Interworking LISP with IPv4 and IPv6
draft-lewis-lisp-interworking-00

Status of this Memo

By submitting this Internet-Draft, each author represents that any applicable patent or other IPR claims of which he or she is aware have been or will be disclosed, and any of which he or she becomes aware will be disclosed, in accordance with Section 6 of BCP 79.

Internet-Drafts are working documents of the Internet Engineering Task Force (IETF), its areas, and its working groups. Note that other groups may also distribute working documents as Internet-Drafts.

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."

The list of current Internet-Drafts can be accessed at http://www.ietf.org/ietf/1id-abstracts.txt.

The list of Internet-Draft Shadow Directories can be accessed at http://www.ietf.org/shadow.html.

This Internet-Draft will expire on June 7, 2008.

Copyright Notice

Copyright (C) The IETF Trust (2007).

Abstract

This document describes various methods for interworking between the Internet and an Internet in which Endpoint Identifiers are separated from to Routing Locators using the Locator/ID Separation Protocol (LISP). Existing Internet hosts that do not participate in the LISP system must still be able to reach sites numbered from this EID prefixes. This new draft describes mechanisms which might be used to provide reachability between these types of sites. A new network
element, the Proxy Tunnel Router may also be deployed to act as a transitional LISP Ingress Tunnel Router for a subnetwork which does not implement LISP but which requires communication to devices that are addressed from LISP prefixes.

Table of Contents

1. Introduction ........................................... 3
2. LISP Interworking Models .............................. 3
3. Definition of Terms .................................... 4
4. Routable EIDs ........................................... 6
   4.1. Impact on Routing Table ............................. 6
   4.2. Requirement for using BGP ........................... 6
   4.3. Limiting the Impact of Routable EIDs ................. 6
   4.4. Use of Routable EIDs for Testing LISP ............... 6
5. Proxy Tunnel Routers ..................................... 7
   5.1. PTR EID announcements ............................... 7
   5.2. Packet Flow with PTRs .............................. 7
   5.3. Scaling PTRs ....................................... 8
   5.4. Impact of the PTRs placement in the network ........... 9
   5.5. Benefit to Networks Deploying PTRs ................... 9
6. LISP-NAT ............................................... 9
   6.1. LISP-NAT for LISP-NR addressed hosts ................. 10
   6.2. LISP Sites with Hosts using RFC 1918 Addresses Sending to non-LISP Sites ....................... 11
   6.3. LISP Sites with Hosts using RFC 1918 Addresses Communicating to Other LISP Sites .................... 11
   6.4. LISP-NAT and multiple EIDs .......................... 12
7. Security Considerations ................................ 12
8. Acknowledgments ...................................... 12
9. IANA Considerations .................................. 12
10. References ........................................... 13
  10.1. Normative References ............................... 13
  10.2. Informative References .............................. 13
Authors’ Addresses ...................................... 13
Intellectual Property and Copyright Statements .......... 15
1. Introduction

This document describes methods for transitioning from today’s Internet to a future Internet in which Endpoint Identifiers (EIDs) are separated from to Routing Locators (RLOCs) using the Locator/ID Separation Protocol (LISP) [LISP].

A key behavior of this separation that EID prefixes are not advertised to the Internet’s Default Free Zone (DFZ). In particular, only RLOCs are carried in the DFZ. Existing Internet hosts who do not participate in the LISP system must still be able to reach sites numbered from this non routed EID space. This draft describes a set of mechanisms which can be used to provide reachability between sites that are LISP-capable and those which are not. A new LISP network element, the Proxy Tunnel Router (PTR) (Section 5) may also be deployed to act as a LISP Ingress Tunnel Router (ITR) for a site which does not implement LISP but which requires communication to devices that are using LISP prefixes advertised into the DFZ.

Note that any successful interworking model should be agnostic to any particular EID-to-RLOC mapping algorithm. This document does not comment on the value of any of the particular mapping system.

The remainder of this document is organized as follows: Section 2 describes the internetworking models considered in this document. Section 3 defines terms used throughout this draft. Section 4 describes the relationship of the EID space and the current Internet. Finally, Section 5 and Section 6 introduce mechanisms which provide successful interworking between the two systems.

