR using the ELP (a, g, b, ETR-B). Here the encapsulation path would be the ITR unicast encapsulates to unicast RLOC 'a'. RTR 'a' multicast encapsulates to delivery group 'g'. The packet gets to all ETRs that have joined delivery group 'g' so they can deliver the multicast packet to joined receivers of application group 'G' in their sites. RTR 'b' is also joined to delivery group 'g'. Since it is in the ELP, it will be the only RTR that unicast encapsulates the multicast packet to ETR 'ETR-B'. Lastly, 'ETR-B' decapsulates and delivers the multicast packet to joined receivers to application group 'G' in its LISP site.

As one can see there are all sorts of opportunities to provide multicast connectivity across a network with non-congruent support

for multicast and different address-families. One can also see how using the mapping database can allow flexible forms of delivery policy, rerouting, and congestion control management in multicast environments.

10. Security Considerations

When an RTR receives a LISP encapsulated packet, it can look at the outer source address to verify that RLOC is the one listed as the previous hop in the ELP list. If the outer source RLOC address appears before the RLOC which matches the outer destination RLOC address, the decapsulating RTR (or ETR if last hop), MAY choose to drop the packet.

11. IANA Considerations

At this time there are no requests for IANA.

12. References

12.1. Normative References

[RFC0791] Postel, J., "Internet Protocol", STD 5, RFC 791, DOI 10.17487/RFC0791, September 1981, <http://www.rfc-editor.org/info/rfc791>.

[RFC1034] Mockapetris, P., "Domain names - concepts and facilities", STD 13, RFC 1034, DOI 10.17487/RFC1034, November 1987, <http://www.rfc-editor.org/info/rfc1034>.

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

[RFC2460] Deering, S. and R. Hinden, "Internet Protocol, Version 6 (IPv6) Specification", RFC 2460, DOI 10.17487/RFC2460, December 1998, <http://www.rfc-editor.org/info/rfc2460>.

[RFC3261] Rosenberg, J., Schulzrinne, H., Camarillo, G., Johnston, A., Peterson, J., Sparks, R., Handley, M., and E. Schooler, "SIP: Session Initiation Protocol", RFC 3261, DOI 10.17487/RFC3261, June 2002, <http://www.rfc-editor.org/info/rfc3261>.

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

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

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

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

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

12.2. Informative References

[I-D.ermagan-lisp-nat-traversal] Ermagan, V., Farinacci, D., Lewis, D., Skriver, J., Maino, F., and C. White, "NAT traversal for LISP", draft-ermagan-lisp-nat-traversal-11 (work in progress), August 2016.

Appendix A. Acknowledgments

The authors would like to thank the following people for their ideas and comments. They are Albert Cabellos, Khalid Raza, and Vina Ermagan, and Gregg Schudel.

Appendix B. Document Change Log

B.1. Changes to draft-farinacci-lisp-te-12.txt

- Posted February 2017.
- Update references and document timer.

B.2. Changes to draft-farinacci-lisp-te-11.txt

- Posted September 2016.
- Update references and document timer.

B.3. Changes to draft-farinacci-lisp-te-10.txt

- Posted March 2016.
- Update references and document timer.

B.4. Changes to draft-farinacci-lisp-te-09.txt

- Posted September 2015.
- Update references and document timer.

B.5. Changes to draft-farinacci-lisp-te-08.txt

- Posted March 2015.
- Update references and document timer.

B.6. Changes to draft-farinacci-lisp-te-07.txt

- Posted September 2014.
- Update references and document timer.

B.7. Changes to draft-farinacci-lisp-te-06.txt

- Posted March 2014.
- Fix Parantap’s affliation to self.

B.8. Changes to draft-farinacci-lisp-te-05.txt

- Posted March 2014.
- Fix text in "Using Recursion" section based on comment Jinghao Wang provided.

B.9. Changes to draft-farinacci-lisp-te-04.txt

- Resubmitted draft due to document timeout.
- Updated Informative References section.

B.10. Changes to draft-farinacci-lisp-te-03.txt

- Update LISP references to their RFC pointers and document timer.

B.11. Changes to draft-farinacci-lisp-te-02.txt

- Update references and document timer.

B.12. Changes to draft-farinacci-lisp-te-01.txt

- Posted July 2012.
- Add the Lookup bit to allow an ELP to be a list of encapsulation and/or mapping database lookup addresses.
- Indicate that ELPs can be used for service chaining.
- Add text to indicate that Map-Notify messages can be sent to new RTRs in a ELP so their map-caches can be pre-populated to avoid mapping database lookup packet loss.
- Fixes to editorial comments from Gregg.

B.13. Changes to draft-farinacci-lisp-te-00.txt

- Initial draft posted March 2012.

Authors' Addresses

Dino Farinacci
lispers.net
San Jose, California
USA

Phone: 408-718-2001
Email: farinacci@gmail.com

Michael Kowal
cisco Systems
111 Wood Avenue South
ISELIN, NJ
USA

Email: mikowal@cisco.com

Parantap Lahiri
USA

Email: parantap.lahiri@gmail.com

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: June 29, 2017

D. Farinacci
lispers.net
D. Lewis
cisco Systems
D. Meyer
1-4-5.net
C. White
Logical Elegance, LLC.
December 26, 2016

LISP Mobile Node
draft-meyer-lisp-mn-16

Abstract

This document describes how a lightweight version of LISP's ITR/ETR functionality can be used to provide seamless mobility to a mobile node. The LISP Mobile Node design described in this document uses standard LISP functionality to provide scalable mobility for LISP mobile nodes.

Status of This Memo

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

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

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

This Internet-Draft will expire on June 29, 2017.

Copyright Notice

Table of Contents

1. Introduction

The Locator/ID Separation Protocol (LISP) [RFC6830] specifies a design and mechanism for replacing the addresses currently used in the Internet with two separate name spaces: Endpoint Identifiers (EIDs), used within sites, and Routing Locators (RLOCs), used by the transit networks that make up the Internet infrastructure. To achieve this separation, LISP defines protocol mechanisms for mapping from EIDs to RLOCs. The mapping infrastructure is comprised of LISP Map-Servers and Map-Resolvers [RFC6833] and is tied together with LISP+ALT [RFC6836].

This document specifies the behavior of a new LISP network element: the LISP Mobile Node. The LISP Mobile Node implements a subset of the standard Ingress Tunnel Router and Egress Tunnel Router functionality [RFC6830]. Design goals for the LISP mobility design include:

- Allowing TCP connections to stay alive while roaming.
- Allowing the mobile node to communicate with other mobile nodes while either or both are roaming.
- Allowing the mobile node to multi-home (i.e., use multiple interfaces concurrently).
- Allowing the mobile node to be a server. That is, any mobile node or stationary node can find and connect to a mobile node as a server.
- Providing shortest path bidirectional data paths between a mobile node and any other stationary or mobile node.
- Not requiring fine-grained routes in the core network to support mobility.
- Not requiring a home-agent, foreign agent or other data plane network elements to support mobility. Note since the LISP mobile node design does not require these data plane elements, there is no triangle routing of data packets as is found in Mobile IP [RFC3344].
- Not requiring new IPv6 extension headers to avoid triangle routing [RFC3775].

The LISP Mobile Node design requires the use of the LISP Map-Server [RFC6836] and LISP Interworking [RFC6832] technology to allow a LISP mobile node to roam and to be discovered in an efficient and scalable

manner. The use of Map-Server technology is discussed further in Section 5.

The protocol mechanisms described in this document apply those cases in which a node's IP address changes frequently. For example, when a mobile node roams, it is typically assigned a new IP address. Similarly, a broadband subscriber may have its address change frequently; as such, a broadband subscriber can use the LISP Mobile Node mechanisms defined in this specification.

The remainder of this document is organized as follows: Section 2 defines the terms used in this document. Section 3 provides a overview of salient features of the LISP Mobile Node design, and Section 4 describes design requirements for a LISP Mobile Node. Section 5 provides the detail of LISP Mobile Node data and control plane operation, and Section 6 discusses options for updating remote caches in the presence of unidirectional traffic flows. Section 7 specifies how the LISP Mobile Node protocol operates. Section 8 specifies multicast operation for LISP mobile nodes. Section 9 and Section 12 outline other considerations for the LISP-MN design and implementation. Finally, Section 13 outlines the security considerations for a LISP mobile node.

2. Definition of Terms

This section defines the terms used in this document.

Stationary Node (SN): A non-mobile node who's IP address changes infrequently. That is, its IP address does not change as frequently as a fast roaming mobile hand-set or a broadband connection and therefore the EID to RLOC mapping is relatively static.

