Network Working Group
Internet-Draft
Intended status: Experimental
Expires: July 31, 2009

D. Lewis
D. Meyer
D. Farinacci
V. Fuller
Cisco Systems, Inc.
January 27, 2009

Interworking LISP with IPv4 and IPv6
draft-lewis-lisp-interworking-02

Status of this Memo

Copyright Notice

Abstract

This document describes techniques for allowing sites running the Locator/ID Separation Protocol (LISP [LISP]) to interoperate with Internet sites not running LISP. A fundamental property of LISP-speaking sites is that they use Endpoint Identifiers (EIDs), rather than traditional IP addresses, in the source and destination fields of all traffic they emit or receive. While EIDs are syntactically identical to IP addresses, routes for them are not carried in the global routing system so an interoperability mechanism is needed for non-LISP-speaking sites to exchange traffic with LISP-speaking sites. This document introduces two such mechanisms: the first uses a new network element, the LISP Proxy Tunnel Router (PTR) (Section 5) to act as a intermediate LISP Ingress Tunnel Router (ITR) for non-LISP-speaking hosts while the second adds Network Address Translation (NAT) functionality to LISP Ingress and LISP Egress Tunnel Routers (xTRs) to substitute routable IP addresses for non-routable EIDs.

Table of Contents

1. Introduction

This document describes two mechanisms for interoperation between LISP [LISP] sites, which use non-globally-routed EIDs, and non-LISP sites: use of PTRs, which create highly-aggregated routes to EID prefixes for non-LISP sites to follow; and the use of NAT by LISP ETRs when communicating with non-LISP hosts.

A key behavior of the separation of Locators and End-Point-IDs is that EID prefixes are not advertised to the Internet's Default Free Zone (DFZ). Specifically, only RLOCs are carried in the Internet's DFZ. Existing Internet sites (and their 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 that can be used to provide reachability between sites that are LISP-capable and those that are not. This document introduces two such mechanisms: the first uses a new network element, the LISP Proxy Tunnel Router (PTR) (Section 5) to act as a intermediate LISP Ingress Tunnel Router (ITR) for non-LISP-speaking hosts while the second adds a form of Network Address Translation (NAT) functionality to Tunnel Routers (xTRs) to substitute routable IP addresses for non-routable EIDs.

More detailed descriptions of these mechanisms and the network elements involved may be found in the following sections:

- Section 2 describes the different cases where interworking mechanisms are needed

- Section 3 defines terms used throughout the document

- Section 4 describes the relationship between the new EID prefix space and the IP address space used by the current Internet

- Section 5 introduces and describes the operation of PTRs

- Section 6 defines how NAT is used by ETRs to translate non-routable EIDs into routable IP addresses.

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

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

Note that while Cases 3 and 4 seem similar, there are subtle differences due to the way communications are originated.

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 because there are no new protocol requirements placed on intermediate non-LISP routers.

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. 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 to not LISP-encapsulate the packet.

Case 4, the most challenging, occurs when a host at a non-LISP site wishes to send traffic to a host at a LISP site. If the source host uses a (non-globally-routable) EID as the destination IP address, the packet is forwarded inside the source site until it reaches a router which cannot forward it, at which point the traffic is dropped. For traffic not to be dropped, either some route must be exist for the destination EID outside of LISP-speaking part of the network or an alternate mechanism must be in place. Section 5 (PTRs) and Section 6 (LISP-NAT) describe two such mechanisms.

Note that case 4 includes packets returning to the LISP Site in case 3.

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 whose addresses are used as both globally routable IP addresses and LISP EIDs.

LISP Non-Routable (LISP-NR) Site: A LISP site whose 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 as 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 routing system, effectively treating them as "Provider Independent (PI)" prefixes. Doing this is undesirable as it defeats one of the primary goals of LISP - to reduce global routing system state.

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, because these EID prefixes will be no 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 the impact of routable EIDs on the Internet routing infrastructure.

4.4. Use of Routable EIDs for Testing LISP

A primary design goal for LISP (and other Locator/ID separation proposals) is to facilitate topological aggregation of addresses and, thus, decrease global routing system state. Another goal is to achieve the benefits of improved aggregation as soon as possible. Advertising routes for LISP EID prefixes into the global routing system is therefore not recommended.

That being said, sites that are already using provider-aggregated prefixes can use these prefixes as LISP EIDs without adding state to the routing system; in other words, such sites do not cause additional prefixes to be advertised. For such sites, connectivity to a non-LISP sites does not require interworking machinery because the "PA" EIDs are already routable.

5. Proxy Tunnel Routers

Proxy Tunnel Routers (PTRs) allow for non-LISP sites to communicate with LISP-NR sites. A PTR is a new network element that shares many characteristics with the LISP ITR. PTRs allow non-LISP sites to send packets to LISP-NR sites without any changes to protocols or equipment at the non-LISP site. PTRs have two primary functions:

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

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 key part of PTR functionality is to advertise routes for highly-aggregated EID prefixes into part of the global routing system. Aggressive aggregation is performed to minimize the number of new announced routes. In addition, careful placement of PTRs can greatly reduce the scope of these new routes. To this end, PTRs should be deployed close to non-LISP-speaking rather than close to LISP sites. Such deployment not only limits the scope of EID-prefix route advertisements, it also also allows traffic forwarding load to be spread among many PTRs.

5.2. Packet Flow with PTRs

Packets from a non-LISP site can reach a LISP-NR site with the aid of a PTR. By advertising a route for a particular EID prefix into the global routing system, traffic destined for that EID prefix is routed to the PTR, which then performs LISP encapsulation. Once encapsulated, traffic packets use the LISP (outer) header's destination address to reach the destination ETR.

