

Network Working Group	L. Jakab
Internet-Draft	A. Cabellos-Aparicio
Intended status: Informational	F. Coras
Expires: August 25, 2011	J. Domingo-Pascual
	Technical University of Catalonia
	D. Lewis
	Cisco Systems
	February 21, 2011

LISP Network Element Deployment Considerations  
draft-jakab-lisp-deployment-02.txt

## [Abstract](#)

This document discusses the different scenarios in which the LISP protocol may be deployed. Changes or extensions to other protocols needed by some of the scenarios are also highlighted.

## [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 August 25, 2011.

## [Copyright Notice](#)

Copyright (c) 2011 IETF Trust and the persons identified as the document authors. All rights reserved.

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

## [Table of Contents](#)

- \*1. [Introduction](#)

- \*2. [Tunnel Routers](#)
- \*2.1. [Customer Edge](#)
- \*2.2. [Provider Edge](#)
- \*2.3. [Split ITR/ETR](#)
- \*2.4. [Inter-Service Provider Traffic Engineering](#)
- \*2.5. [Tunnel Routers Behind NAT](#)
- \*2.5.1. [ITR](#)
- \*2.5.2. [ETR](#)
- \*2.6. [Summary and Feature Matrix](#)
- \*3. [Map-Resolvers and Map-Servers](#)
- \*3.1. [Map-Servers](#)
- \*3.2. [Map-Resolvers](#)
- \*4. [Proxy Tunnel Routers](#)
- \*4.1. [P-ITR](#)
- \*4.1.1. [LISP+BGP](#)
- \*4.1.2. [Mapping Service Provider P-ITR Service](#)
- \*4.1.3. [Tier 1 P-ITR Service](#)
- \*4.1.4. [Migration Summary](#)
- \*4.1.5. [Content Provider Load Balancing](#)
- \*4.2. [P-ETR](#)
- \*5. [Security Considerations](#)
- \*6. [IANA Considerations](#)
- \*7. [Acknowledgements](#)
- \*8. [References](#)
- \*8.1. [Normative References](#)
- \*8.2. [Informative References](#)

## **1. Introduction**

The Locator/Identifier Separation Protocol (LISP) addresses the scaling issues of the global Internet routing system by separating the current addressing scheme into Endpoint IDentifiers (EIDs) and Routing LOCators (RLOCs). The main protocol specification [\[I-D.ietf-lisp\]](#) describes how the separation is achieved, which new network elements are introduced, and details the packet formats for the data and control planes.

While the boundary between the core and edge is not strictly defined, one widely accepted definition places it at the border routers of stub autonomous systems, which may carry a partial or complete default-free zone (DFZ) routing table. The initial design of LISP took this location as a baseline for protocol development. However, the applications of LISP go beyond of just decreasing the size of the DFZ routing table, and include improved multihoming and ingress traffic engineering (TE) support for edge networks, and even individual hosts. Throughout the draft we will use the term LISP site to refer to these networks/hosts behind a LISP Tunnel Router. We formally define it as:

**LISP site:** A single host or a set of network elements in an edge network under the administrative control of a single organization, delimited from other networks by LISP Tunnel Router(s).

Since LISP is a protocol which can be used for different purposes, it is important to identify possible deployment scenarios and the additional requirements they may impose on the protocol specification and other protocols. The main specification [\[I-D.ietf-lisp\]](#) mentions positioning of tunnel routers, but without an in-depth discussion. This document fills that gap, by exploring the most common cases. While the theoretical combinations of device placements are quite numerous, the more practical scenarios are given preference in the following.

Additionally, this document is intended as a guide for the operational community for LISP deployments in their networks. It is expected to evolve as LISP deployment progresses, and the described scenarios are better understood or new scenarios are discovered.

Each subsection considers an element type, discussing the impact of deployment scenarios on the protocol specification. For definition of terms, please refer to the appropriate documents (as cited in the respective sections).

Comments and discussions about this memo should be directed to the LISP working group mailing list: [lisp@ietf.org](mailto:lisp@ietf.org).

## **2. Tunnel Routers**

LISP is a map-and-encap protocol, with the main goal of improving global routing scalability. To achieve its goal, it introduces several new network elements, each performing specific functions necessary to

separate the edge from the core. The device that is the gateway between the edge and the core is called Tunnel Router (xTR), performing one or both of two separate functions:

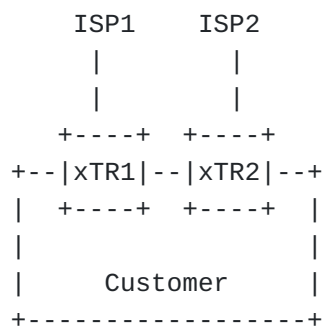
1. Encapsulating packets originating from an end host to be transported over intermediary (transit) networks towards the other end-point of the communication
2. Decapsulating packets entering from intermediary (transit) networks, originated at a remote end host.

