Network Working Group
Internet-Draft
Intended status: Experimental
Expires: March 13, 2020

J. Leong
D. Lewis
G. Schudel
A. Smirnov
Cisco Systems
C. Cassar
Tesla
I. Kouvelas
Arista Networks Inc.
September 10, 2019

LISP RLOC Membership Distribution
draft-kouvelas-lisp-rloc-membership-02

Abstract

The Locator/ID Separation Protocol (LISP) operation is based on EID to RLOC mappings that are exchanged through a mapping system. The mapping system can use the RLOCs included in mapping registrations to construct the complete set of RLOC addresses across all xTRs that are members of the LISP deployment. This set can then be made available by the mapping system to all the member xTRs. An xTR can use the RLOC set to optimise protocol operation as well as to implement new functionality. This document describes the use of the LISP reliable transport session between an xTR and a Map-Server to communicate the contents of the RLOC membership set.

Table of Contents

1. Introduction

The Locator/ID Separation Protocol (LISP) registration process between an xTR and a Map-Server is defined in [I-D.ietf-lisp-rfc6833bis]. In each registration message the xTR communicates mapping records providing the list of routing locators (RLOCs) that can be used to reach the endpoint identifier (EID) space

behind the xTR. By gleaning the RLOCs from all such registrations, the map-server constructs the set of RLOCs across all the received registrations. This set represents all the RLOCs used to encapsulate traffic and is the complete RLOC membership of the LISP network (limitations described below).

The gleaned RLOC membership set is communicated to the member xTRs where it can be used to implement new functionality as well as to optimise protocol operation. As one example in deployments where the RLOC network provides guarantees against RLOC source address spoofing the membership can be used as a decapsulation filter to prevent injection of traffic by non-members. As a second example, a possible optimisation to existing functionality can use changes to the RLOC membership set to validate the xTR map-cache contents and trigger updates for out-of-date mappings.

Distribution of the RLOC membership set is practical in VPN use cases [I-D.lewis-lisp-vpns] where the number of member xTRs and their RLOCs is bounded thus limiting both the number of membership elements that must be distributed as well as the number of members that the set must be distributed to. In a VPN use case the membership set is specific to each VPN identified through the LISP Instance ID (IID). It is reasonable to expect that all member xTRs for a specific VPN can register against a pair of redundant Map-Servers. The complete membership set will therefore be available on those Map-Servers. Alternatively, registration can be across a small set of Map-Servers that synchronise the RLOC membership set between them (outside the scope of this document). In the general case the RLOC membership knowledge is split across a distributed mapping system [I-D.ietf-lisp-ddt] and its collection and distribution would hit scale limits.

Membership gleaning at the Map-Server assumes symmetric ITR and ETR deployments. All encapsulating ITRs also have to be configured as ETRs registering against the Map-Servers. This is a common way of deploying LISP xTRs. To allow members that do not own EID space (such as exclusive ITRs and proxy routers) to be included in the membership set the registration mechanism must be extended.

Note that automatic membership gleaning at the Map-Server through registrations is just one mechanism that can be used to discover the RLOC set to be distributed. This document focuses on the membership set distribution mechanism.

The LISP extension in [I-D.kouvelas-lisp-reliable-transport] introduces a reliable transport session between the xTR and the MS. The membership set communication described in this document is based on message exchange over the reliable transport.

2. Requirements Notation

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 [RFC2119].

3. Membership Distribution Overview

The RLOC membership set distribution from the Map-Server to the xTR is initiated on demand by the xTR. Unless the xTR specifically subscribes to receive the RLOC membership set no action is taken by the Map-Server. The granularity at which a Map-Server gleans membership, and that an xTR can request its distribution, is per EID address family and instance ID. This matches the VPN EID space segmentation model allowing separate communication of the membership of different VPNs. It also allows for each EID address family to have a different xTR membership.

The Map-Server SHOULD only allow the distribution of the RLOC membership set for an EID instance and address family to xTRs that are valid members of the set being distributed. An xTR that has a reliable transport session established with the Map-Server and is registering EID prefixes with the Map-Server but not for the specific instance ID and EID address family, SHOULD NOT be sent the RLOC membership set.

The set of member RLOCs for an EID address family and instance ID is dynamic and changes as new registrations are received by the Map-Server and as registration state times out. When membership distribution is initiated by the xTR, the complete RLOC set contents is communicated. In parallel updates to the membership set begin being communicated. The membership set updates continue for the duration of the reliable transport session or until the xTR unsubscribes from the membership distribution.