What follows 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 240.0.0.0/24. For the purposes of this example, this prefix and no covering aggregate is present in the global routing system. In other words, if a packet with this destination were to reach a router in the "Default Free Zone", it would be dropped.

A full protocol exchange example follows:

1. Source host makes a DNS lookup EID for destination, and gets 240.1.1.1 in return.

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

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

4. The PE has route to 240.0.0.0/24 and the next hop is the PTR.

5. The PTR has or acquires a mapping for 240.1.1.1 and LISP encapsulates, the packet now has a destination address of the RLOC. The source address of this encapsulated packet is the PTR's RLOC.

6. The 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 240.1.1.1 host in the destination LISP site.

8. Packets from host 240.1.1.1 will flow back through the LISP site's ITR. Such packets are not encapsulated because the ITR knows that the destination (the original source) is a non-LISP site. The ITR knows this because it can check the LISP mapping database for the destination EID, and on a failure determine that the destination site is not LISP enabled.

9. Packets are then routed natively and directly to the destination (original source) site.

Note that in this example the return path is asymmetric, 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).

The asymmetric nature of traffic flows allows the PTR to be relatively simple - it will only have to encapsulate LISP packets.

5.3. Scaling PTRs

PTRs attract traffic by announcing the LISP EID namespace into parts of the non-LISP-speaking global routing system. There are several ways that a network could control how traffic reaches a particular PTR to prevent it from receiving more traffic than it can handle:

First, the PTR's aggregate routes might be selectively announced, giving a coarse way to control the quantity of traffic attracted by that PTR.

Second, the same address might be announced by multiple PTRs in order to share the traffic using IP Anycast. The asymmetric nature of traffic flows allows 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 approaches that a network could take in placing PTRs. Placing the PTR near the ingress of traffic allows for the communication between the non-LISP site and the LISP site to have the least "stretch" (i.e. the least number of forwarding hops when compared to an optimal path between the sites).

Some proposals, 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 if the EID prefix route advertisement points are changed (such a change might occur if a site adds, removes, or replaces one or more ISPs connections).

5.5. Benefit to Networks Deploying PTRs

When traffic destined for LISP-NR site arrives and is encapsulated at a PTR, a new LISP packet header is pre-pended. This causes the packet's destination to be set to the destination site RLOC. Because

traffic is thus routed towards RLOCs, it can potentially better follow the network's traffic engineering policies (such as closest exit routing). This also means that providers who are not default-free and do not deploy PTRs end up sending more traffic to expensive transit links rather than to RLOC addresses, to which they may have settlement-free peering. For large transit providers, deploying PTRs may attract more traffic, and therefore more revenue, from their customers.

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 site 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:

Site NR-EID	Site R-EID	Site's RLOC	Translation Pool
220.1.1.0/24	128.200.1.0/24	128.200.1.1	128.200.1.2 - 128.200.1.254

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 128.200.1.2.

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 also has been provided the LISP-R EID of 192.0.2.0/24, and uses the first address (192.0.2.1) as the site's RLOC. The rest of this PA space (192.0.2.2 to 192.0.2.254) 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.

The ITR then rewrites the source address of the packet from 192.168.1.2 to 192.0.2.2, 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 192.0.2.2.

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 addresses 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.

6.5. LISP-NAT and PTRs Together

With LISP-NAT, there are two EIDs possible for a given host, the LISP-R EID and the LISP-NR EID. When a site has two addresses that a host might use for global reachability, name-to-address directories may need to be modified.

This problem, global addressability, exists for NAT in general, but

the specific issue described above is unique to LOC/ID split schemes. Some schemes [ref: 6-1 proxy] have suggested running a separate DNS instance for legacy types of EIDs. This solves the problem but introduces complexity for the site. Alternatively, using PTRs can mitigate this problem, because the LISP-NR EID can hbe reached in all cases.

In summary, there are two options for interworking LISP with IPv4 and V6. In the NAT case the LISP site can use NAT and manage the transition on its own. In the PTR case, we add a new network element called a PTR that can relieve that burden on the site, with the downside of potentially adding stretch to sites trying to reach the LISP site.

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.

When LISP Sites reply to non-LISP sites and rely on PTRs to enable Interworking, packets will be sourced from addresses not recognized by their Internet Service Provider's Unicast Reverse Path Forwarding (uRPF) enabled on the Provider Edge Router. Several options are available to the service provider. For example they could enable a less strict version of uRPF, where they only look for the existence of the the EID prefix in the routing table. Another, more secure, option is to add a static route for the customer on the PE router, but not redistribute this route into the provider's routing table.

8. Acknowledgments

Thanks goes to Christian Vogt, Lixia Zhang and Robin Whittle 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 Considersations

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

10. References

10.1. Normative References

[LISP] Farinacci, D., Fuller, V., Oran, D., and D. Meyer, "Locator/ID Separation Protocol (LISP)", draft-farinacci-lisp-11 (work in progress), July 2008.

[LISP-ALT] Farinacci, D., Fuller, V., and D. Meyer, "LISP Alternative Topology (LISP-ALT)", draft-fuller-lisp-alt-03 (work in progress), April 2008.

[RFC1918] Rekhter, Y., Moskowitz, R., Karrenberg, D., Groot, G., and E. Lear, "Address Allocation for Private Internets", BCP 5, RFC 1918, February 1996.

[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

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

[RFC2434] Narten, T. and H. Alvestrand, "Guidelines for Writing an IANA Considerations Section in RFCs", BCP 26, RFC 2434, October 1998.

[RFC2993] Hain, T., "Architectural Implications of NAT", RFC 2993, November 2000.

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

Vince Fuller
Cisco Systems, Inc.

Email: vaf@cisco.com