The first function is performed by an Ingress Tunnel Router (ITR), the second by an Egress Tunnel Router (ETR).

Section 8 of the main LISP specification [\[I-D.ietf-lisp\]](#) has a short discussion of where Tunnel Routers can be deployed and some of the associated advantages and disadvantages. This section adds more detail to the scenarios presented there, and provides additional scenarios as well.

## 2.1. Customer Edge

LISP was designed with deployment at the core-edge boundary in mind, which can be approximated as the set of DFZ routers belonging to non-transit ASes. For the purposes of this document, we will consider this boundary to be consisting of the routers connecting LISP sites to their upstreams. As such, this is the most common expected scenario for xTRs, and this document considers it the reference location, comparing the other scenarios to this one.



From the LISP site perspective the main advantage of this type of deployment (compared to the one described in the next section) is having direct control over its ingress traffic engineering. This makes it is easy to set up and maintain active/active, active/backup, or more complex TE policies, without involving third parties.

Being under the same administrative control, reachability information of all ETRs is easier to synchronize, because the necessary control traffic can be allowed between the locators of the ETRs. A correct synchronous global view of the reachability status is thus available,

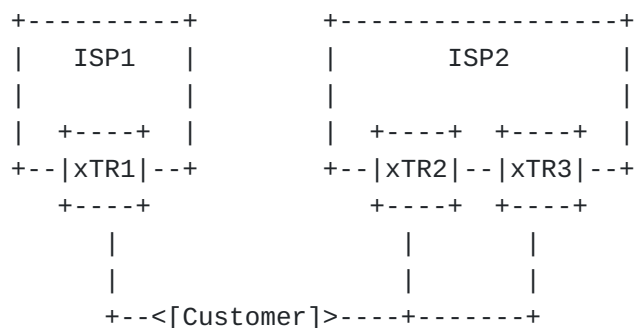
and the Loc-Status-Bits can be set correctly in the LISP data header of outgoing packets.

By placing the tunnel router at the edge of the site, existing internal network configuration does not need to be modified. Firewall rules, router configurations and address assignments inside the LISP site remain unchanged. This helps with incremental deployment and allows a quick upgrade path to LISP. For larger sites with many external connections, distributed in geographically diverse PoPs, and complex internal topology, it may however make more sense to both encapsulate and decapsulate as soon as possible, to benefit from the information in the IGP to choose the best path (see [Section 2.3](#) for a discussion of this scenario).

Another thing to consider when placing tunnel routers are MTU issues. Since encapsulating packets increases overhead, the MTU of the end-to-end path may decrease, when encapsulated packets need to travel over segments having close to minimum MTU. Some transit networks are known to provide larger MTU than the typical value of 1500 bytes of popular access technologies used at end hosts (e.g., IEEE 802.3 and 802.11). However, placing the LISP router connecting to such a network at the customer edge could possibly bring up MTU issues, depending on the link type to the provider as opposed to the following scenario.

## 2.2. Provider Edge

The other location at the core-edge boundary for deploying LISP routers is at the internet service provider edge. The main incentive for this case is that the customer does not have to upgrade the CE router(s), or change the configuration of any equipment. Encapsulation/decapsulation happens in the provider's network, which may be able to serve several customers with a single device. For large ISPs with many residential/business customers asking for LISP this can lead to important savings, since there is no need to upgrade the software (or hardware, if it's the case) at each client's location. Instead, they can upgrade the software (or hardware) on a few PE routers serving the customers. This scenario is depicted in [Figure 2](#).



While this approach can make transition easy for customers and may be cheaper for providers, the LISP site loses one of the main benefits of

LISP: ingress traffic engineering. Since the provider controls the ETRs, additional complexity would be needed to allow customers to modify their mapping entries.

The problem is aggravated when the LISP site is multihomed. Consider the scenario in [Figure 2](#): whenever a change to TE policies is required, the customer contacts both ISP1 and ISP2 to make the necessary changes on the routers (if they provide this possibility). It is however unlikely, that both ISPs will apply changes simultaneously, which may lead to inconsistent state for the mappings of the LISP site (e.g., weights for the same priority don't sum 100). Since the different upstream ISPs are usually competing business entities, the ETRs may even be configured to compete, either to attract all the traffic or to get no traffic. The former will happen if the customer pays per volume, the latter if the connectivity has a fixed price. A solution could be to have the mappings in the Map-Server(s), and have their operator give control over the entries to customer, much like in today's DNS.

Additionally, since xTR1, xTR2, and xTR3 are in different administrative domains, locator reachability information is unlikely to be exchanged among them, making it difficult to set Loc-Status-Bits correctly on encapsulated packets.

