Locator/ID Separation Protocol Architecture

Abstract

The dramatic growth of the Internet has created scaling challenges that both adversely affect network capital expenditures and operating expenses, and impede the implementation of desirable features such as multi-homing, mobility, and IPv6 transition. The Locator/ID Separation Protocol (LISP) is a revolutionary new routing architecture that improves the scalability of the routing system while enabling enterprises and service providers to simplify multihoming, facilitate scalable any-to-any WAN connectivity, support data center virtual machine mobility, and reduce operation complexities. This document provides an overview of the LISP architecture and describes the functions of various components in the LISP solution.

Introduction and Problem Background

Since the public Internet first became part of the global infrastructure, its dramatic growth has created a number of scaling challenges. Among the most fundamental of these is helping ensure that the routing and addressing system continues to function efficiently even as the number of connected devices continues to increase.

A basic observation, made during early network research and development work and as documented in [1] and [2], is that use of a single address field for both device identification and routing is problematic; to effectively identify a device as a network session endpoint, an address should not change even if the device moves, such as from a home to a work location, or if the organization with which the device is associated changes its network connectivity, perhaps from one service provider to another. However, it is not feasible for the routing system to track billions of devices with such flexibly assigned addresses, so a device needs an address that is tightly coupled to its topological location to enable routing to operate efficiently.

To provide improved routing scalability while also facilitating flexible address assignment for multi-homing, provider independence, and mobility, the Locator Identifier Separation Protocol (LISP) was created. LISP describes a change to the Internet architecture in which IP addresses are replaced by routing locators (RLOCs) for routing through the global Internet and by endpoint identifiers (EIDs) for identifying network sessions between devices.

The key to using these RLOCs and EIDs is the mapping between them. A device (S1) communicating with another device (D1) will create a packet with the EID of S1 as the source IP address and the EID of D1 as the destination IP address. At the edge of the network, when this packet needs to be routed to the Internet, an ingress tunnel router (ITR; see “Ingress Tunnel Router” later in this document) maps the destination EID to a destination RLOC and then encapsulates the original packet with an additional header that has the source IP address of the ITR RLOC and the destination IP address of the RLOC of an egress tunnel router (ETR; see “Egress Tunnel Router” later in this document) that connects to D1. LISP specifies the way that the mappings between EIDs and RLOCs are defined, exchanged, and used.

Note that RLOCs and EIDs are syntactically identical to IP addresses, allowing LISP to be incrementally deployed and operated in a backward-compatible manner. This feature means that that although only LISP-capable devices can take advantage of the new features of the protocol, these devices can interoperate with non-LISP-capable devices.

Components of the LISP Architecture

Although the acronym "LISP" refers to the protocol used for exchanging EID-to-RLOC mapping information, the overall architecture includes two additional components: a mapping database, which consists of devices and procedures for distributing mapping information, and additional devices for implementing interworking between LISP-capable and non-LISP-capable parts of the Internet.

LISP Message Definitions

LISP, under development in the LISP working group of the IETF, is described in draft-ietf-lisp-09, a working document [3] that will initially be published as an experimental RFC. After a suitable period of prototyping and early operational deployment, the LISP developers expect it to proceed on the IETF standards track.

The LISP specification defines both an encapsulation format for transporting data between hosts and a set of protocol messages that are used by the LISP infrastructure for managing EID-to-RLOC mapping information. Both data and control messages use User Datagram Protocol (UDP) transport to facilitate passage through firewalls and distribution across parallel link aggregation group (LAG) paths through the Internet. Encapsulated user data packets are transported using UDP port 4341, and LISP control packets are transported using UDP port 4342.

User Data

A LISP-encapsulated user data packet consists of an IP header, with RLOCs in the source and destination IP address fields; followed by a UDP header, with a destination port of 4341 and a source port determined by a hash of the source and destination addresses and port fields from the original user data header (for LAG distribution); followed by the original user data packet. A LISP-encapsulated user data packet is created by an ITR when it receives a user data packet addressed to a LISP EID. The packet is sent to the ETR that owns that EID and decapsulated for delivery to the destination.