Endpoint ID (EID): This is the traditional LISP EID [RFC6830], and is the address that a LISP mobile node uses as its address for transport connections. A LISP mobile node never changes its EID, which is typically a /32 or /128 prefix and is assigned to a loopback interface. Note that the mobile node can have multiple EIDs, and these EIDs can be from different address families.

Routing Locator (RLOC): This is the traditional LISP RLOC, and is in general a routable address that can be used to reach a mobile node. Note that there are cases in which an mobile node may receive an address that it thinks is an RLOC (perhaps via DHCP) which is either an EID or an RFC 1918 address [RFC1918]. This could happen if, for example, if the mobile node roams into a LISP domain or a domain behind a Network Address Translator (NAT)) See Section 10 for more details.

Ingress Tunnel Router (ITR): An ITR is a router that accepts an IP packet with a single IP header (more precisely, an IP packet that does not contain a LISP header). The router treats this "inner" IP destination address as an EID and performs an EID-to-RLOC mapping lookup. The router then prepends an "outer" IP header with one of its globally routable RLOCs in the source address field and the result of the mapping lookup in the destination address field. Note that this destination RLOC may be an intermediate, proxy device that has better knowledge of the EID-to-RLOC mapping closer to the destination EID. In general, an ITR receives IP packets from site end-systems on one side and sends LISP-encapsulated IP packets toward the Internet on the other side. A LISP mobile node, however, when acting as an ITR LISP encapsulates all packet that it originates.

Egress Tunnel Router (ETR): An ETR is a router that accepts an IP packet where the destination address in the "outer" IP header is one of its own RLOCs. The router strips the "outer" header and forwards the packet based on the next IP header found. In general, an ETR receives LISP-encapsulated IP packets from the Internet on one side and sends decapsulated IP packets to site end-systems on the other side. A LISP mobile node, when acting as an ETR, decapsulates packets that are then typically processed by the mobile node.

Proxy Ingress Tunnel Router (PITR): PITRs are used to provide interconnectivity between sites that use LISP EIDs and those that do not. They act as a gateway between the Legacy Internet and the LISP enabled Network. A given PITR advertises one or more highly aggregated EID prefixes into the public Internet and acts as the ITR for traffic received from the public Internet. Proxy Ingress Tunnel Routers are described in [RFC6832].

Proxy Egress Tunnel Router (PETR): An infrastructure element used to decapsulate packets sent from mobile nodes to non-LISP sites. Proxy Egress Tunnel Routers are described in [RFC6832].

LISP Mobile Node (LISP-MN): A LISP capable fast roaming mobile hand-set.

Map-cache: A data structure which contains an EID-prefix, its associated RLOCs, and the associated policy. Map-caches are typically found in ITRs and PITRs.

Negative Map-Reply: A Negative Map-Reply is a Map-Reply that contains a coarsely aggregated non-LISP prefix. Negative Map-Replies are typically generated by Map-Resolvers, and are used to inform an ITR (mobile or stationary) that a site is not a LISP

site. A LISP mobile node encapsulate packets to destinations covered by the negative Map-Reply are encapsulated to a PETR.

Roaming Event: A Roaming Event occurs when there is a change in a LISP mobile node's RLOC set.

3. Design Overview

The LISP-MN design described in this document uses the Map-Server/ Map-Resolver service interface in conjunction with a light-weight ITR/ETR implementation in the LISP-MN to provide scalable fast mobility. The LISP-MN control-plane uses a Map-Server as an anchor point, which provides control-plane scalability. In addition, the LISP-MN data-plane takes advantage of shortest path routing and therefore does not increase packet delivery latency.

4. Design Requirements

This section outlines the design requirements for a LISP-MN, and is divided into User Requirements (Section 4.1) and Network Requirements (Section 4.2).

4.1. User Requirements

This section describes the user-level functionality provided by a LISP-MN.

Transport Connection Survivability: The LISP-MN design must allow a LISP-MN to roam while keeping transport connections alive.

Simultaneous Roaming: The LISP-MN design must allow a LISP-MN to talk to another LISP-MN while both are roaming.

Multihoming: The LISP-MN design must allow for simultaneous use of multiple Internet connections by a LISP-MN. In addition, the design must allow for the LISP mobile node to specify ingress traffic engineering policies as documented in [RFC6830]. That is, the LISP-MN must be able to specify both active/active and active/ passive policies for ingress traffic.

Shortest Path Data Plane: The LISP-MN design must allow for shortest path bidirectional traffic between a LISP-MN and a stationary node, and between a LISP-MN and another LISP-MN (i.e., without triangle routing in the data path). This provides a low-latency data path between the LISP-MN and the nodes that it is communicating with.

4.2. Network Requirements

This section describes the network functionality that the LISP-MN design provides to a LISP-MN.

Routing System Scalability: The LISP-MN design must not require injection of fine-grained routes into the core network.

Mapping System Scalability: The LISP-MN design must not require additional state in the mapping system. In particular, any mapping state required to support LISP mobility must BE confined to the LISP-MN's Map-Server and the ITRs which are talking to the LISP-MN.

Component Reuse: The LISP-MN design must use existing LISP infrastructure components. These include map server, map resolver, and interworking infrastructure components.

Home Agent/Foreign Agent: The LISP-MN design must not require the use of home-agent or foreign-agent infrastructure components [RFC3344].

Readdressing: The LISP-MN design must not require TCP connections to be reset when the mobile node roams. In particular, since the IP address associated with a transport connection will not change as the mobile node roams, TCP connections will not reset.

5. LISP Mobile Node Operation

The LISP-MN design is built from three existing LISP components: A lightweight LISP implementation that runs in an LISP-MN, and the existing Map-Server [RFC6833] and Interworking [RFC6832] infrastructures. A LISP mobile node typically sends and receives LISP encapsulated packets (exceptions include management protocols such as DHCP).

The LISP-MN design makes a single mobile node look like a LISP site as described in in [RFC6830] by implementing ITR and ETR functionality. Note that one subtle difference between standard ITR behavior and LISP-MN is that the LISP-MN encapsulates all non-local, non-LISP site destined outgoing packets to a PETR.

When a LISP-MN roams onto a new network, it receives a new RLOC. Since the LISP-MN is the authoritative ETR for its EID-prefix, it must Map-Register it's updated RLOC set. New sessions can be established as soon as the registration process completes. Sessions that are encapsulating to RLOCs that did not change during the roaming event are not affected by the roaming event (or subsequent

mapping update). However, the LISP-MN must update the ITRs and PITRs that have cached a previous mapping. It does this using the techniques described in Section 6.

5.1. Addressing Architecture

A LISP-MN is typically provisioned with one or more EIDs that it uses for all transport connections. LISP-MN EIDs are provisioned from blocks reserved from mobile nodes much the way mobile phone numbers are provisioned today (such that they do not overlap with the EID space of any enterprise). These EIDs can be either IPv4 or IPv6 addresses. For example, one EID might be for a public network while another might be for a private network; in this case the "public" EID will be associated with RLOCs from the public Internet, while the "private" EID will be associated with private RLOCs. It is anticipated that these EIDs will change infrequently if at all, since the assignment of a LISP-MN's EID is envisioned to be a subscription time event. The key point here is that the relatively fixed EID allows the LISP-MN's transport connections to survive roaming events. In particular, while the LISP-MN's EIDs are fixed during roaming events, the LISP-MN's RLOC set will change. The RLOC set may be comprised of both IPv4 or IPv6 addresses.

A LISP-MN is also provisioned with the address of a Map-Server and a corresponding authentication key. Like the LISP-MN's EID, both the Map-Server address and authentication key change very infrequently (again, these are anticipated to be subscription time parameters). Since the LISP LISP-MN's Map-Server is configured to advertise an aggregated EID-prefix that covers the LISP-MN's EID, changes to the LISP-MN's mapping are not propagated further into the mapping system [RFC6836]. It is this property that provides for scalable fast mobility.

A LISP-MN is also be provisioned with the address of a Map-Resolver. A LISP-MN may also learn the address of a Map-Resolver though a dynamic protocol such as DHCP [RFC2131].

Finally, note that if, for some reason, a LISP-MN's EID is re-provisioned, the LISP-MN's Map-Server address may also have to change in order to keep LISP-MN's EID within the aggregate advertised by the Map-Server (this is discussed in greater detail in Section 5.2).