2. LISP Interworking Models

There are 4 unicast connectivity cases which describe how sites can communicate with each other:

1. Non-LISP site to Non-LISP site
2. LISP site to LISP site
3. LISP site to Non-LISP site
4. Non-LISP site to LISP site

The first case is the Internet as we know it today and as such will not be discussed further here. The second case is documented in [LISP] and hence there are no new interworking requirements (since there are no new protocol requirements placed on intermediate non-
LISP routers).

Cases 3 and 4, while seemingly similar, are subtly different and are explained below.

In case 3 (LISP site to Non-LISP site), a LISP site can send packets to a non-LISP site because the non-LISP site prefixes are routable. The non-LISP site need not do anything new to receive packets. The only action the LISP site needs to take is to know when not to LISP-encapsulate packets. This can be achieved via two mechanisms:

1. First, at the ITR in the source site, if the destination of an IP packet is found to match a prefix from the BGP routing table, then the site is directly reachable by the BGP core that exists and operates today.

2. Second, if (from the perspective of the ITR at the source site) the destination address of an IP address is not found in the EID-to-RLOC mapping database, the ITR could infer that it is not a LISP-capable site, and decide not to LISP-encapsulate the packet.

The most challenging case, case 4, occurs when a host at a non-LISP source site sends a packet to a host in a LISP site. If the source host chooses a non-routable EID address, the packet continue to be forwarded (inside the domain) until it reaches a router which cannot forward it, at which point the packet is dropped. So either the EID that the destination site is making available for other sites is routed outside of LISP or an alternate mechanisms need to be in place to solve this.

This specification describes two alternate interworking mechanisms, Proxy Tunnel Router (PTR) (Section 5) and LISP-NAT (Section 6).

3. Definition of Terms

Endpoint ID (EID): A 32- or 128-bit value used in the source and destination fields of the first (most inner) LISP header of a packet. A packet that is emitted by a system contains EIDs in its headers and LISP headers are prepended only when the packet reaches an Ingress Tunnel Router (ITR) on the data path to the destination EID.

EID-Prefix Aggregate: A set of EID-prefixes said to be aggregatable in the [RFC4632] sense. That is, an EID-Prefix aggregate is defined to be a single contiguous power-of-two EID-prefix block. Such a block is characterized by a prefix and a length.
Routing Locator (RLOC): An IP address of a LISP tunnel router. It is the output of a EID-to-RLOC mapping lookup. An EID maps to one or more RLOCs. Typically, RLOCs are numbered from topologically-aggregatable blocks and 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 Provider Aggregatable (PA) addresses.

EID-to-RLOC Mapping: A binding between an EID and the RLOC-set that can be used to reach the EID. We use the term "mapping" in this document to refer to a EID-to-RLOC mapping.

EID Prefix Reachability: An EID prefix is said to be "reachable" if one or more of its locators are reachable. That is, an EID prefix is reachable if the ETR (or its proxy) is reachable.

Default Mapping: A Default Mapping is a mapping entry for EID-prefix 0.0.0.0/0. It maps to a locator-set used for all EIDs in the Internet. If there is a more specific EID-prefix in the mapping cache it overrides the Default Mapping entry. The Default Mapping route can be learned by configuration or from a Map-Reply message [LISP].

LISP Routable (LISP-R) Site: A LISP site who’s addresses are used as both globally routable IP addresses and LISP EIDs.

LISP Non-Routable (LISP-NR) Site: A LISP site who’s addresses are EIDs only, these EIDs are not found in the legacy Internet routing table.

LISP Proxy Tunnel Router (PTR): PTRs are used to provide interconnectivity between sites which use LISP EIDs and those which do not. They act a gateway between the Legacy Internet and the LISP enabled Network. A given PTR advertises one or more highly aggregated EID prefixes into the public Internet and acts as the ITR for traffic received from the public Internet. LISP Proxy Tunnel Routers are described in Section 5.

LISP Network Address Translation (LISP-NAT): Network Address Translation between EID space assigned to a site and RLOC space also assigned to that site. LISP Network Address Translation is described in Section 6.

EID Sub Namespace: A power-of-two block of aggregatable locators set aside for LISP interworking.
4. Routable EIDs

An obvious way to achieve interworking between LISP and non-LISP hosts is to simply announce EID prefixes into the DFZ (much like today’s Provider Independent (PI) addresses are). Of course, making EIDs routable in this way is not desirable as it has the potential to increase the size of the DFZ table rather than reduce its size (a primary goal of LISP).