Control Messages

LISP control messages are used by LISP network elements to exchange EID-to-RLOC mapping information:

- Map-Request: This message is sent by an ITR to the mapping database (see the next section, "EID-to-RLOC Mapping Database") when it needs to send a packet to a destination EID for which is has no cached RLOC.
- Map-Reply: This message is returned to an ITR by an ETR or map server (see the next section, "EID-to-RLOC Mapping Database") in response to a Map-Request message. A Map-Reply message contains the EID prefix that matches the requested destination EID along with a set of RLOCs that can be used as the destination IP addresses for encapsulating user data. Additional information regarding priority and traffic-distribution across multiple RLOCs is also returned.
- Map-Register: This message is sent by an ETR to a map server to define an EID prefix that it owns as well as the RLOCs that should be used for exchanging Map-Request and Map-Reply messages. The registration request includes the EID prefix, prefix length, RLOCs associated with the prefix, and priorities and traffic-sharing weights of each RLOC. Map-Register messages are sent periodically to maintain the registration state between an ETR and its map servers.
- Encapsulated control message: This message is a Map-Request message that is encapsulated and sent from an ITR to a map resolver (see the next section, "EID-to-RLOC Mapping Database").

EID-to-RLOC Mapping Database

ETR Database Entries

The database that stores information for mapping from an EID to a set of RLOCs is called the EID-to-RLOC mapping database. Unlike in conventional databases, mapping entries are not centralized in a set of database servers. Instead, the authoritative source of the mapping information for a particular site is configured and stored in the ETR that connects it to the Internet. The full database, therefore, is distributed among all the ETRs at all LISP sites. Each ITR maintains a cache of the mapping database entries that it needs at a particular time; except in the case of a very unusual situation in which a single LISP site is concurrently communicating with all other LISP sites, no copy of the full mapping database exists in any individual system.

One mapping database entry exists for each defined EID. The entry is manually configured on each ETR that provides connectivity for the EID. The database entry consists of an EID prefix and prefix length and the set of ETR RLOCs that can be used to reach that EID prefix. Each RLOC entry includes a priority value (a lower value is preferred) and a weight value indicating the percentage of traffic that should use that RLOC. All ETRs that provide connectivity for an EID prefix must be configured with the same set of RLOCs, priority values, and weight values. Percentages should be configured to add up to a total of 100.

Connecting the Databases Together

To publish its mapping database entries, an ETR uses the Map-Register message to establish an association with one or more map servers. Map servers, in turn, are configured into an alternative logical topology (ALT) that is interconnected in a partial mesh using generic routing encapsulation (GRE) tunnels and Border Gateway Protocol (BGP). An individual map server is responsible for one or more bit-aligned prefixes; each of these prefixes is then subdivided into longer prefixes (smaller blocks of EIDs) that are assigned to individual LISP sites and configured into the ETRs for the sites. In some cases, aggregation may be performed at intermediate points in the EID prefix tree by dedicated ALT routers, which connect to the partial GRE and BGP mesh but have no ETR clients. Finally, map resolvers, which ITRs are configured to use to assist them in sending an Encapsulated Map-Request message to an ETR that can answer it, are also connected to the ALT network.

Database Lookup Procedures

For an ITR that needs to find one or more RLOCs for a given EID, the database lookup procedure is quite simple: it creates an Encapsulated Map-Request message and sends it to one of its configured map resolvers. The map resolver looks up the EID in its ALT routing table; if no match is found, then the destination is not an EID, and a negative Map-Reply message indicating that that LISP encapsulation should not be used for the destination is returned to the ITR. If a match does occur, then the Map-Request message is forwarded to the RLOC, which indicates the location of either an ETR that can answer it or a map server that is topologically closer to such an ETR.

More complete definitions of the LISP mapping database and its components can be found in draft-ietf-lisp-ms-06, which describes the LISP map server [4], and draft-ietf-lisp-alt-05, which describes LISP ALT [5].

EID-to-RLOC Mapping Database

A transition strategy is an important consideration when proposing any major change to or enhancement of the Internet architecture. Successful new features have tended to be those that both can be incrementally deployed and that offer benefits to early adopters prior to ubiquitous availability. LISP has been designed with this strategy in mind. To that end, two, complementary mechanisms are specified for interworking between LISP-capable and non-LISP-capable devices.

The first mechanism involves the deployment of proxy ingress tunnel routers (PITRs), which advertise highly aggregated blocks of EIDs to the global routing system to attract traffic destined to normally nonroutable EIDs. Upon receipt of traffic to an EID, a PITR performs the same EID-to-RLOC mapping and LISP encapsulation functions as any other ITR. Non-LISP sites are then able to send traffic to LISP sites simply by treating LISP EIDs as routable IP addresses. PITR deployment does have two potentially undesirable properties: first, it makes LISP sites dependent on third parties (those who deploy and operate PITRs), which requires some degree of multiorganizational coordination; and second, advertisement of EID prefixes to the global routing system does add to the overall size of the routing system. Furthermore, the presence of EID prefixes in the global routing system somewhat complicates the way that a LISP-capable site determines whether to use LISP-encapsulation to communicate with another LISP-capable site.