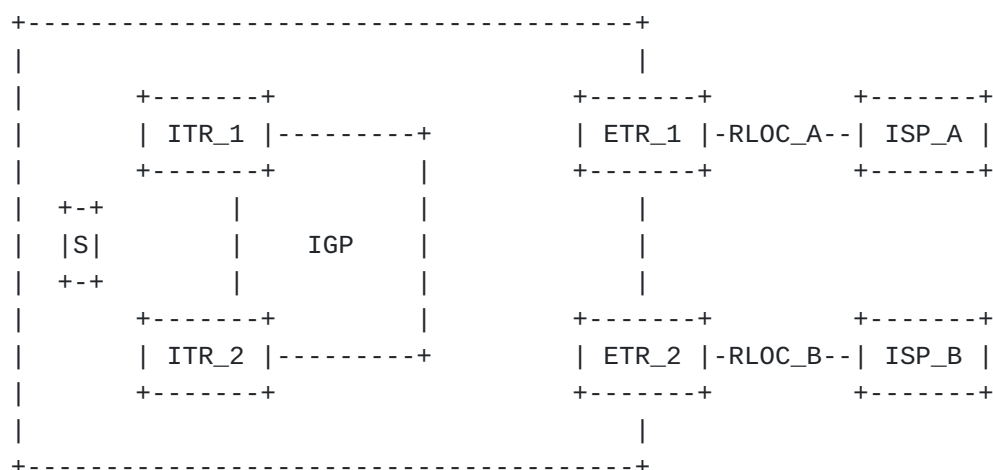
Compared to the customer edge scenario, deploying LISP at the provider edge might have the advantage of diminishing potential MTU issues, because the tunnel router is closer to the core, where links typically have higher MTUs than edge network links.

### [2.3. Split ITR/ETR](#)

In a simple LISP deployment, xTRs are located at the border of the LISP site (see [Section 2.1](#)). In this scenario packets are routed inside the domain according to the EID. However, more complex networks may want to route packets according to the destination RLOC. This would enable them to choose the best egress point.

The LISP specification separates the ITR and ETR functionality and considers that both entities can be deployed in separated network equipment. ITRs can be deployed closer to the host (i.e., access routers). This way packets are encapsulated as soon as possible, and packets exit the network through the best egress point in terms of BGP policy. In turn, ETRs can be deployed at the border routers of the network, and packets are decapsulated as soon as possible. Again, once decapsulated packets are routed according to the EID, and can follow the best path according to internal routing policy.

In the following figure we can see an example. The Source (S) transmits packets using its EID and in this particular case packets are encapsulated at ITR\_1. The encapsulated packets are routed inside the domain according to the destination RLOC, and can egress the network through the best point (i.e., closer to the RLOC's AS). On the other hand, inbound packets are received by ETR\_1 which decapsulates them. Then packets are routed towards S according to the EID, again following the best path.



This scenario has a set of implications:

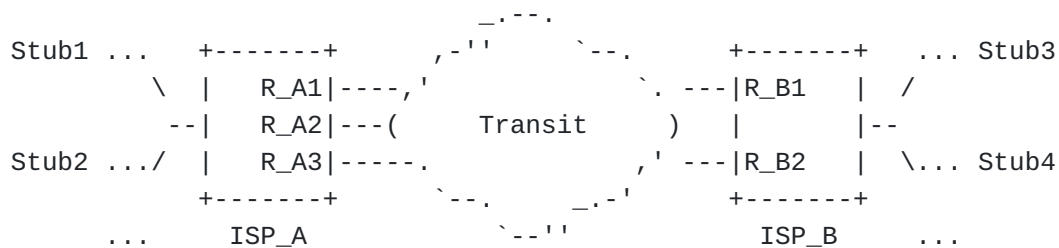
- \*The site must carry at least partial BGP routes in order to choose the best egress point, increasing the complexity of the network. However, this is usually already the case for LISP sites that would benefit from this scenario.
- \*If the site is multihomed to different ISPs and any of the upstream ISPs is doing uRPF filtering, this scenario may become impractical. ITRs need to determine the exit ETR, for setting the correct source RLOC in the encapsulation header. This adds complexity and reliability concerns.
- \*In LISP, ITRs set the reachability bits when encapsulating data packets. Hence, ITRs need a mechanism to be aware of the liveness of ETRs.
- \*ITRs encapsulate packets and in order to achieve efficient communications, the MTU of the site must be large enough to accommodate this extra header.
- \*In this scenario, each ITR is serving fewer hosts than in the case when it is deployed at the border of the network. It has been shown that cache hit ratio grows logarithmically with the amount of users [\[cache\]](#). Taking this into account, when ITRs are deployed closer to the host the effectiveness of the mapping cache may be lower (i.e., the miss ratio is higher). Another consequence of this is that the site will transmit a higher amount of Map-Requests, increasing the load on the distributed mapping database.