5.2. Control Plane Operation

A roaming event occurs when the LISP-MN receives a new RLOC. Because the new address is a new RLOC from the LISP-MN's perspective, it must update its EID-to-RLOC mapping with its Map-Server; it does this using the Map-Register mechanism described in [RFC6830].

A LISP-MN may want the Map-Server to respond on its behalf for a variety of reasons, including minimizing control traffic on radio links and minimizing battery utilization. A LISP-MN may instruct its Map-Server to proxy respond to Map-Requests by setting the Proxy-Map-Reply bit in the Map-Register message [RFC6830]. In this case the Map-Server responds with a non-authoritative Map-Reply so that an ITR or PITR will know that the ETR didn't directly respond. A Map-Server will proxy reply only for "registered" EID-prefixes using the registered EID-prefix mask-length in proxy replies.

Because the LISP-MN's Map-Server is pre-configured to advertise an aggregate covering the LISP-MN's EID prefix, the database mapping change associated with the roaming event is confined to the Map-Server and those ITRs and PITRs that may have cached the previous mapping.

5.3. Data Plane Operation

A key feature of LISP-MN control-plane design is the use of the Map-Server as an anchor point; this allows control of the scope to which changes to the mapping system must be propagated during roaming events.

On the other hand, the LISP-MN data-plane design does not rely on additional LISP infrastructure for communication between LISP nodes (mobile or stationary). Data packets take the shortest path to and from the LISP-MN to other LISP nodes; as noted above, low latency shortest paths in the data-plane is an important goal for the LISP-MN design (and is important for delay-sensitive applications like gaming and voice-over-IP). Note that a LISP-MN will need additional interworking infrastructure when talking to non-LISP sites [RFC6832]; this is consistent with the design of any host at a LISP site which talks to a host at a non-LISP site.

In general, the LISP-MN data-plane operates in the same manner as the standard LISP data-plane with one exception: packets generated by a LISP-MN which are not destined for the mapping system (i.e., those sent to destination UDP port 4342) or the local network are LISP encapsulated. Because data packets are always encapsulated to a RLOC, packets travel on the shortest path from LISP-MN to another LISP stationary or LISP-MN. When the LISP mobile node is sending packets to a stationary or LISP-MN in a non-LISP site, it sends LISP-encapsulated packets to a PETR which then decapsulates the packet and forwards it to its destination.

6. Updating Remote Caches

A LISP-MN has five mechanisms it can use to cause the mappings cached in remote ITRs and PITRs to be refreshed:

Map Versioning: If Map Versioning [RFC6834] is used, an ETR can detect if an ITR is using the most recent database mapping. In particular, when mobile node's ETR decapsulates a packet and detects the Destination Map-Version Number is less than the current version for its mapping, in invokes the SMR procedure described in [RFC6830]. In general, SMRs are used to fix the out of sync mapping while Map-Versioning is used to detect they are out of sync. [RFC6834] provides additional details of the Map Versioning process.

Data Driven SMRs: An ETR may elect to send SMRs to those sites it has been receiving encapsulated packets from. This will occur when an ITR is sending to an old RLOC (for which there is one-to-one mapping between EID-to-RLOC) and the ETR may not have had a chance to send an SMR the ITR.

Setting Small TTL on Map Replies: The ETR (or Map Server) may set a small Time to Live (TTL) on its mappings when responding to Map Requests. The TTL value should be chosen such that changes in mappings can be detected while minimizing control traffic. In this case the ITR is a SN and the ETR is the MN.

Piggybacking Mapping Data: If an ITR and ETR are co-located, an ITR may elect to send Map-Requests with piggybacked mapping data to those sites in its map cache or to which it has recently encapsulated data in order to inform the remote ITRs and PITRs of the change.

Temporary PITR Caching: The ETR can keep a cache of PITRs that have sent Map-Requests to it. The cache contains the RLOCs of the PITRs so later when the locator-set of a LISP-MN changes, SMR messages can be sent to all RLOCs in the PITR cache. This is an example of a control-plane driven SMR procedure.

7. Protocol Operation

There are five distinct connectivity cases considered by the LISP-MN design. The five mobility cases are:

LISP Mobile Node to a Stationary Node in a LISP Site.

LISP Mobile Node to a Non-LISP Site.

LISP Mobile Node to a LISP Mobile Node.

Non-LISP Site to a LISP Mobile Node.

LISP Site to a LISP Mobile Node.

The remainder of this section covers these cases in detail.

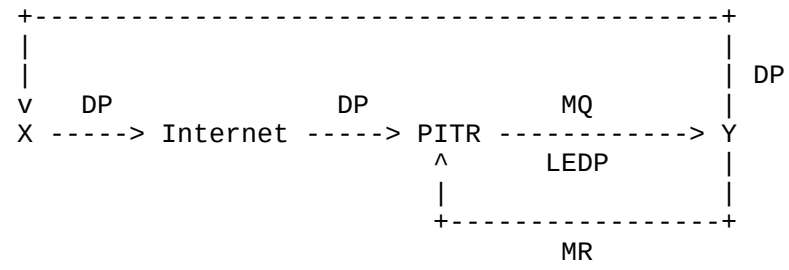
7.1. LISP Mobile Node to a Stationary Node in a LISP Site

After a roaming event, a LISP-MN must immediately register its new EID-to-RLOC mapping with its configured Map-Server(s). This allows LISP sites sending Map-Requests to the LISP-MN to receive the current mapping. In addition, remote ITRs and PITRs may have cached mappings that are no longer valid. These ITRs and PITRs must be informed that the mapping has changed. See Section 6 for a discussion of methods for updating remote caches.

7.1.1. Handling Unidirectional Traffic

A problem may arise when traffic is flowing unidirectionally between LISP sites. This can arise in communication flows between PITRs and LISP sites or when a site’s ITRs and ETRs are not co-located. In these cases, data-plane techniques such as Map-Versioning and Data-Driven SMRs can’t be used to update the remote caches.

For example, consider the unidirectional packet flow case depicted in Figure 1. In this case X is a non-LISP enabled SN (i.e., connected to the Internet) and Y is a LISP MN. Data traffic from X to Y will flow through a PITR. When Y changes its mapping (for example, during a mobility event), the PITR must update its mapping for Y. However, since data traffic from Y to X is unidirectional and does not flow though the PITR, it can not rely data traffic from Y to X to signal a mapping change at Y. In this case, the Y must use one or more of the techniques described in Section 6 to update the PITR’s cache. Note that if Y has only one RLOC, then the PITR has to know when to send a Map-Request based on its existing state; thus it can only rely on the TTL on the existing mapping.



DP: Data Packet
LEDP: LISP Encapsulated Data Packet
MQ: Map Request
MR: Map Reply

Figure 1: Unidirectional Packet Flow

7.2. LISP Mobile Node to a Non-LISP Stationary Node

LISP-MNs use the LISP Interworking infrastructure (specifically a PETR) to reach non-LISP sites. In general, the PETR will be co-located with the LISP-MN's Map-Server. This ensures that the LISP packets being decapsulated are from sources that have Map-Registered to the Map-Server. Note that when a LISP-MN roams it continues to uses its configured PETR and Map-Server which can have the effect of adding stretch to packets sent from a LISP-MN to a non-LISP destination.

7.3. LISP Mobile Node to LISP Mobile Node

LISP-MN to LISP-MN communication is an instance of LISP-to-LISP communication with three sub-cases:

- Both LISP-MNs are stationary (Section 7.1).
- Only one LISP-MN is roaming (Section 7.3.1).
- Both LISP-MNs are roaming. The case is analogous to the case described in Section 7.3.1.

7.3.1. One Mobile Node is Roaming

In this case, the roaming LISP-MN can find the stationary LISP-MN by sending Map-Request for its EID-prefix. After receiving a Map-Reply, the roaming LISP-MN can encapsulate data packets directly to the non-roaming LISP-MN node.

The roaming LISP-MN, on the other hand, must update its Map-Server with the new mapping data as described in Section 7.1. It should also use the cache management techniques described in Section 6 to provide for timely updates of remote caches. Once the roaming LISP-MN has updated its Map-Server, the non-roaming LISP-MN can retrieve the new mapping data (if it hasn't already received an updated mapping via one of the mechanisms described in Section 6) and the stationary LISP-MN can encapsulate data directly to the roaming LISP-MN.

7.4. Non-LISP Site to a LISP Mobile Node

When a stationary ITR is talking to a non-LISP site, it may forward packets natively (unencapsulated) to the non-LISP site. This will occur when the ITR has received a negative Map Reply for a prefix covering the non-LISP site's address with the Natively-Forward action bit set [RFC6830]. As a result, packets may be natively forwarded to non-LISP sites by an ITR (the return path will through a PITR, however, since the packet flow will be non-LISP site to LISP site).