To avoid the coordination and complexity challenges associated with PITR deployment and use, a LISP-capable site can elect to deploy the second defined mechanism: LISP Network Address Translation (LISP-NAT). LISP-NAT is deployed by configuring the site’s ITRs to use LISP encapsulation for EIDs when communicating with other LISP-capable sites, but to perform NAT processing to replace EIDs with RLOCs when communicating with non-LISP-capable sites. LISP-NAT operates without the need for PITRs and has the advantage of being deployable without the need for coordination with or dependency on PITR operators. It does, though, share the same limitations as any other form of NAT.

See draft-ietf-lisp-interworking-01 [6] for a more complete description of the LISP interworking mechanisms.

Summary of LISP Network Element Functions

The LISP architecture defines seven new network infrastructure components. In some cases, a single physical device can implement more than one of these logical components.

Ingress Tunnel Router

An ITR is responsible for finding EID-to-RLOC mappings for all traffic destined for LISP-capable sites. When it receives a packet destined for an EID, it first looks for the EID in its mapping cache. If it finds a match, it encapsulates the packet inside a LISP header, with one of its RLOCs as the IP source address and one of the RLOCs from the mapping cache entry as the IP destination. It then routes the packet normally.

If no entry is found in its mapping cache, the ITR sends a Map-Request message to one of its configured map resolvers. It then discards the original packet. When it receives a response to its Map-Request message, it creates a new mapping cache entry with the contents of the Map-Reply message. When another packet, such as a retransmission for the original, discarded packet arrives, the mapping cache entry is used for encapsulation and forwarding. Note that the Map-Reply message may indicate that the destination is not an EID; if that occurs, a negative mapping cache entry is created, which causes packets to either be discarded or forwarded natively when the cache entry is matched.

The ITR function is usually implemented in the customer premises equipment (CPE) router and does not require hardware changes on software-switched platforms such as a Cisco® Integrated Services Router (ISR). The same CPE router will often provide both ITR and ETR functions; such a configuration is referred to as an xTR.

Egress Tunnel Router

An ETR connects a site to the LISP-capable part of the Internet, publishes EID-to-RLOC mappings for the site, responds to Map-Request messages, and decapsulates and delivers LISP-encapsulated user data to end systems at the site.

During operation, an ETR sends periodic Map-Register messages to all its configured map servers. The Map-Register messages contain all the EID-to-RLOC entries, which the ETR owns: that is, all the EID-numbered networks that are connected to the ETR's site.

When an ETR receives a Map-Request message, it verifies that the request matches an EID for which is responsible, constructs an appropriate Map-Reply message containing its configured mapping information, and sends this message to the ITR whose RLOCs are listed in the Map-Request message.

When an ETR receives a LISP-encapsulated packet that is directed to one of its RLOCs, it decapsulates the packet, verifies that the inner header is destined for an EID-numbered end system at its site, and then forwards the packet to the end system using site-internal routing.

Like the ITR function, the ETR function is usually implemented in a LISP site's CPE routers, typically as part of xTR function.

Map Server

A LISP map server (sometimes referred to as MS) implements the mapping database distribution. It does this by accepting registration requests from its client ETRs, aggregating the EID prefixes that they successfully register, and advertising the aggregated prefixes to the ALT with BGP. To do this, it is configured with a partial mesh of GRE tunnels and BGP sessions to other map server systems or ALT routers. Since a map server does not forward user data traffic, it does not have high-performance switching capability and is well-suited for implementation on a general-purpose computing server rather than on special-purpose router hardware. Both map server and map resolver functions are typically implemented on a common system; such a system is referred to as a map resolver/map server (MR/MS).

Map Resolver

Like a map server, a LISP map resolver (sometimes referred to as MR) connects to the ALT using a partial mesh of GRE tunnels and BGP sessions. It accepts Encapsulated Map-Request messages sent by ITRs, decapsulates them, and then forwards them over the ALT toward the ETRs responsible for the EIDs being requested.

ALT Router

An ALT router, which may not be present in all mapping database deployments, connects through GRE tunnels and BGP sessions, map servers, map resolvers, and other ALT routers. Its only purpose is to accept EID prefixes advertised by devices that form a hierarchically distinct part of the EID numbering space and then advertise an aggregated EID prefix that covers that space to other parts of the ALT. Just as in the global Internet routing system, such aggregation is performed to reduce the number of prefixes that need to be propagated throughout the entire network. A map server or combined MR/MS may also perform such aggregation, thus implementing the functions of an ALT router.