## **2.4. Inter-Service Provider Traffic Engineering**

With LISP, two LISP sites can route packets among them and control their ingress TE policies. Typically, LISP is seen as applicable to

stub networks, however the LISP protocol can also be applied to transit networks recursively.

Consider the scenario depicted in [Figure 4](#). Packets originating from the LISP site Stub1, client of ISP\_A, with destination Stub4, client of ISP\_B, are LISP encapsulated at their entry point into the ISP\_A's network. The external IP header now has as the source RLOC an IP from ISP\_A's address space and destination RLOC from ISP\_B's address space. One or more ASes separate ISP\_A from ISP\_B. With a single level of LISP encapsulation, Stub4 has control over its ingress traffic. However, ISP\_B only has the current tools (such as BGP prefix deaggregation) to control on which of his own upstream or peering links should packets enter. This is either not feasible (if fine-grained per-customer control is required, the very specific prefixes may not be propagated) or increases DFZ table size.



A solution for this is to apply LISP recursively. ISP\_A and ISP\_B may reach a bilateral agreement to deploy their own private mapping system. ISP\_A then encapsulates packets destined for the prefixes of ISP\_B, which are listed in the shared mapping system. Note that in this case the packet is double-encapsulated. ISP\_B's ETR removes the outer, second layer of LISP encapsulation from the incoming packet, and routes it towards the original RLOC, the ETR of Stub4, which does the final decapsulation.

If ISP\_A and ISP\_B agree to share a private distributed mapping database, both can control their ingress TE without the need of disaggregating prefixes. In this scenario the private database contains RLOC-to-RLOC bindings. The convergence time on the TE policies updates is fast, since ISPs only have to update/query a mapping to/from the database.

This deployment scenario includes two important recommendations. First, it is intended to be deployed only between two ISPs (ISP\_A and ISP\_B in [Figure 4](#)). If more than two ISPs use this approach, then the xTRs deployed at the participating ISPs must either query multiple mapping systems, or the ISPs must agree on a common shared mapping system. Second, the scenario is only recommended for ISPs providing connectivity to LISP sites, such that source RLOCs of packets to be reencapsulated belong to said ISP. Otherwise the participating ISPs must register prefixes they do not own in the above mentioned private mapping system. Failure to follow these recommendations may lead to operational and security issues when deploying this scenario.



Besides these recommendations, the main disadvantages of this deployment case are:

- \*Extra LISP header is needed. This increases the packet size and, for efficient communications, it requires that the MTU between both ISPs can accomodate double-encapsulated packets.
- \*The ISP ITR must encapsulate packets and therefore must know the RLOC-to-RLOC binding. These bindings are stored in a mapping database and may be cached in the ITR's mapping cache. Cache misses lead to an extra lookup latency, unless NERD [\[I-D.lear-lisp-nerd\]](#) is used for the lookups.
- \*The operational overhead of maintaining the shared mapping database.

## **[2.5.](#) Tunnel Routers Behind NAT**

NAT in this section refers to IPv4 network address and port translation.

### **[2.5.1.](#) ITR**

Packets encapsulated by an ITR are just UDP packets from a NAT device's point of view, and they are handled like any UDP packet, there are no additional requirements for LISP data packets.

Map-Requests sent by an ITR, which create the state in the NAT table have a different 5-tuple in the IP header than the Map-Reply generated by the authoritative ETR. Since the source address of this packet is different from the destination address of the request packet, no state will be matched in the NAT table and the packet will be dropped. To avoid this, the NAT device has to do the following:

- \*Send all UDP packets with source port 4342, regardless of the destination port, to the RLOC of the ITR. The most simple way to achieve this is configuring 1:1 NAT mode from the external RLOC of the NAT device to the ITR's RLOC (Called "DMZ" mode in consumer broadband routers).
- \*Rewrite the ITR-AFI and "Originating ITR RLOC Address" fields in the payload.

This setup supports a single ITR behind the NAT device.

### **[2.5.2.](#) ETR**

An ETR placed behind NAT is reachable from the outside by the Internet-facing locator of the NAT device. It needs to know this locator (and configure a loopback interface with it), so that it can use it in the

Map-Replies. Thus support for dynamic locators for the mapping database is needed in LISP equipment.

Again, only one ETR behind the NAT device is supported.

An implication of the issues described above is that LISP sites with xTRs can not be behind carrier based NATs, since two different sites would collide on the port forwarding.

