

Internet Engineering Task Force
Internet-Draft
Intended status: Experimental
Expires: July 11, 2013

D. Farinacci
cisco Systems
M. Napierala
AT&T
January 7, 2013

LISP Control-Plane Multicast Signaling
draft-farinacci-lisp-mr-signaling-01

Abstract

This document describes an alternate method to signal multicast tree building among xTRs in multicast capable LISP sites. This approach avoids the need to run a multicast routing protocol on the LISP overlay.

Status of this Memo

This Internet-Draft is submitted in full conformance with the provisions of [BCP 78](#) and [BCP 79](#).

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

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

This Internet-Draft will expire on July 11, 2013.

Copyright Notice

Copyright (c) 2013 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.	Requirements Language	3
2.	Introduction	4
3.	Definition of Terms	6
4.	Overview	7
5.	Join-Request/Leave-Request Encoding Formats	9
6.	Replication Considerations	11
7.	Interworking Considerations	12
8.	Security Considerations	13
9.	IANA Considerations	14
10.	References	15
10.1.	Normative References	15
10.2.	Informative References	15
Appendix A.	Acknowledgments	16
Appendix B.	Document Change Log	17
B.1.	Changes to draft-farinacci-lisp-mr-signaling-01.txt	17
B.2.	Changes to draft-farinacci-lisp-mr-signaling-00.txt	17
Authors'	Addresses	18

1. Requirements Language

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

2. Introduction

The Locator/ID Separation Architecture [[LISP](#)] provides a mechanism to separate out Identification and Location semantics from the current definition of an IP address. By creating two namespaces, an Endpoint ID (EID) namespace used by sites and a Routing Locator (RLOC) namespace used by core routing, the core routing infrastructure can scale by doing topological aggregation of routing information.

Since LISP creates a new namespace, a mapping function must exist to map a site's EID prefixes to its associated locators. For unicast packets, both the source address and destination address must be mapped. For multicast packets, only the source address needs to be mapped. The destination group address doesn't need to be mapped because the semantics of an IPv4 or IPv6 group address are logical in nature and not topology-dependent. Therefore, this specification focuses on the procedures of how to map a source EID address of a multicast flow during distribution tree setup and packet delivery.

The LISP Multicast specification [[LISP-MCAST](#)] addresses the following procedures:

1. How a multicast source host in a LISP site sends multicast packets to receivers inside of its site as well as to receivers in other sites that are LISP enabled.
2. How inter-domain (or inter-LISP sites) multicast distribution trees are built and how forwarding of multicast packets leaving a source site toward receivers sites is performed.
3. What protocols are affected and what changes are required to such multicast protocols.
4. How ASM-mode (Any Source Multicast), SSM-mode (Single Source Multicast), and Bidir-mode (Bidirectional Shared Trees) service models will operate.
5. How multicast packet flow will occur for multiple combinations of LISP and non-LISP capable source and receiver sites.

The distribution tree mechanism in [[LISP-MCAST](#)] specifies the use of the PIM multicast routing protocol and how PIM is used between xTRs connecting multicast capable source sites and receiver sites together.

This specification will describe an alternate method for connecting multicast capable sites together by using Map-Requests instead of the PIM protocol. The advantages this brings is a single LISP control-

plane mechanism used for both unicast and multicast packet flow.

This specification does not describe how (S-EID,G) multicast distribution tree state is built in multicast receiver sites and in multicast source sites. This specification also does not describe how (S-RLOC,G) or (S-RLOC,DG) multicast distribution tree state is built in the core network. This specification defines how (S-EID,G) state is propagated from multicast receiver site resident ETRs to multicast ITRs. This signaling is needed so the (S-EID,G) state can be propagated from the ITR to the source host in the multicast source site.

3. Definition of Terms

Note that all the definitions that apply in [[LISP-MCAST](#)] apply in this specification as well. And the following definitions are specific to this specification.

Join-Request: This is a reference to a Map-Request that allows the joining to a multicast tree by an ETR to an ITR (or PITR) for a given (S-EID,G) entry.

Leave-Request: This is a reference to a Map-Request that allows the leaving from a multicast tree by an ETR to an ITR (or PITR) for a given (S-EID,G) entry.

LISP-RE: RE stands for "Replication Engineering" which is a term used to design algorithms, protocols, and networks to optimize where multicast packet replication is performed in the network.

Unicast Replication: Is the notion of replicating a multicast packet at an ITR (or PITR) by encapsulating it into a unicast packet. That is, the oif-list of a multicast routing table entry can not only have interfaces present for link-layer replication and for multicast encapsulation but also for unicast encapsulation.

Delivery Group: This is the outer packet's (or encapsulating header's) destination address when encapsulating a multicast packet inside of a multicast packet. There is a one-to-one and one-to-many relationship between the inner header's destination group address and the outer header's destination group address. Notation for a delivery group entry is (S-RLOC,DG).