A LISP-MN behaves differently when talking to non-LISP sites. In particular, the LISP-MN always encapsulates packets to a PETR. The PETR then decapsulates the packet and forwards it natively to its destination. As in the stationary case, packets from the non-LISP site host return to the LISP-MN through a PITR. Since traffic forwarded through a PITR is unidirectional, a LISP-MN should use the cache management techniques described in Section 7.1.1.

7.5. LISP Site to LISP Mobile Node

When a LISP-MN roams onto a new network, it needs to update the caches in any ITRs that might have stale mappings. This is analogous to the case in that a stationary LISP site is renumbered; in that case ITRs that have cached the old mapping must be updated. This is done using the techniques described in Section 6.

When a LISP router in a stationary site is performing both ITR and ETR functions, a LISP-MN can update the stationary site's map-caches using techniques described in Section 6. However, when the LISP router in the stationary site is performing is only ITR functionality, these techniques can not be used because the ITR is not receiving data traffic from the LISP-MN. In this case, the LISP-MN should use the technique described in Section 7.1.1. In particular, a LISP-MN should set the TTL on the mappings in its Map-Replies to be in 1-2 minute range.

8. Multicast and Mobility

Since a LISP-MN performs both ITR and ETR functionality, it should also perform a lightweight version of multicast ITR/ETR functionality described in [RFC6831]. When a LISP-MN originates a multicast packet, it will encapsulate the packet with a multicast header, where the source address in the outer header is one of it's RLOC addresses and the destination address in the outer header is the group address from the inner header. The interfaces in which the encapsulated packet is sent on is discussed below.

To not require PIM functionality in the LISP-MN as documented in [RFC6831], the LISP-MN resorts to using encapsulated IGMP for joining groups and for determining which interfaces are used for packet origination. When a LISP-MN joins a group, it obtains the map-cache entry for the (S-EID,G) it is joining. It then builds a IGMP report encoding (S-EID,G) and then LISP encapsulates it with UDP port 4341. It selects an RLOC from the map-cache entry to send the encapsulated IGMP Report.

When other LISP-MNs are joining an (S-EID,G) entry where the S-EID is for a LISP-MN, the encapsulated IGMP Report will be received by the LISP-MN multicast source. The LISP-MN multicast source will remember the interfaces the encapsulated IGMP Report is received on and build an outgoing interface list for it's own (S-EID,G) entry. If the list is greater than one, then the LISP-MN is doing replication on the source-based tree for which it is the root.

When other LISP routers are joining (S-EID,G), they are instructed to send PIM encapsulated Join-Prune messages. However, to keep the LISP-MN as simple as possible, the LISP-MN will not be able to process encapsulated PIM Join-Prune messages. Because the map-cache entry will have a MN-bit indicating the entry is for a LISP-MN, the LISP router will send IGMP encapsulated IGMP Reports instead.

When the LISP-MN is sending a multicast packet, it can operate in two modes, multicast-origination-mode or unicast-origination-mode. When in multicast-origination-mode, the LISP-MN multicast-source can encapsulate a multicast packet in another multicast packet, as described above. When in unicast-origination-mode, the LISP-MN multicast source encapsulates the multicast packet into a unicast packet and sends a packet to each encapsulated IGMP Report sender.

These modes are provided depending on whether or not the mobile node's network it is currently connected can support IP multicast.

9. RLOC Considerations

This section documents cases where the expected operation of the LISP-MN design may require special treatment.

9.1. Mobile Node's RLOC is an EID

When a LISP-MN roams into a LISP site, the "RLOC" it is assigned may be an address taken from the site's EID-prefix. In this case, the LISP-MN will Map-Register a mapping from its statically assigned EID to the "RLOC" it received from the site. This scenario creates another level of indirection: the mapping from the LISP-MN's EID to a site assigned EID. The mapping from the LISP-MN's EID to the site assigned EID allow the LISP-MN to be reached by sending packets using the mapping for the EID; packets are delivered to site's EIDs use the same LISP infrastructure that all LISP hosts use to reach the site.

A packet egressing a LISP site destined for a LISP-MN that resides in a LISP site will have three headers: an inner header that is built by the host and is used by transport connections, a middle header that is built by the site's ITR and is used by the destination's ETR to find the current topological location of the LISP-MN, and an outer header (also built by the site's ITR) that is used to forward packets between the sites.

Consider a site A with EID-prefix 1.0.0.0/8 and RLOC A and a site B with EID-prefix 2.0.0.0/8 and RLOC B. Suppose that a host S in site A with EID 1.0.0.1 wants to talk to a LISP LISP-MN MN that has registered a mapping from EID 240.0.0.1 to "RLOC" 2.0.0.2 (where 2.0.0.2 allocated from site B's EID prefix, 2.0.0.0/8 in this case). This situation is depicted in Figure 2.

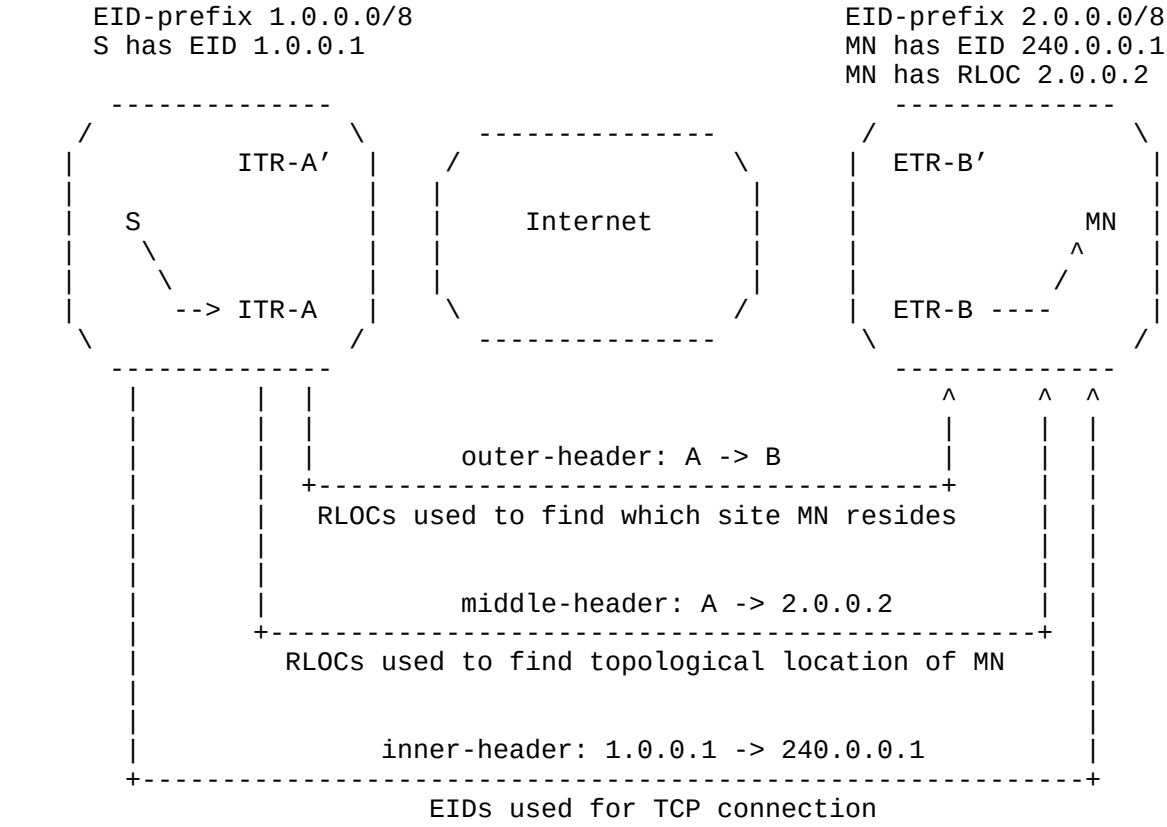


Figure 2: Mobile Node Roaming into a LISP Site

In this case, the inner header is used for transport connections, the middle header is used to find topological location of the LISP-MN (the LISP-MN Map-Registers the mapping 240.0.0.1 -> 2.0.0.2 when it roams into site B), and the outer header is used to move packets between sites (A and B in Figure 2).

In summary, when a LISP-MN roams into a LISP site and receives a new address (e.g., via DHCP) that is part of the site's EID space, the following sequence occurs:

1. The LISP-MN in the LISP site (call it Inside) registers its new RLOC (which is actually part of the sites EID prefix) to its map-server. Call its permanent EID E and the EID it DHCPs D. So it registers a mapping that looks like E->D.