## **2.6. Summary and Feature Matrix**

Feature	CE	PE	Split	Rec.
-----				
Control of ingress TE	x	-	x	x
No modifications to existing int. network infrastructure	x	x	-	-
Loc-Status-Bits sync	x	-	x	x
MTU/PMTUD issues minimized	-	x	-	x

## **3. Map-Resolvers and Map-Servers**

### **3.1. Map-Servers**

The Map-Server learns EID-to-RLLOC mapping entries from an authoritative source and publishes them in the distributed mapping database. These entries are learned through authenticated Map-Register messages sent by authoritative ETRs. Also, upon reception of a Map-Request, the Map-Server verifies that the destination EID matches an EID-prefix for which it is responsible for, and then re-encapsulates and forwards it to a matching ETR. Map-Server functionality is described in detail in [\[I-D.ietf-lisp-ms\]](#).

The Map-Server is provided by a Mapping Service Provider (MSP). A MSP can be any of the following:

- \*EID registrar. Since the IPv4 address space is nearing exhaustion, IPv4 EIDs will come from already allocated Provider Independent (PI) space. The registrars in this case remain the current five Regional Internet Registries (RIRs). In the case of IPv6, the possibility of reserving a /16 block as EID space is currently under consideration [\[I-D.meyer-lisp-eid-block\]](#). If granted by IANA, the community will have to determine the body responsible for allocations from this block, and the associated policies. For already allocated IPv6 prefixes the principles from IPv4 should be applied.

- \*Third parties. Participating in the LISP mapping system is similar to participating in global routing or DNS: as long as there is at least another already participating entity willing to forward the newcomer's traffic, there is no barrier to entry. Still, just like routing and DNS, LISP mappings have the issue of trust, with efforts underway to make the published information

verifiable. When these mechanisms will be deployed in the LISP mapping system, the burden of providing and verifying trust should be kept away from MSPs, which will simply host the secured mappings. This will keep the low barrier of entry to become an MSP for third parties.

In all cases, the MSP configures its Map-Server(s) to publish the prefixes of its clients in the distributed mapping database and start encapsulating and forwarding Map-Requests to the ETRs of the AS. These ETRs register their prefix(es) with the Map-Server(s) through periodic authenticated Map-Register messages. In this context, for some LISP end sites, there is a need for mechanisms to:

- \*Automatically distribute EID prefix(es) shared keys between the ETRs and the EID-registrar Map-Server.

- \*Dynamically obtain the address of the Map-Server in the ETR of the AS.

The Map-Server plays a key role in the reachability of the EID-prefixes it is serving. On the one hand it is publishing these prefixes into the distributed mapping database and on the other hand it is encapsulating and forwarding Map-Requests to the authoritative ETRs of these prefixes. ITRs encapsulating towards EIDs under the responsibility of a failed Map-Server will be unable to look up any of their covering prefixes. The only exception are the ITRs that already contain the mappings in their local cache. In this case ITRs can reach ETRs until the entry expires (typically 24 hours). For this reason, redundant Map-Server deployments are desirable. A set of Map-Servers providing high-availability service to the same set of prefixes is called a redundancy group. ETRs are configured to send Map-Register messages to all Map-Servers in the redundancy group. To achieve fail-over (or load-balancing, if desired), current known BGP practices can be used on the LISP+ALT BGP overlay network.

Additionally, if a Map-Server has no reachability for any ETR serving a given EID block, it should not originate that block into the mapping system.

### **3.2. Map-Resolvers**

A Map-Resolver is a network infrastructure component which accepts LISP encapsulated Map-Requests, typically from an ITR, and finds the appropriate EID-to-RLOC mapping by either consulting its local cache or by consulting the distributed mapping database. Map-Resolver functionality is described in detail in [\[I-D.ietf-lisp-ms\]](#).

Anyone with access to the distributed mapping database can set up a Map-Resolver and provide EID-to-RLOC mapping lookup service. In the case of the LISP+ALT mapping system, the Map-Resolver needs to become part of the ALT overlay so that it can forward packets to the

appropriate Map-Servers. For more detail on how the ALT overlay works, see [\[I-D.ietf-lisp-alt\]](#)

For performance reasons, it is recommended that LISP sites use Map-Resolvers that are topologically close to their ITRs. ISPs supporting LISP will provide this service to their customers, possibly restricting access to their user base. LISP sites not in this position can use open access Map-Resolvers, if available. However, regardless of the availability of open access resolvers, the MSP providing the Map-Server(s) for a LISP site should also make available Map-Resolver(s) for the use of that site.

In medium to large-size ASes, ITRs must be configured with the RLOC of a Map-Resolver, operation which can be done manually. However, in Small Office Home Office (SOHO) scenarios a mechanism for autoconfiguration should be provided.

One solution to avoid manual configuration in LISP sites of any size is the use of anycast RLOCs for Map-Resolvers similar to the DNS root server infrastructure. Since LISP uses UDP encapsulation, the use of anycast would not affect reliability. LISP routers are then shipped with a preconfigured list of well know Map-Resolver RLOCs, which can be edited by the network administrator, if needed.