4.1. Impact on Routing Table

If EID prefixes are announced into the DFZ, the impact is similar to the case in which LISP has not been deployed, since these EID prefixes will not be any more aggregatable than existing PI addressing. This behavior is not desirable, and such a mechanism is not viewed as a viable long term solution.

4.2. Requirement for using BGP

Non-LISP sites today use BGP to (among other things) enable ingress traffic engineering. Relaxing this requirement is another primary design goal of LISP.

4.3. Limiting the Impact of Routable EIDs

Two schemes are proposed to limit the impact of having EIDs announced in the current global Internet routing table:

Section 5 discusses the LISP Proxy Tunnel Router, an approach that provides ITR functionality to bridge LISP-capable and non-LISP-capable sites.

Section 6 discusses another approach, LISP-NAT, in which NAT [RFC2993] is combined with ITR functionality to limit the impact of routable EIDs on the Internet routing infrastructure.

4.4. Use of Routable EIDs for Testing LISP

Given one of the primary design goals for LISP (and other Locator/ID separation schemes) is to be able to take advantage of topological aggregation of addresses. Note that since a primary design goal of LISP is to see benefits of aggregation as soon as possible, it is highly desirable to remove EID-prefixes from the global routing system.

That being said, sites that are using PA address blocks that are aggregated by their service provider, can use these addresses as EIDs without disadvantage to the the routing system (i.e., no additional
prefixes need to be advertised into the DFZ). In this case, interworking from a LISP site with PA addresses to a non-LISP site can easily be achieved, since such addresses are already routable. However, as mentioned above, it is a design goal of LISP to reduce the size of the DFZ routing table, and as such if follows that removal of EID prefixes (of any kind) from the DFZ follows is also a goal.

5. Proxy Tunnel Routers

Proxy Tunnel Routers (PTRs) allow for non-LISP sites to communicate with LISP-NR sites. PTRs have two primary functions:

- Originating EID Advertisements: PTRs advertise highly aggregated EID-prefix space on behalf. LISP sites to non-LISP sites can reach such LISP sites.

- Encapsulating Legacy Internet Traffic: PTRs also encapsulate non-LISP Internet traffic into LISP packets and route them towards their destination RLOCs.

5.1. PTR EID announcements

A critical characteristic of PTR functionality is to advertise aggressively aggregated EID-prefixes. Not only can the number of advertised routes be minimized, but in addition the number of places where the EID-prefix routes are stored can also be reduced by careful PTR placement. Rather than deploying PTRs close the LISP sites, they get deployed close to the non-LISP sites. In this way, load can be spread among many PTRs while at the same time as scoping the EID-prefix advertisements.

5.2. Packet Flow with PTRs

Packets from non-LISP sites can reach LISP-NR sites by the aid of PTRs. Packets from non-LISP sites traverse the Internet until reaching a PTR, where they are encapsulated. Once encapsulated, the packets use the LISP (outer) header’s destination address in order to reach the destination’s ETR.

Following is an example of the path a packet would take when using a PTR. In this example, the LISP-NR site is given the EID prefix of 140.0.0.0/24. Let us also say, for the purposes of this example, that this prefix, or a larger aggregate is not found in today’s Internet routing system. In other words, if a packet with this destination reached a router in the Default Free Zone, it would be dropped. In this example PTR is configured to announce this route...
(very likely a much larger aggregate), and so sink the traffic to the
PTR.

1. Host makes a DNS lookup EID for destination, and gets 140.1.1.1
   in return.

2. Host has a default route to customer Edge (CE) router and
   forwards the packet to the CE

3. The Site’s CE has a default route to its Provider Edge (PE)
   router, and forwards the packet to the PE.

4. The PE has route to 140.0.0.0/24 and the next hop is the PTR.

5. The PTR has a mapping for 140.0.0.0/24 and LISP encapsulates, the
   packet now has a destination address of the RLOC.

6. PTR looks up the RLOC, and forwards LISP packet to the next hop.

7. The ETR decapsulates the packet and delivers the packet to the
   140.1.1.1 host in the destination LISP site.

8. Packets from host 140.1.1.1 will flow back through the LISP
   site’s ITR. In this case, the packet is not encapsulated because
   the ITR knows the destination is a non-LISP site so the packet is
   delivered natively and directly to the destination site (i.e. the
   PTR does not get return traffic)

Note that in this example the return path is asymmetrical, so return
traffic will not go back through the PTR. This is because the
LISP-NR site’s ITR will discover that the originating site is not a
LISP site, and not encapsulate the returning packet (see [LISP] for
details of ITR behavior).