2. The MN which is outside (call it Outside) sends a map request for inside's EID (E) and receives D (plus its policy). Outside realizes that D is an EID and sends a map request for D. This will return the site's RLOCs (by its ETR). Call this R.

3. Outside then double encapsulates the outbound packet with the inner destination being D and the outer destination being R.

4. The packet then finds its way to R, which strips the outer header and the packet is routed to D in the domain to Inside. Inside decapsulates the packet to serve the inner header to the application.

Note that both D and R could be returned to Inside in one query, so as not to incur the additional RTT.

10. LISP Mobile Nodes behind NAT Devices

When a LISP-MN resides behind a NAT device, it will be allocated a private RLOC address. The private RLOC address is used as the source address in the outer header for LISP encapsulated packets. The NAT device will translate the source address and source UDP port in the LISP encapsulated packet. The NAT device will keep this translated state so when packets arrive from the public side of the NAT, they can be translated back to the stored state. For remote LISP ITRs, PITRs, and RTRs, will need to know the translated RLOC address and port so they can encapsulate to the LISP-MN traversing the NAT device.

Procedures a LISP-MN should follow when it resides behind a NAT, will follow the LISP xTRs procedures in specification [I-D.ermagan-lisp-nat-traversal].

11. Mobility Example

This section provides an example of how the LISP-MN is integrated into the base LISP Design [RFC6830].

11.1. Provisioning

The LISP-MN needs to be configured with the following information:

An EID, assigned to its loopback address

A key for map-registration

An IP address of a Map-Resolver (this could be learned dynamically)

An IP address of its Map-Server and Proxy ETR

11.2. Registration

After a LISP roams to a new network, it must immediately register its new mapping this new RLOC (and associated priority/weight data) with its Map-Server.

The LISP-MN may chose to set the 'proxy' bit in the map-register to indicate that it desires its Map-Server to answer map-requests on its behalf.

12. LISP Implementation in a Mobile Node

This section will describe a possible approach for developing a lightweight LISP-MN implementation. A LISP-MN will implement a LISP sub-layer inside of the IP layer of the protocol stack. The sub-layer resides between the IP layer and the link-layer.

For outgoing unicast packets, once the header that contains EIDs is built and right before an outgoing interface is chosen, a LISP header is prepended to the outgoing packet. The source address is set to the local RLOC address (obtained by DHCP perhaps) and the destination address is set to the RLOC associated with the destination EID from the IP layer. To obtain the RLOC for the EID, the LISP-MN maintains a map-cache for destination sites or destination LISP-MNs to which it is currently talking. The map-cache lookup is performed by doing a longest match lookup on the destination address the IP layer put in the first IP header. Once the new header is prepended, a route table lookup is performed to find the interface in which to send the packet or the default router interface is used to send the packet.

When the map-cache does not exist for a destination, the mobile node may queue or drop the packet while it sends a Map-Request to it's configured Map-Resolver. Once a Map-Reply is returned, the map-cache entry stores the EID-to-RLOC state. If the RLOC state is empty in the Map-Reply, the Map-Reply is known as a Negative Map-Reply in which case the map-cache entry is created with a single RLOC, the RLOC of the configured Map-Server for the LISP-MN. The Map-Server that serves the LISP-MN also acts as a Proxy ETR (PETR) so packets can get delivered to hosts in non-LISP sites to which the LISP-MN is sending.

For incoming unicast packets, the LISP sub-layer simply decapsulates the packets and delivers to the IP layer. The loc-reach-bits can be processed by the LISP sub-layer. Specifically, the source EID from the packet is looked up in the map-cache and if the loc-reach-bits

settings have changed, store the loc-reach-bits from the packet and note which RLOCs for the map-cache entry should not be used.

In terms of the LISP-MN detecting which RLOCs from each stored map-cache entry is reachable, it can use any of the Locator Reachability Algorithms from [RFC6830].

A background task that runs off a timer should be run so the LISP-MN can send periodic Map-Register messages to the Map-Server. The Map-Register message should also be triggered when the LISP-MN detects a change in IP address for a given interface. The LISP-MN should send Map-Registers to the same Map-Register out each of it's operational links. This will provide for robustness on radio links with which the mobile node is associated.

A LISP-MN receives a Map-Request when it has Map-Registered to a Map-Server with the Proxy-bit set to 0. This means that the LISP-MN wishes to send authoritative Map-Replies for Map-Requests that are targeted at the LISP-MN. If the Proxy-bit is set when the LISP-MN registers, then the Map-Server will send non-authoritative Map-Replies on behalf of the LISP-MN. In this case, the Map-Server never encapsulates Map-Requests to the LISP-MN. The LISP-MN can save resources by not receiving Map-Requests (note that the LISP-MN will receive SMRs which have the same format as Map-Requests).

To summarize, a LISP sub-layer should implement:

- Encapsulating and decapsulating data packets.
- Sending and receiving of Map-Request control messages.
- Receiving and optionally sending Map-Replies.
- Sending Map-Register messages periodically.

The key point about the LISP sub-layer is that no other components in the protocol stack need changing; just the insertion of this sub-layer between the IP layer and the interface layer-2 encapsulation/decapsulation layer.

13. Security Considerations

Security for the LISP-MN design builds upon the security fundamentals found in LISP [RFC6830] for data-plane security and the LISP Map Server [RFC6833] registration security. Security issues unique to the LISP-MN design are considered below.

13.1. Proxy ETR Hijacking

The Proxy ETR (or PETR) that a LISP-MN uses as its destination for non-LISP traffic must use the security association used by the registration process outlined in Section 5.2 and explained in detail in the LISP-MS specification [RFC6833]. These measures prevent third party injection of LISP encapsulated traffic into a Proxy ETR. Importantly, a PETR must not decapsulate packets from non-registered RLOCs.

13.2. LISP Mobile Node using an EID as its RLOC

For LISP packets to be sent to a LISP-MN which has an EID assigned to it as an RLOC as described in Section 9.1), the LISP site must allow for incoming and outgoing LISP data packets. Firewalls and stateless packet filtering mechanisms must be configured to allow UDP port 4341 and UDP port 4342 packets.

14. Acknowledgments

Albert Cabellos, Noel Chiappa, Pierre Francois, Michael Menth, Andrew Partan, Chris White and John Zwiebel provided insightful comments on the mobile node concept and on this document. A special thanks goes to Mary Nickum for her attention to detail and effort in editing early versions of this document.

15. IANA Considerations

This document creates no new requirements on IANA namespaces [RFC5226].

16. References

16.1. Normative References

[RFC1918] Rekhter, Y., Moskowitz, B., Karrenberg, D., de Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, DOI 10.17487/RFC1918, February 1996, <http://www.rfc-editor.org/info/rfc1918>.

[RFC2131] Droms, R., "Dynamic Host Configuration Protocol", RFC 2131, DOI 10.17487/RFC2131, March 1997, <http://www.rfc-editor.org/info/rfc2131>.

[RFC3344] Perkins, C., Ed., "IP Mobility Support for IPv4", RFC 3344, DOI 10.17487/RFC3344, August 2002, <http://www.rfc-editor.org/info/rfc3344>.

[RFC3775] Johnson, D., Perkins, C., and J. Arkko, "Mobility Support in IPv6", RFC 3775, DOI 10.17487/RFC3775, June 2004, <http://www.rfc-editor.org/info/rfc3775>.

[RFC5226] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 5226, DOI 10.17487/RFC5226, May 2008, <http://www.rfc-editor.org/info/rfc5226>.

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

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

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

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

[RFC6834] Iannone, L., Saucez, D., and O. Bonaventure, "Locator/ID Separation Protocol (LISP) Map-Versioning", RFC 6834, DOI 10.17487/RFC6834, January 2013, <http://www.rfc-editor.org/info/rfc6834>.

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

16.2. Informative References

[I-D.ermagan-lisp-nat-traversal] Ermagan, V., Farinacci, D., Lewis, D., Skriver, J., Maino, F., and C. White, "NAT traversal for LISP", draft-ermagan-lisp-nat-traversal-11 (work in progress), August 2016.