The use of anycast also helps improving mapping lookup performance. Large MSPs can increase the number and geographical diversity of their Map-Resolver infrastructure, using a single anycasted RLOC. Once LISP deployment is advanced enough, very large content providers may also be interested running this kind of setup, to ensure minimal connection setup latency for those connecting to their network from LISP sites. While Map-Servers and Map-Resolvers implement different functionalities within the LISP mapping system, they can coexist on the same device. For example, MSPs offering both services, can deploy a single Map-Resolver/Map-Server in each PoP where they have a presence, as long as load on the two functions is comparable.

## **[4. Proxy Tunnel Routers](#)**

### **[4.1. P-ITR](#)**

Proxy Ingress Tunnel Routers (P-ITRs) are part of the non-LISP/LISP transition mechanism, allowing non-LISP sites to reach LISP sites. They announce via BGP certain EID prefixes (aggregated, whenever possible) to attract traffic from non-LISP sites towards EIDs in the covered range. They do the mapping system lookup, and encapsulate received packets towards the appropriate ETR. Note that for the reverse path LISP sites can reach non-LISP sites simply by not encapsulating traffic. See [\[I-D.ietf-lisp-interworking\]](#) for a detailed description of P-ITR functionality.

The success of new protocols depends greatly on their ability to maintain backwards compatibility and interoperate with the protocol(s) they intend to enhance or replace, and on the incentives to deploy the necessary new software or equipment. A LISP site needs an interworking

mechanism to be reachable from non-LISP sites. A P-ITR can fulfill this role, enabling early adopters to see the benefits of LISP, similar to tunnel brokers helping the transition from IPv4 to IPv6. A site benefits from new LISP functionality (proportionally with existing global LISP deployment) when going LISP, so it has the incentives to deploy the necessary tunnel routers. In order to be reachable from non-LISP sites it has two options: keep announcing its prefix(es) with BGP (see next subsection), or have a P-ITR announce prefix(es) covering them.

If the goal of reducing the DFZ routing table size is to be reached, the second option is preferred. Moreover, the second option allows LISP-based ingress traffic engineering from all sites. However, the placement of P-ITRs greatly influences performance and deployment incentives. The following subsections present the LISP+BGP transition strategy and then possible P-ITR deployment scenarios. They use the loosely defined terms of "early transition phase" and "late transition phase", which refer to time periods when LISP sites are a minority and a majority respectively.

#### **4.1.1. LISP+BGP**

For sites wishing to go LISP with their PI prefix the least disruptive way is to upgrade their border routers to support LISP, register the prefix into the LISP mapping system, but keep announcing it with BGP as well. This way LISP sites will reach them over LISP, while legacy sites will be unaffected by the change. The main disadvantage of this approach is that no decrease in the DFZ routing table size is achieved. Still, just increasing the number of LISP sites is an important gain, as an increasing LISP/non-LISP site ratio will slowly decrease the need for BGP-based traffic engineering that leads to prefix deaggregation. That, in turn, may lead to a decrease in the DFZ size in the late transition phase.

This scenario is not limited to sites that already have their prefixes announced with BGP. Newly allocated EID blocks could follow this strategy as well during the early LISP deployment phase, if the costs of setting up BGP routing are lower than using P-ITR services, or the expected performance is better. Since this leads to an increase in the DFZ size, one of the following scenarios should be preferred for new allocations.

#### **4.1.2. Mapping Service Provider P-ITR Service**

In addition to publishing their clients' registered prefixes in the mapping system, MSPs with enough transit capacity can offer them P-ITR service as a separate service. This service is especially useful for new PI allocations, to sites without existing BGP infrastructure, that wish to avoid BGP altogether. The MSP announces the prefix into the DFZ, and the client benefits from ingress traffic engineering without

prefix deaggregation. The downside of this scenario is path stretch, which is greater than 1.

Routing all non-LISP ingress traffic through a third party which is not one of its ISPs is only feasible for sites with modest amounts of traffic (like those using the IPv6 tunnel broker services today), especially in the first stage of the transition to LISP, with significant numbers of legacy sites. When the LISP/non-LISP site ratio becomes high enough, this approach can prove increasingly attractive.

Compared to LISP+BGP, this approach avoids DFZ bloat caused by prefix deaggregation for traffic engineering purposes, resulting in slower routing table increase in the case of new allocations and potential decrease for existing ones. Moreover, MSPs serving different clients with adjacent aggregable prefixes may lead to additional decrease, but quantifying this decrease is subject to future research study.