Alternatively, return traffic could be symmetrical. There are two
cases in which traffic could be symmetrical:

1. The ETR in the LISP site gleans the PTR’s RLOCs for the EID, and
2. The PTR could put the non-LISP site’s prefix in the mapping
database and use itself and a partner PTR as the RLOCs.

5.3. Scaling PTRs

PTRs sink traffic by announcing the LISP EID namespace. There are
several ways that an network could control the way traffic reaches a
PTR in order to prevent a given PTR from receiving more traffic than
it has the capacity to handle.
First, the PTR’s aggregate routes might be selectively announced in order to have a course way to control the typical amount of traffic sent towards a particular PTR.

Second, the same address might be announced by multiple PTRs in order to share the traffic by using the IP Anycast technique. The asymmetric nature of traffic flows allows for the PTR to be relatively simple - it will only have to encapsulate LISP packets.

5.4. Impact of the PTRs placement in the network

There are several that a network using PTRs would place the PTR as close as possible to non-LISP sites. First, placing the PTR near the ingress of traffic allows for the communication between the non-LISP site and the LISP site to have the last amount of stretch.

Some proposals (see for example CRIO [CRIO]) have suggested grouping PTRs near an arbitrary subset of ETRs and announcing a ‘local’ subset of EID space. This model cannot guarantee minimum stretch (the ratio of actual path length to optimal path length) if there are changes to where EIDs are injected into the system (for example, if a site changes ISPs).

5.5. Benefit to Networks Deploying PTRs

When traffic destined for LISP-NR site arrives and is encapsulated at a PTR, the new (appended) packet header is addressed to the destination’s RLOC. One benefit of this behavior is that this traffic will be routed towards RLOCs and be better able to follow the network’s traffic engineering policies (such as closest exit routing).

6. LISP-NAT

LISP Network Address Translation (LISP-NAT) is a limited form of NAT [RFC2993]. LISP-NAT is designed to enable the interworking of non-LISP sites and LISP-NR sites by ensuring that the LISP-NR’s sites addresses are always routable. LISP-NAT accomplishes this by translating a host’s source address from an ‘inner’ value to an ‘outer’ value and keeping this translation in a table that it can reference for subsequent packets.

In addition, existing RFC 1918 [RFC1918] sites can use LISP-NAT to talk to both LISP or non-LISP sites.

The basic concept of LISP-NAT is that when transmitting a packet, the
ITR replaces a non-routable EID source address with a routable source address, which enables packets to return to the site.

There are two main cases that involve LISP-NAT:

1. Hosts at LISP sites that use non-routable global EIDs speaking to non-LISP sites using global addresses.

2. Hosts at LISP sites that use RFC 1918 private EIDs speaking to other sites, who may be either LISP or non-LISP.

Note that LISP-NAT is not needed in the case of LISP-R (routable global EIDs) sources. This is because the LISP-R source’s address is routable, and return packets will be able to natively reach the site.

6.1. LISP-NAT for LISP-NR addressed hosts

LISP-NAT allows a host with a LISP-NR EID to communicate with non-LISP hosts by translating the LISP-NR EID to a globally unique address. This globally unique address may be a either a PI or PA address.

An example of this translation follows. For this example, a site has been assigned a LISP-NR EID of 220.1.1.0/24. In order to utilize LISP-NAT, the site has also been provided the PA EID of 128.200.1.0/24, and uses the first address (128.200.1.1) as the site’s RLOC. The rest of this PA space (128.200.1.2 to 128.200.1.254) is used as a translation pool for this site’s hosts who need to communicate with non-LISP hosts.

The translation table might look like the following:

<table>
<thead>
<tr>
<th>Site NR-EID</th>
<th>Site R-EID</th>
<th>Site’s RLOC</th>
<th>Translation Pool</th>
</tr>
</thead>
<tbody>
<tr>
<td>220.1.1.0/24</td>
<td>128.200.1.0/24</td>
<td>128.200.1.1</td>
<td>128.200.1.2 - 128.200.1.254</td>
</tr>
</tbody>
</table>

Figure 1: Example Translation Table

The Host 220.1.1.2 sends a packet destined for a non-LISP site to its default route (the ITR). The ITR receives the packet, and determines that the destination is not a LISP site. How the ITR makes this determination is up to the ITRs implementation of the EID-to-RLOC mapping system used (see, for example [LISP-ALT]).