Authors' Addresses

Dino Farinacci
lispers.net
San Jose, CA 95134
USA

Email: farinacci@gmail.com

Darrel Lewis
cisco Systems
Tasman Drive
San Jose, CA 95134
USA

Email: darlewis@cisco.com

David Meyer
1-4-5.net
USA

Email: dmm@1-4-5.net

Chris White
Logical Elegance, LLC.
San Jose, CA 95134
USA

Email: chris@logicalelegance.com

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: July 21, 2017

V. Moreno
Cisco Systems
D. Farinacci
lispers.net
January 17, 2017

LISP Virtual Private Networks (VPNs)
draft-moreno-lisp-vpn-00

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 http://datatracker.ietf.org/drafts/current/.

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

This Internet-Draft will expire on July 21, 2017.

Table of Contents

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 [I-D.ietf-lisp-signal-free-multicast].

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-Provider-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) [I-D.ietf-lisp-ddt]. The LISP IID is used in the data plane of the LISP header [RFC6830], as well as in the LISP control plane [I-D.ietf-lisp-lcaf].

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:

- 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.

- 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 [I-D.ietf-lisp-lcaf]. 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 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 Provider 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 provider VPN. There may be multiple provider VPNs in a LISP network and a VPN may subscribe to multiple provider VPNs. A subscriber VPN may act as a provider VPN to provide reachability across subscriber VPNs, this effectively merges the subscriber VPNs together, a scenario that is usually better achieved by creating a single 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 4.1.3.

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 4.1.3. 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.

4.1.3. 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-Register, Map-Reply or Map-Notify messages as defined in [RFC6830].

The additional RLOC containing the Home-IID should use AFI = 16387 (LCAF) with a List type as described in Section 4.1.3.1.