#### **4.1.3. Tier 1 P-ITR Service**

The ideal location for a P-ITR is on the traffic path, as close to non-LISP site as possible, to minimize or completely eliminate path stretch. However, this location is far away from the networks that most benefit from the P-ITR services (i.e., LISP sites, destinations of encapsulated traffic) and have the most incentives to deploy them. But the biggest challenge having P-ITRs close to the traffic source is the large number of devices and their wide geographical diversity required to have a good coverage, in addition to considerable transit capacity. Tier 1 service providers fulfill these requirements and have clear incentives to deploy P-ITRs: to attract more traffic from their customers. Since a large fraction is multihomed to different providers with more than one active link, they compete with the other providers for traffic.

To operate the P-ITR service, the ISP announces an aggregate of all known EID prefixes (a mechanism will be needed to obtain this list) downstream to their customers with BGP. First, the performance concerns of the MSP P-ITR service described in the previous section are now addressed, as P-ITRs are on-path, eliminating path stretch (except when combined with LISP+BGP, see below). Second, thanks to the direction of the announcements, the DFZ routing table size is not affected.

The main downside of this approach is non-global coverage for the announced prefixes, caused by the downstream direction of the announcements. As a result, a LISP site will be only reachable from customers of service providers running P-ITRs, unless one of the previous approaches is used as well. Due to this issue, it is unlikely that existing BGP speakers migrating to LISP will withdraw their announcements to the DFZ, resulting in a combination of this approach with LISP+BGP. At the same time, smaller new LISP sites still depend on MSP for global reachability. The early transition phase thus will keep the status quo in the DFZ routing table size, but offers the benefits of increasingly better ingress traffic engineering to early adopters. As the number of LISP destinations increases, traffic levels from those non-LISP, large multihomed clients who rely on BGP path length for

provider selection (such as national/regional ISPs), start to shift towards the Tier 1 providing P-ITRs. The competition is then incentivised to deploy their own service, thus improving global P-ITR coverage. If all Tier 1 providers have P-ITR service, the LISP+BGP and MSP alternatives are not required for global reachability of LISP sites. Still, LISP+BGP user may still want to keep announcing their prefixes for security reasons (i.e., preventing hijacking). DFZ size evolution in this phase depends on that choice, and the aggregability of all LISP prefixes. As a result, it may decrease or stay at the same level.

For performance reasons, and to simplify P-ITR management, it is desirable to minimize the number of non-aggregable EID prefixes. In IPv6 this can be easily achieved if a large prefix block is reserved as LISP EID space [\[I-D.meyer-lisp-eid-block\]](#). If the EID space is not fragmented, new LISP sites will not cause increase in the DFZ size, unless they do LISP+BGP.

#### **4.1.4. Migration Summary**

Phase	LISP+BGP	MSP	Tier-1
-----+-----+-----+-----			
Early transition	no change	slowdown increase	no change
Late transition	may decrease	slowdown increase	may decrease
LISP Internet		considerable decrease	

The following table presents the expected effects of the different transition scenarios during a certain phase on the DFZ routing table size:

It is expected that a combination of these scenarios will exist during the migration period, in particular existing sites choosing LISP+BGP, new small sites choosing MSP, and competition between Tier 1 providers bringing optimized service. If all Tier 1 ISPs have P-ITR service in place, the other scenarios can be deprecated, greatly reducing DFZ size.

#### **4.1.5. Content Provider Load Balancing**

By deploying P-ITRs in strategic locations, traffic engineering could be improved beyond what is currently offered by DNS, by adjusting percentages of traffic flow to certain data centers, depending on their load. This can be achieved by setting the appropriate priorities, weights and loc-status-bits in mappings. And since the P-ITRs are controlled by the content provider, changes can take place instantaneously.



## **4.2. P-ETR**

In contrast to P-ITRs, P-ETRs are not required for the correct functioning of all LISP sites. There are two cases, where they can be of great help:

- \*LISP sites with unicast reverse path forwarding (uRPF) restrictions, and

- \*LISP sites without native IPv6 communicating with LISP nodes with IPv6-only locators.

In the first case, uRPF filtering is applied at their upstream PE router. When forwarding traffic to non-LISP sites, an ITR does not encapsulate packets, leaving the original IP headers intact. As a result, packets will have EIDs in their source address. Since we are discussing the transition period, we can assume that a prefix covering the EIDs belonging to the LISP site is advertized to the global routing tables by a P-ITR, and the PE router has a route towards it. However, the next hop will not be on the interface towards the CE router, so non-encapsulated packets will fail uRPF checks.

To avoid this filtering, the affected ITR encapsulates packets towards the locator of the P-ETR for non-LISP destinations. Now the source address of the packets, as seen by the PE router is the ITR's locator, which will not fail the uRPF check. The P-ETR then decapsulates and forwards the packets.