(S-EID,G): This is multicast state notation which is signaled from ETR to ITR in Map-Request messages. 'G' is the group address multicast hosts send and/or join to. 'S-EID' can be a host EID that sends multicast packets. 'S-EID' can be a Rendezvous Point (RP) that resides in the LISP multicast site so (*,G) state can be signaled from ETR to ITR. All of PIM (S,G), (*,G), and (S,G,RP-bit) state can be conveyed via the Multicast Info Type format [[LISP-LCAF](#)] in Map-Request messages.

4. Overview

In [[LISP-MCAST](#)], there is a two step approach for an ETR to join (S-EID,G) to an ITR. In the first step, the ETR must look up which ITR is associated with S-EID. That is performed with a mapping database lookup and having the ETR select an ITR from the list of high priority RLOCs. In the second step, a unicast PIM join must be sent by the ETR to the ITR.

In the design here within, we transmit the join and leave semantic in a Map-Request message. In this case, both the S-EID lookup and the fact the ETR wants to join S-EID for a particular multicast group can be conveyed in one message exchange. The advantages of this are:

1. Less message overhead used for signaling.
2. State signaling comes together in a single message. If an ETR has a map-cache entry for the S-EID, it also knows that the Join for (S-EID,G) reliably made it to the ITR. If there is message loss both pieces of state fate-share the loss.
3. The Map-Reply is used as an acknowledgement where as with unicast PIM Join-Prune messages, they must be sent periodically which may create scalability problems in networks with a lot of multicast state.

Here is the basic procedure that a multicast ETR or multicast PETR uses to convey (S-EID,G) join state to a multicast ITR or multicast PITR:

1. When an ETR creates (S-EID,G) from a site based PIM Join message and the oif-list goes non-empty, a Join-Request is sent. If a map-cache entry exists for S-EID, then the Map-Request is sent to the highest multicast priority RLOC. If a map-cache entry does not exist, the Map-Request is sent to the mapping database system.
2. When a Map-Reply is not returned, the Map-Request is retransmitted. When a Map-Reply is returned, the ETR can be assured that the ITR will replicate packets to the ETR.
3. When unicast replication is performed, no additional action needs to be performed by the ETR.
4. When multicast replication is performed, the ETR must send a PIM Join message for (S-RLOC,G) or (S-RLOC,DG) into the core as specified in [[LISP-MCAST](#)]. See [Section 6](#) for details when ITR unicast and/or multicast replication is done and how it is

decided.

An ETR must detect when an ITR has reloaded or cleared its state so that the ETR can resend Join-Requests for all the (S-EID,G) state it has cached. Procedures for how to achieve this will be discussed in future versions of this specification.

Here is the basic procedure a multicast ETR or multicast PETR uses to convey (S-EID,G) leave state to a multicast ITR or multicast PITR:

1. When an ETR (S-EID,G) oif-list state goes empty, a Leave-Request is sent. If a map-cache entry exists for S-EID, then the Map-Request is sent to the highest multicast priority RLOC. If a map-cache entry does not exist, the Map-Request is sent to the mapping database system.
2. When a Map-Reply is not returned, the Map-Request is retransmitted. When a Map-Reply is returned, the ETR can be assured that the ITR will no longer replicate packets to the ETR.
3. When unicast replication is performed, no additional action needs to be performed by the ETR.
4. When multicast replication is performed, the ETR must send a PIM Leave message for (S-RLOC,G) or (S-RLOC,DG) into the core network as specified in [[LISP-MCAST](#)].

5. Join-Request/Leave-Request Encoding Formats

A Join-Request and Leave-Request are encoded as follows:

- o (S-EID,G) is encoded in the destination EID-prefix field of a Map-Request [[LISP](#)].
- o The encoding format of the destination EID-prefix is in LCAF format Type 'Multicast Info' [[LISP-LCAF](#)]. The J-bit and L-bit indicate if the Map-Request is a Join-Request or a Leave-Request, respectively.
- o (S-RLOC,DG) is encoded in the Source EID Address field of the Map-Request. It is also encoded in the same LCAF Type 'Multicast Info'.
- o If S-RLOC is not known, then AFI=0 is encoded in the Source Address field of the LCAF type.
- o If S-RLOC is known, then the RLOC of the ITR is encoded in the Source Address field of the LCAF type.
- o If a Delivery Group is being requested by the ETR, then DG is encoded in the Group Address field of the LCAF type.
- o If a unicast replication is being requested by the ETR, then ETR encodes a unicast RLOC address in the Group Address field of the LCAF type.

A Map-Reply is returned for a Join-Request or a Leave-Request with the following format encoding:

1. The destination EID-prefix encoding in the Map-Request is copied and encoded in the Source EID Address field of the Map-Reply. This is so the ETR can match Map-Replies with Map-Requests. The nonce field may be used for this purpose as well. The address encoding is (S-EID,G).
2. The destination EID-prefix in the Map-Reply is the multicast information the ITR is conveying to ETR. It can be (ITR-RLOC,DG) or (ITR-RLOC,ETR-RLOC).
3. When the ETR requested unicast replication, then the returned destination EID-prefix contains (ITR-RLOC,ETR-RLOC)
4. When the ETR requested a DG for multicast replication, then the returned destination EID-prefix contains (ITR-RLOC,DG).