4.1.3.1. 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 [I-D.ietf-lisp-lcaf]. 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>
```

4.1.3.2. 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 [I-D.ietf-lisp-signal-free-multicast] 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 [I-D.ietf-lisp-signal-free-multicast] 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 [I-D.ietf-lisp-signal-free-multicast].

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 [I-D.ietf-lisp-signal-free-multicast] 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 need to carry the IID for the VPNs of senders. Since the sender’s VPN is not known, the ETR must send the SMR to the sending RLOC but replicated to all VPNs defined in the Extranet Policy. Multicast optimizations could be used to minimize the amount of traffic replicated when sending these SMRs and potentially replicate only at the ITR. An SMR traveling from an Extranet Subscriber VPN to an Extranet VPN will usually be less susceptible to being replicated many times than an SMR traveling in the opposite direction (provider to subscriber).

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 [I-D.ietf-lisp-crypto] 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 [I-D.ietf-lisp-crypto] 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, 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, <http://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, <http://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, <http://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, <http://www.rfc-editor.org/info/rfc4607>.

8.2. Informative References

[I-D.farinacci-lisp-crypto] Farinacci, D., "LISP Data-Plane Confidentiality", draft-farinacci-lisp-crypto-01 (work in progress), July 2014.

[I-D.farinacci-lisp-mr-signaling] Farinacci, D. and M. Napierala, "LISP Control-Plane Multicast Signaling", draft-farinacci-lisp-mr-signaling-06 (work in progress), February 2015.

[I-D.ietf-lisp-crypto]
Farinacci, D. and B. Weis, "LISP Data-Plane Confidentiality", draft-ietf-lisp-crypto-10 (work in progress), October 2016.

[I-D.ietf-lisp-ddt]
Fuller, V., Lewis, D., Ermagan, V., Jain, A., and A. Smirnov, "LISP Delegated Database Tree", draft-ietf-lisp-ddt-08 (work in progress), September 2016.

[I-D.ietf-lisp-lcaf]
Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", draft-ietf-lisp-lcaf-22 (work in progress), November 2016.

[I-D.ietf-lisp-sec]
Maino, F., Ermagan, V., Cabellos-Aparicio, A., and D. Saucez, "LISP-Security (LISP-SEC)", draft-ietf-lisp-sec-12 (work in progress), November 2016.

[I-D.ietf-lisp-signal-free-multicast]
Moreno, V. and D. Farinacci, "Signal-Free LISP Multicast", draft-ietf-lisp-signal-free-multicast-02 (work in progress), October 2016.

[RFC6407] Weis, B., Rowles, S., and T. Hardjono, "The Group Domain of Interpretation", RFC 6407, DOI 10.17487/RFC6407, October 2011, <http://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, <http://www.rfc-editor.org/info/rfc6830>.

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

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

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

Authors' Addresses

Victor Moreno
Cisco Systems
170 Tasman Drive
San Jose, California 95134
USA

Email: vimoreno@cisco.com

Dino Farinacci
lispers.net
San Jose, CA 95120
USA

Email: farinacci@gmail.com

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

M. Portoles
V. Ashtaputre
V. Moreno
F. Maino
Cisco Systems
D. Farinacci
lispers.net

October 8, 2016

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

Abstract

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

Requirements Language

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

Status of This Memo

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

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

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

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

Table of Contents

1. Introduction

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

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

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

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

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

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

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).

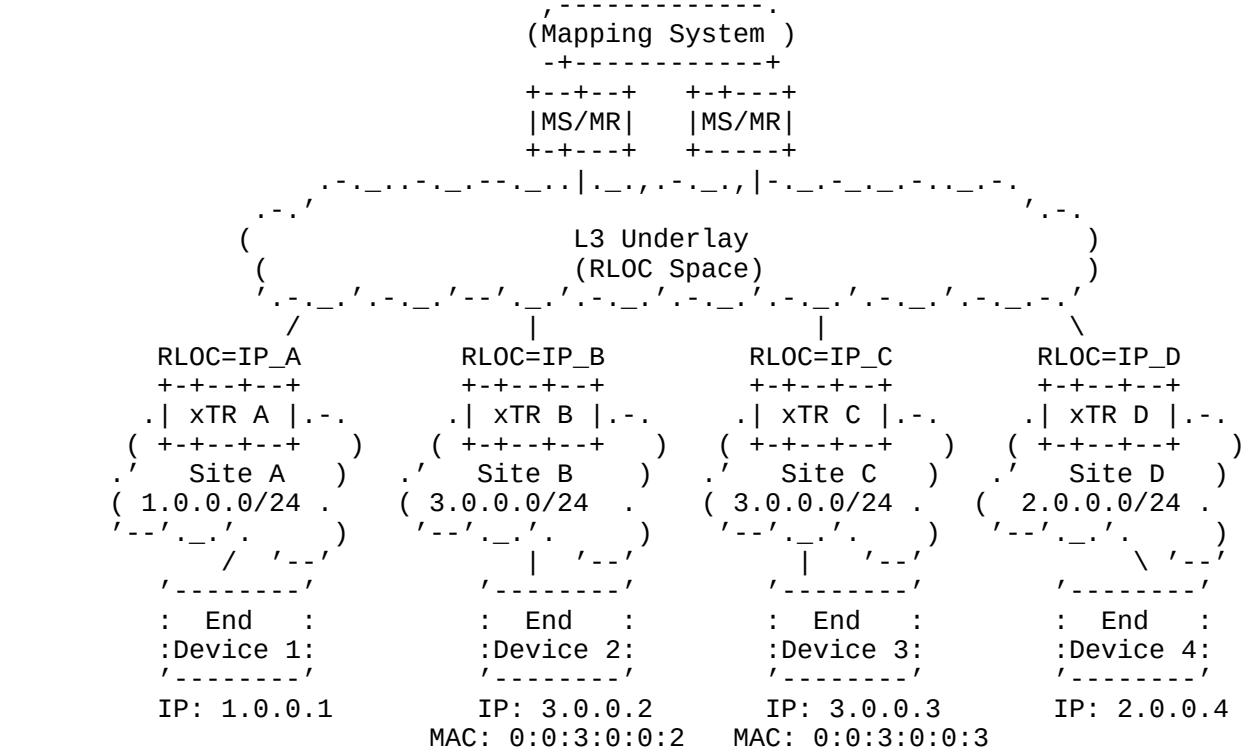


Figure 1: Reference System Architecture with unified L2 and L3 overlays

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.

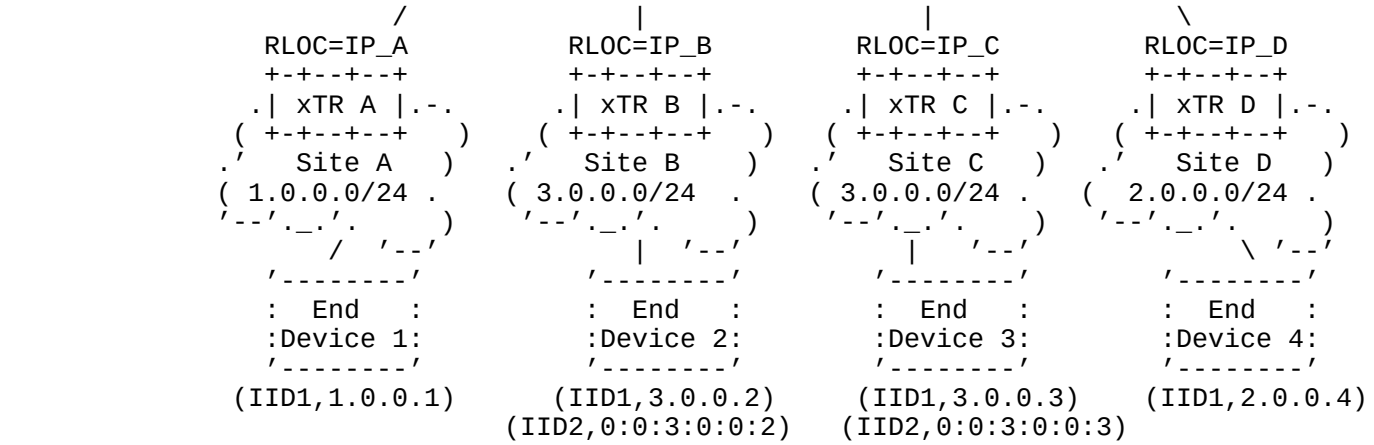


Figure 2: Reference Mobility Architecture

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 the original specification [RFC6830]. We describe the packet flow here for completeness

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

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

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

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

4.2. Implementation Considerations

4.2.1. L3 Segmentation

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

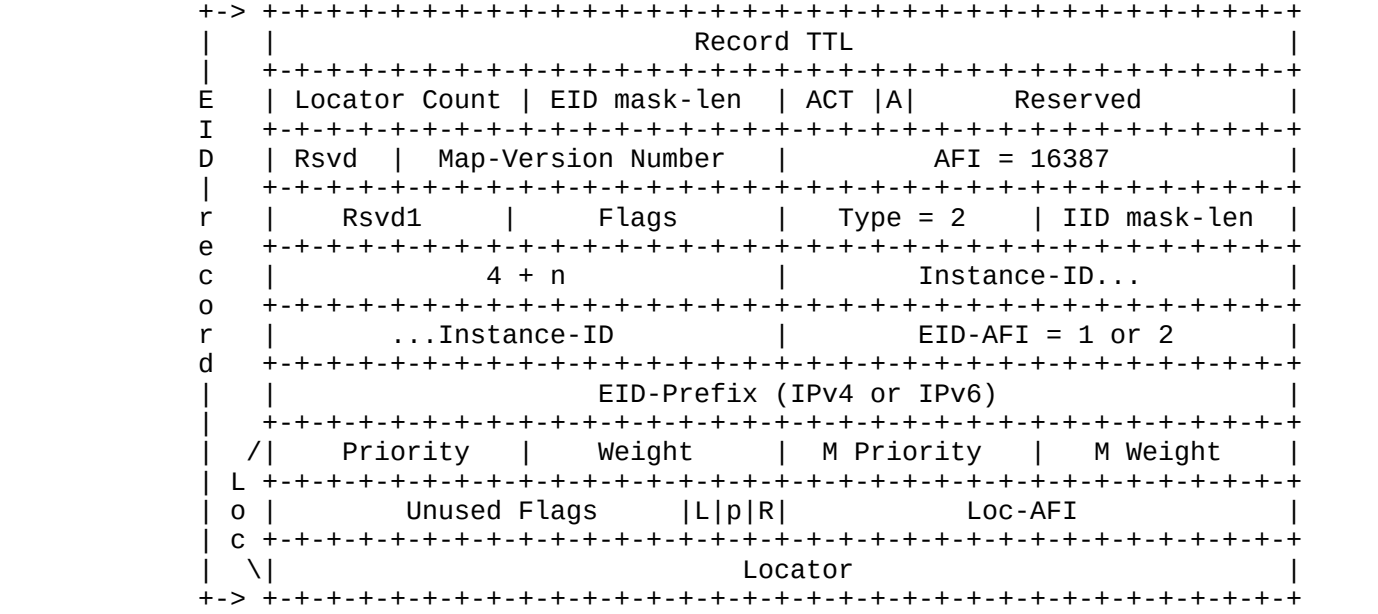
Instance-IDs can be used to support L3 overlay segmentation, such as in extended VRFs or multi-VPN overlays.

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 [I-D.ietf-lisp-lcaf] and the specific EID record to be used is illustrated in the following figure:



The L3 EID record follows the structure as described in [RFC6830].

4.2.3. LISP Mapping System support

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

In order to support system convergence after mobility, when the Map-Server receives a new registration for a specific EID, it MUST send a Map-Notify to the entire RLOC set in the site that last registered

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

4.2.4. Using SMRs to Track Moved-Away Hosts

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

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

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

- 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], [I-D.ietf-lisp-signal-free-multicast] and [I-D.coras-lisp-re].

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

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

5. L2 Overlays and Mobility Support

5.1. Reference Architecture and packet flows

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

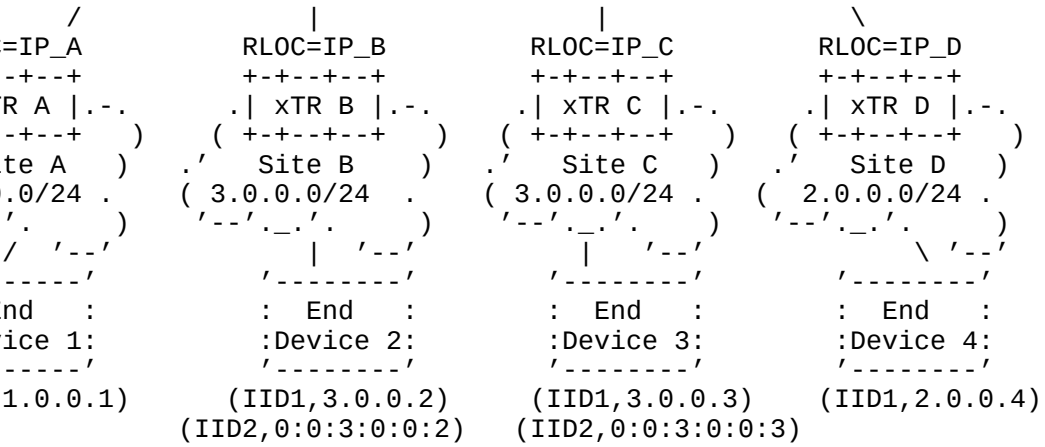


Figure 3: Reference Mobility Architecture

5.1.1. Bridged Traffic Flow: L2 Overlay use

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

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