The second use case is IPv4-to-IPv6 transition. Service providers using older access network hardware, which only supports IPv4 can still offer IPv6 to their clients, by providing a CPE device running LISP, and P-ETR(s) for accessing IPv6-only non-LISP sites and LISP sites, with IPv6-only locators. Packets originating from the client LISP site for these destinations would be encapsulated towards the P-ETR's IPv4 locator. The P-ETR is in a native IPv6 network, decapsulating and forwarding packets. For non-LISP destination, the packet travels natively from the P-ETR. For LISP destinations with IPv6-only locators, the packet will go through a P-ITR, in order to reach its destination. For more details on P-ETRs see the [\[I-D.ietf-lisp-interworking\]](#) draft. P-ETRs can be deployed by ISPs wishing to offer value-added services to their customers. As is the case with P-ITRs, P-ETRs too may introduce path stretch. Because of this the ISP needs to consider the tradeoff of using several devices, close to the customers, to minimize it, or few devices, farther away from the customers, minimizing cost instead. Since the deployment incentives for P-ITRs and P-ETRs are different, it is likely they will be deployed in separate devices, except for the CDN case, which may deploy both in a single device.

In all cases, the existence of a P-ETR involves another step in the configuration of a LISP router. CPE routers, which are typically configured by DHCP, stand to benefit most from P-ETRs. To enable



autoconfiguration of the P-ETR locator, a DHCP option would be required.

As a security measure, access to P-ETRs should be limited to legitimate users by enforcing ACLs.

## 5. Security Considerations

Security implications of LISP deployments are to be discussed in a separate document.

## 6. IANA Considerations

This memo includes no request to IANA.

## 7. Acknowledgements

Many thanks to Margaret Wasserman for her contribution to the IETF76 presentation that kickstarted this work. The authors would also like to thank Damien Saucez, Luigi Iannone, Joel Halpern, Vince Fuller, Dino Farinacci, Terry Manderson, Noel Chiappa, and everyone else who provided input.

## 8. References

### 8.1. Normative References

[I-D.ietf-lisp]	Farinacci, D, Fuller, V, Meyer, D and D Lewis, " <a href="#">Locator/ID Separation Protocol (LISP)</a> ", Internet-Draft draft-ietf-lisp-09, October 2010.
[I-D.ietf-lisp-interworking]	Lewis, D, Meyer, D, Farinacci, D and V Fuller, " <a href="#">Interworking LISP with IPv4 and IPv6</a> ", Internet-Draft draft-ietf-lisp-interworking-01, August 2010.
[I-D.ietf-lisp-ms]	Fuller, V and D Farinacci, " <a href="#">LISP Map Server</a> ", Internet-Draft draft-ietf-lisp-ms-06, October 2010.
[I-D.ietf-lisp-alt]	Fuller, V, Farinacci, D, Meyer, D and D Lewis, " <a href="#">LISP Alternative Topology (LISP+ALT)</a> ", Internet-Draft draft-ietf-lisp-alt-05, October 2010.

### 8.2. Informative References

[I-D.lear-lisp-nerd]	Lear, E, " <a href="#">NERD: A Not-so-novel EID to RLOC Database</a> ", Internet-Draft draft-lear-lisp-nerd-08, March 2010.
----------------------	---

<b>[I-D.meyer-lisp-eid-block]</b>	Lewis, D, Meyer, D and V Fuller, " <a href="#">LISP EID Block</a> ", Internet-Draft draft-meyer-lisp-eid-block-01, May 2008.
<b>[cache]</b>	Jung, J., Sit, E., Balakrishnan, H. and R. Morris, "DNS performance and the effectiveness of caching", 2002.

### Authors' Addresses

Lorand Jakab Jakab Technical University of Catalonia C/Jordi Girona, s/n BARCELONA, 08034 Spain EMail: [ljakab@ac.upc.edu](mailto:ljakab@ac.upc.edu)

Albert Cabellos-Aparicio Cabellos-Aparicio Technical University of Catalonia C/Jordi Girona, s/n BARCELONA, 08034 Spain EMail: [acabello@ac.upc.edu](mailto:acabello@ac.upc.edu)

Florin Coras Coras Technical University of Catalonia C/Jordi Girona, s/n BARCELONA, 08034 Spain EMail: [fcoras@ac.upc.edu](mailto:fcoras@ac.upc.edu)

Jordi Domingo-Pascual Domingo-Pascual Technical University of Catalonia C/Jordi Girona, s/n BARCELONA, 08034 Spain EMail: [jordi.domingo@ac.upc.edu](mailto:jordi.domingo@ac.upc.edu)

Darrel Lewis Lewis Cisco Systems 170 Tasman Drive San Jose, CA 95134 USA EMail: [darlewis@cisco.com](mailto:darlewis@cisco.com)