The ITR then rewrites the source address of the packet from 220.1.1.2 to 128.200.1.2, which is the first available address in the LISP-R EID space available to it. The ITR keeps this translation in a table in order to reverse this process when receiving packets destined to
Finally, when the ITR forwards this packet without encapsulating it, it uses the entry in its LISP-NAT table to translate the returning packets’ destination IPs to the proper host.

6.2. LISP Sites with Hosts using RFC 1918 Addresses Sending to non-LISP Sites

In the case where RFC 1918 addressed hosts desire to communicate with non-LISP hosts the LISP-NAT implementation acts much like an existing IPv4 NAT device. The ITR providing the NAT service must use LISP-R EIDs for its global pool as well as providing all the standard NAT functions required today.

The source of the packet must be translated to a LISP-R EID in a manner similar to Section 6, and this packet must be forwarded to the ITR’s next hop for the destination, without LISP encapsulation.

6.3. LISP Sites with Hosts using RFC 1918 Addresses Communicating to Other LISP Sites

LISP-NAT allows a host with a RFC 1918 address to communicate with LISP hosts by translating the RFC 1918 address to a LISP EID. After translation, the communication between source and destination ITR and ETRs continues as described in [LISP].

An example of this translation and encapsulation follows. For this example, a host has been assigned a RFC 1918 address of 192.168.1.2. In order to utilize LISP-NAT, the site has also been provided the LISP-R EID of 128.200.1.0/24, and uses the first address (128.200.1.1) as the site’s RLOC. The rest of this PA space (128.200.1.2 to 128.200.1.254) is used as a translation pool for this site’s hosts who need to communicate with both non-LISP and LISP hosts.

The Host 192.168.1.2 sends a packet destined for a non-LISP site to its default route (the ITR). The ITR receives the packet, and determines that the destination is a LISP site. How the ITR makes this determination is up to the ITRs implementation of the EID/RLOC mapping system. (reference LISP-ALT for this behavior)

The ITR then rewrites the source address of the packet from 192.168.1.2 to 128.200.1.2, which is the first available address in the LISP EID space available to it. The ITR keeps this translation in a table in order to reverse this process when receiving packets destined to 128.200.1.2.
The ITR then LISP encapsulates this packet (see [LISP] for details). The ITR uses the site’s RLOC as the LISP outer header’s source, and the translation address as the LISP inner header’s source. Once it decapsulates returning traffic, it uses the entry in its LISP-NAT table to translate the returning packet’s destination IP address and then forward to the proper host.

6.4. LISP-NAT and multiple EIDs

When a site has two address that a host might use for global reachability, care must be chosen on which EID is found in DNS. For example, whether applications such as DNS use the LISP-R EID or the LISP-NR EID. This problem exists for NAT in general, but the specific issue described above is unique to LISP. Using PTRs can mitigate this problem, since the LISP-NR EID can be reached in all cases.

7. Security Considerations

Like any LISP ITR, PTRs will have the ability to inspect traffic at the time that they encapsulate. More work needs to be done to see if this ability can be exploited by the control plane along the lines of Remote Triggered BGP Black Holes. XXX:Reference?

As with traditional NAT, LISP-NAT will hide the actual host ID behind the RLOCs used as the NAT pool.

8. Acknowledgments

The authors would like to acknowledge Vince Fuller for his contributions and review of this draft. In addition, thanks goes to Christian Vogt and Lixia Zhang who have made insightful comments with respect to interworking and transition mechanisms.

A special thanks goes to Scott Brim for his initial brainstorming of these ideas and also for his careful review.

9. IANA Considerations

This document creates no new requirements on IANA namespaces [RFC2434].

10. References
10.1. Normative References


[RFC4632] Fuller, V. and T. Li, "Classless Inter-domain Routing (CIDR): The Internet Address Assignment and Aggregation Plan", BCP 122, RFC 4632, August 2006.

10.2. Informative References

[CRI0] Zhang, X., Francis, P., Wang, J., and K. Yoshida, "CRI0: Scaling IP Routing with the Core Router-Integrated Overlay".


Authors’ Addresses

Darrel Lewis
Cisco Systems, Inc.

Email: darlewis@cisco.com

David Meyer
Cisco Systems, Inc.

Email: dmm@cisco.com
Dino Farinacci
Cisco Systems, Inc.

Email: dino@cisco.com