- o ITR B does a L2 lookup in its local map-cache for the destination MAC 0:0:3:0:0:3. When the lookup of 0:0:3:0:0:3 is a miss, the ITR sends a Map-Request to the mapping database system looking up for EID=<IID2,0:0:3:0:0:3>.

- o The mapping systems forwards the Map-Request to ETR C, that has registered the EID-to-RLOC mapping for EID=<IID2,0:0:3:0:0:3>. Alternatively, depending on the mapping system configuration, a Map-Server which is part of the mapping database system MAY send a Map-Reply directly to ITR B.

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

- 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.

5.2. Implementation Considerations

5.2.1. L2 Segmentation

As with L3 overlays, LISP support of L2 segmentation is structured around the propagation and use of Instance-IDs, and handled as part of the EID in control plane operations. The encoding is described in [I-D.ietf-lisp-lcaf] and its use in [I-D.ietf-lisp-ddt]. Instance-IDs are unique to a Mapping System and MAY be used to identify the overlay type (e.g., L2 or L3 overlay).

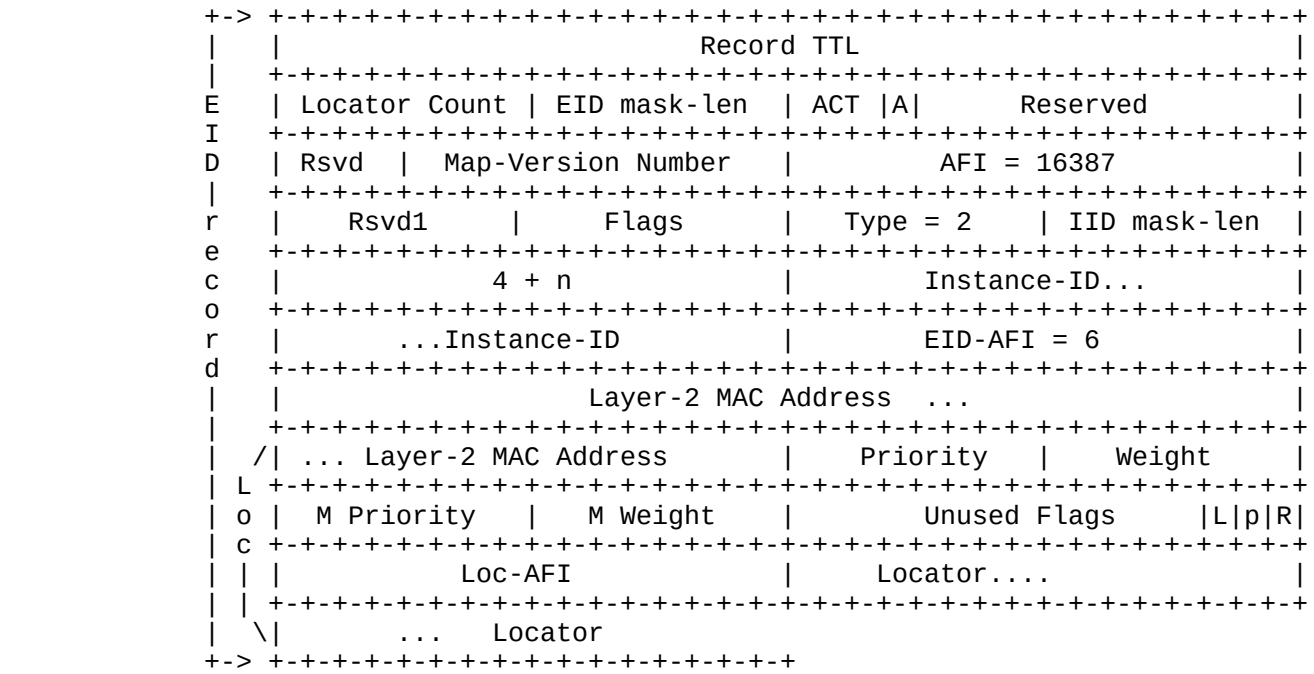
An Instance-ID can be used for L2 overlay segmentation. An important aspect of L2 segmentation is the mapping of VLANs to IIDs. In this case a Bridge Domain (which is the L2 equivalent to a VRF as a forwarding context) maps to an IID, a VLAN-ID may map 1:1 to a bridge domain or different VLAN-IDs on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

5.2.2. L2 Database-Mappings

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

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

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



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

5.2.3. Interface to the LISP Mapping System

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

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

5.2.4. SMR support to track L2 hosts that moved away

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

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

- 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 [I-D.ietf-lisp-signal-free-multicast], where S is 0000-0000-0000/0 and G is ffff-ffff-ffff/48. This strategy allows and ITR to know which ETRs are part of the L2 overlay and it can head-end replicate traffic to.

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

5.2.6. L2 Unknown Unicast Support

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

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

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

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

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

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

5.3. Support to ARP resolution through Mapping System

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

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

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

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

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

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

- 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

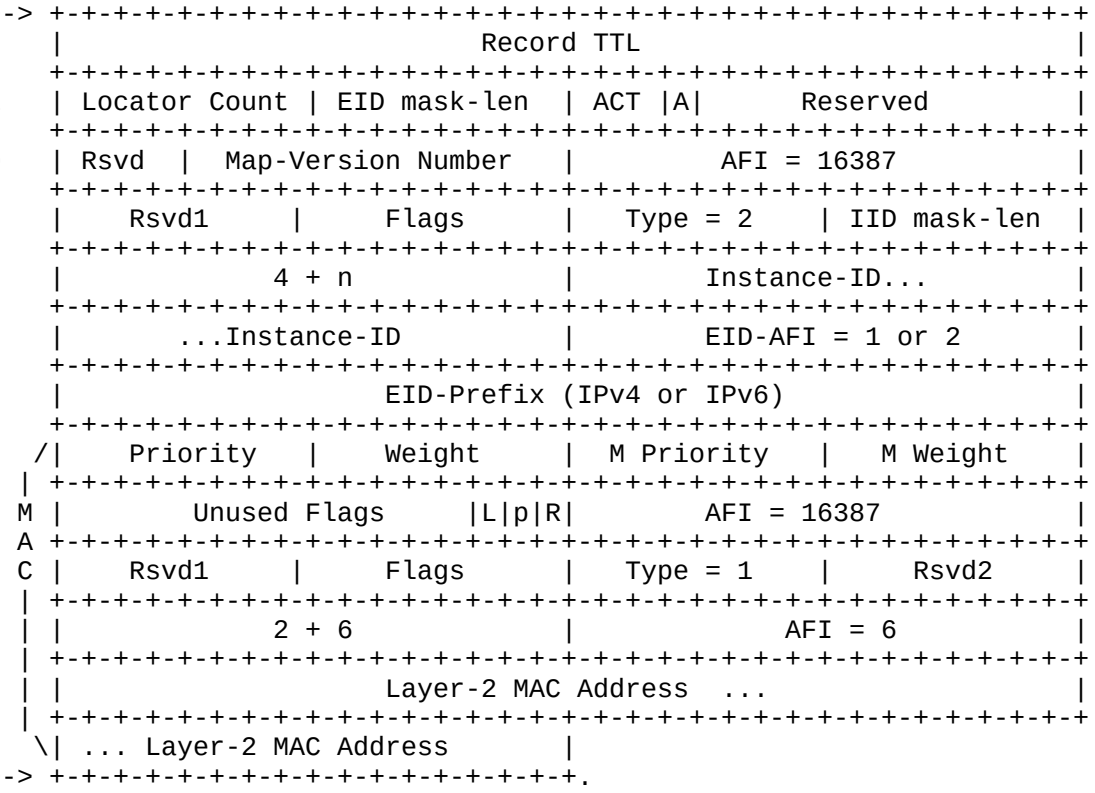
When an end-host is attached or detected in an ETR that provides L2-overlay services and also supports ARP resolution using the LISP control-plane, an additional mapping entry is registered to the mapping system:

- 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 [I-D.ietf-lisp-lcaf]. This option would also allow adding additional attributes to the locator record, while maintaining compatibility with legacy devices.

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



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

Locator Count: This value SHOULD be set to 1.

Instance-ID: This is the IID used to provide segmentation of the L2-overlays, L3 overlays and ARP tables.

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

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

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

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

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

5.3.3. Implementation Considerations

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

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

- LISP encap: This is the encapsulation format defined in the original LISP specification [RFC6830]. The encapsulation allows encapsulating both L2 and L3 packets. The Instance-ID used in the EIDs directly maps to the Instance-ID that the LISP header carries. At the ETR, after decapsulation, the IID MAY be used to decide between L2 processing or L3 processing.

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

7. IANA Considerations

This memo includes no request to IANA.

8. Acknowledgements

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

9. References

9.1. Normative References

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

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

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

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

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

9.2. Informative References

[I-D.coras-lisp-re]
Coras, F., Cabellos-Aparicio, A., Domingo-Pascual, J., Maino, F., and D. Farinacci, "LISP Replication Engineering", draft-coras-lisp-re-08 (work in progress), November 2015.

[I-D.ietf-lisp-ddt]
Fuller, V., Lewis, D., Ermagan, V., Jain, A., and A. Smirnov, "LISP Delegated Database Tree", draft-ietf-lisp-ddt-08 (work in progress), September 2016.

[I-D.ietf-lisp-lcaf]
Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", draft-ietf-lisp-lcaf-16 (work in progress), October 2016.

[I-D.ietf-lisp-signal-free-multicast]
Moreno, V. and D. Farinacci, "Signal-Free LISP Multicast", draft-ietf-lisp-signal-free-multicast-01 (work in progress), April 2016.

[I-D.ietf-nvo3-vxlan-gpe]
Kreeger, L. and U. Elzur, "Generic Protocol Extension for VXLAN", draft-ietf-nvo3-vxlan-gpe-02 (work in progress), April 2016.

[I-D.kouvelas-lisp-map-server-reliable-transport]
Cassar, C., Kouvelas, I., Lewis, D., Arango, J., and J. Leong, "LISP Map Server Reliable Transport", draft-kouvelas-lisp-map-server-reliable-transport-02 (work in progress), August 2016.

Authors' Addresses

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

Email: mportole@cisco.com

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

Email: vrushali@cisco.com

Victor Moreno
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
USA

Email: vimoreno@cisco.com

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

Email: fmaino@cisco.com

Dino Farinacci
lispers.net
San Jose, CA
USA

Email: farinacci@gmail.com