Proxy Ingress Tunnel Router

A PITR implements ITR mapping database lookups and LISP encapsulation functions on behalf of non-LISP-capable sites. PITRs are typically deployed near major Internet exchange points (IXPs) or in Internet service provider (ISP) networks to allow non-LISP customers of those facilities to connect to LISP sites. In addition to implementing ITR functions, a PITR also advertises some or all the non-routable EID prefix space to the part of the non-LISP-capable Internet that it serves. This advertising is performed so that the non-LISP sites will route traffic toward the PITR for encapsulation and forwarding to LISP sites. Note that these advertisements are intended to be highly-aggregated, with many EID prefixes covered by each prefix advertised by a PITR.

Proxy Egress Tunnel Router

A PETR implements ETR functions on behalf of non-LISP sites. A PETR is typically used when a LISP site needs to send traffic to non-LISP sites but cannot do so because its access network (the service provider to which it connects) will not accept nonroutable EIDs as packet sources. When dual-stacked, a PETR may also serve as a mechanism for LISP sites with EIDs within one address family and RLOCs within a different address family to communicate with each other. The PETR function is commonly offered by devices that also act as PITRs; such devices are referred to as PxTRs.

See draft-ietf-lisp-interworking-01 [6] for a more complete description of the PITR and PETR LISP interworking mechanisms.

Applicability and Use Cases

Although LISP was initially designed to offer simplified, scalable multihoming of Internet sites with improved traffic-engineering capabilities, the level of indirection it provides for IP addressing also enables a variety of other applications, including:

- Simplified and cost-effective multihoming, including ingress traffic engineering
- IP address portability, including no renumbering when changing providers or adding multihoming
- IP address (host) mobility, including session persistence across mobility events
- IPv6 transition simplification, including incremental deployment of IPv6 using existing IPv4 infrastructure (or IPv4 over IPv6)
- Simplified multi-tenancy and large-scale VPNs
- Operation and network simplification

More information about LISP capabilities, use cases, and deployments can be found at http://www.cisco.com/go/lisp, http://lisp4.cisco.com, and http://lisp6.cisco.com.

Conclusion

LISP provides an exciting new mechanism for enterprises and service providers to improve routing system scalability and introduce new capabilities to their networks. LISP is a consolidated, robust architecture that enables functions that can help businesses achieve new revenue streams, reduce capital expenditures and operating expenses, simplify network design, reduce dependency on a variety of tools, reduce system load, and improve network scalability. LISP does not require changes within sites or to the Internet core. LISP can be gradually introduced into an existing IP network without affecting the network endpoints or hosts. LISP is a Cisco innovation that is being promoted as an open standard. Through its participation in standards bodies such as the IETF LISP Working Group, Cisco is committed to the development of the LISP architecture. LISP capabilities are currently supported on a range of Cisco routing and switching platforms.

For more information about LISP please visit http://www.cisco.com/go/lisp. For general LISP solution questions, including deployment guidance, contact your local Cisco account representative or send an email to lisp-support@cisco.com. A public LISP network is available now. Visit http://lisp4.cisco.com or http://lisp6.cisco.com to use the LISP network and read more about supported features and functions.

References

1. J. Chiappa, “Endpoints and Endpoint names: A Proposed Enhancement to the Internet Architecture,” Internet-Draft, http://www.chiappa.net/~jnc/tech/endpoints.txt, 1999.
2. D. Meyer, L. Zhang, and K. Fall, “Report from the IAB Workshop on Routing and Addressing,” RFC 4984, September 2007.

3. D. Farinacci, V. Fuller, D. Meyer, and D. Lewis, “Locator/ID Separation Protocol (LISP),” draft-ietf-lisp-09 (work in progress), April 2010.
4. D. Farinacci and V. Fuller, “LISP Map Server,” draft-ietf-lisp-ms-06 (work in progress), October 2010.
5. D. Farinacci, V. Fuller, D. Meyer, and D. Lewis, “LISP Alternative Topology (LISP-ALT),” draft-ietf-lisp-alt-05 (work in progress), October 2010.
6. D. Lewis, D. Meyer, D. Farinacci, and V. Fuller, “Interworking LISP with IPv4 and IPv6,” draft-ietf-lisp-interworking-01 (work in progress), August 2010.

For More Information

LISP Documentation: LISP Command Reference Guide, LISP Configuration Guide, and LISP Lab Test Guide, at http://lisp4.cisco.com and http://lisp6.cisco.com.

Cisco marketing information about LISP: http://www.cisco.com/go/lisp

LISP Beta Network information: http://www.lisp4.net and http://www.lisp6.net

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **www.cisco.com/go/offices.**

Printed in USA C11-652502-00 03/11