5. When the ITR overrides a requested (ITR-RLLOC,ETR-RLLOC) with a returned (ITR-RLLOC,DG), then the ETR must send a join (or leave) for (ITR-RLLOC,DG) into the core network.
6. When the ITR overrides a join-requested (ITR-RLLOC,DG1) with a returned value of (ITR-RLLOC,DG2), then the ETR must send a Join-Request for (ITR-RLLOC,DG2) and send a Leave-Request for (ITR-RLLOC,DG1) into the core network.
7. When the ITR with RLLOC 'RLLOC-ITR1' returns (RLLOC-ITR2,DG) in a Map-Reply, the ETR must send a Join-Request to RLLOC-ITR2 and send a Leave-Request to RLLOC-ITR1 for (RLLOC-ITR1,DG). Same action is performed when (RLLOC-ITR2,ETR-RLLOC) is returned for a join-requested value of (RLLOC-ITR1,ETR-RLLOC).

6. Replication Considerations

When an ITR processes a received multicast packet sourced by a host in its site, the oif-list for the (S-EID,G) entry it maintains can have the following entries:

1. An interface entry that leads to multicast receivers inside of the site.
2. An encapsulation entry that can be targeted to a Delivery Group or a unicast RLOC.

The oif-list entries can be created by the signaling mechanisms defined in [[LISP-MCAST](#)] using the PIM protocol or by the signaling mechanisms in this specification using Map-Requests.

Another option is to have an external orchestration system program the mapping database explicitly so ETR signaling to the ITR can be reduced or even eliminated. Also by the use of Explicit-Locator-Paths (ELPs) [[LISP-TE](#)], LISP-RE capabilities can be explored. For more details see [[LISP-RE](#)].

Since an oif-list can contain either a Delivery Group or a unicast RLOC as a destination address for the outer header, a question is raised where the decision is made to use one or the other, or both.

It is desirable to use multicast routing in the core network where it is available. However, if ETRs are attached to a multicast capable core network, the ITR may not be. In this case, unicast RLOC encapsulation will be necessary to deliver multicast packets directly to the ETR. It will left to the network administrator to configure the decision on Delivery Group versus unicast RLOCs is done by the ETRs, the ITR, or an orchestration system directly programming the mapping database. This specification allows and permits for the ETR to request the encapsulation destination address as well as allowing the ITR to override it.

7. Interworking Considerations

The Map-Request multicast signaling between ETR(s) and an ITR described in this specification is also used by ETR(s) to multicast PITRs which are deployed to support non-LISP multicast source sites. This is true for multicast PETRs that signal to an ITR or mPITR which support non-LISP multicast receiver sites.

8. Security Considerations

The security concerns for LISP multicast are mainly the same as for the base LISP specification [[LISP](#)] and the LISP multicast specification [[LISP-MCAST](#)], including PIM-ASM [[RFC4601](#)].

Where there are security concerns with respect to unicast PIM messages, as discussed in [[LISP-MCAST](#)], the same may also be true for multicast signaling with Map-Request messages.

9. IANA Considerations

At this time there are no requests for IANA.

10. References

10.1. Normative References

- [LISP] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis,
 "Locator/ID Separation Protocol (LISP)",
 [draft-ietf-lisp-24.txt](#) (work in progress).

- [LISP-MCAST] Farinacci, D., Meyer, D., Zwiebel, J., and S. Venaas,
 "LISP for Multicast Environments",
 [draft-ietf-lisp-multicast-14.txt](#) (work in progress).

- [RFC2119] Bradner, S., "Key words for use in RFCs to Indicate
 Requirement Levels", [BCP 14](#), [RFC 2119](#), March 1997.

- [RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas,
 "Protocol Independent Multicast - Sparse Mode (PIM-SM):
 Protocol Specification (Revised)", [RFC 4601](#), August 2006.

10.2. Informative References

- [LISP-LCAF] Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical
 Address Format", [draft-ietf-lisp-lcaf-00.txt](#) (work in
 progress).

- [LISP-RE] Coras, F., Cabellos-Aparicio, A., Domingo-Pascual, J.,
 Maino, F., and D. Farinacci, "LISP Replication
 Engineering", [draft-coras-lisp-re-01.txt](#) (work in
 progress).

- [LISP-TE] Farinacci, D., Lahiri, P., and M. Kowal, "LISP Traffic
 Engineering Use-Cases", [draft-farinacci-lisp-te-02.txt](#)
 (work in progress).

[Appendix A](#). Acknowledgments

The authors would like to thank the following people for their participation in conversations on the topic. They are Gregg Schudel, Florin Coras, Darrel Lewis, and Fabio Maino.

Appendix B. Document Change Log

B.1. Changes to [draft-farinacci-lisp-mr-signaling-01.txt](#)

- o Refreshing references and document timer.

B.2. Changes to [draft-farinacci-lisp-mr-signaling-00.txt](#)

- o Initial draft posted July 2012.

Authors' Addresses

Dino Farinacci
cisco Systems
Tasman Ave.
San Jose, California
USA

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

Maria Napierala
AT&T
Middletown, NJ
USA

Email: mnapierala@att.com