4. Membership Message Format

The membership distribution exchange between the xTR and Map-Server over the reliable transport session relies on a number of new messages defined below. The use of these messages is described in the following sections. The table below lists the messages. All messages carry the EID address family and instance ID for the membership distribution. Some messages additionally carry extra fields that are listed in the table. The new messages are:

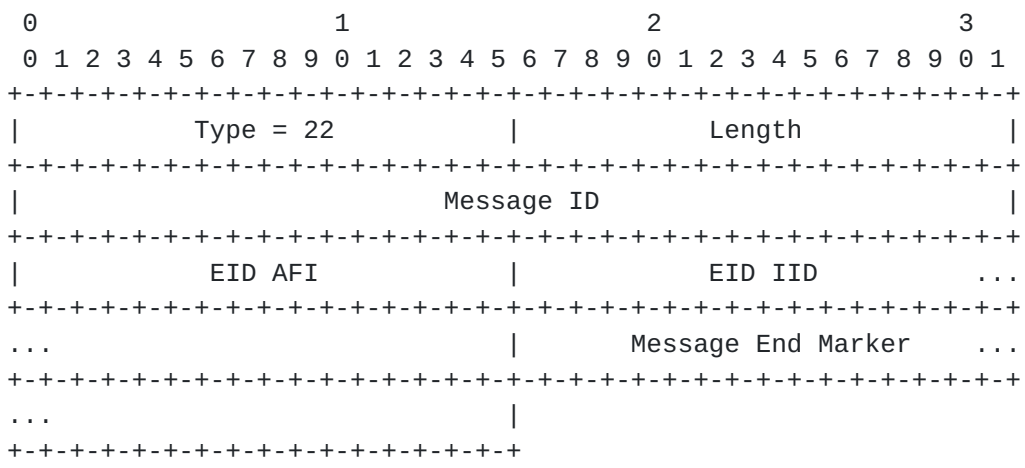
Type	Message	Direction	Additional fields
22	Subscribe	xTR -> MS	
23	Subscribe ACK	MS -> xTR	Subscribe ID
24	Subscribe NACK	MS -> xTR	Subscribe ID, Error
25	Unsubscribe	xTR -> MS	
26	Element Add	MS -> xTR	Site-ID, RLOC
27	Element Delete	MS -> xTR	Site-ID, RLOC
28	Refresh Request	xTR -> MS	
29	Refresh Begin	MS -> xTR	Request ID
30	Refresh End	MS -> xTR	Request ID

Table 1: Reliable transport membership distribution TLVs

The rest of this section provides the format of each of the messages in the table. For a description of the Type, Length, Message ID and Message End Marker fields refer to [I-D.kouvelas-lisp-reliable-transport].

4.1. Membership Subscribe

The Membership subscribe message is sent by the xTR to the Map-Server to initiate RLOC membership set distribution for a specific EID AFI and instance ID.



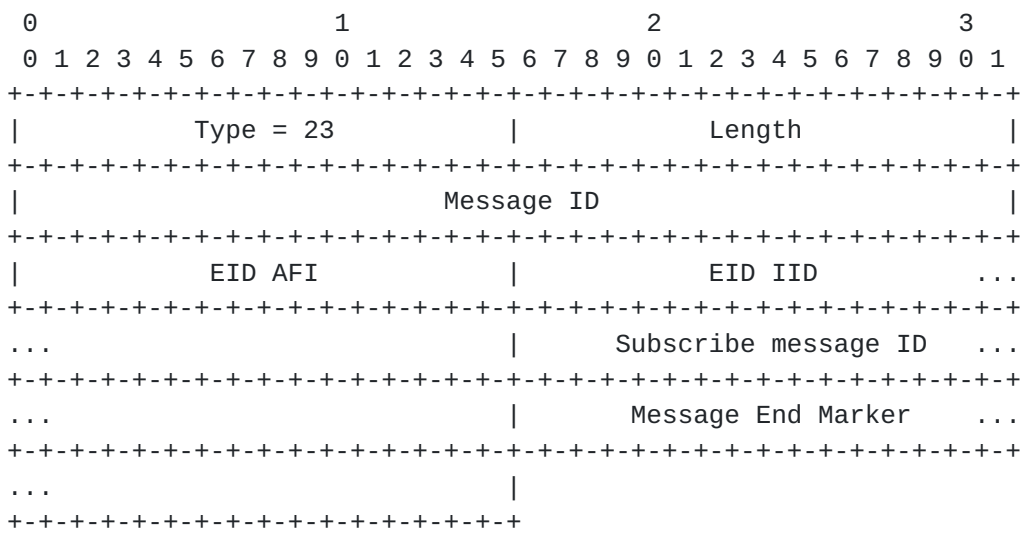
Membership subscribe message format

- EID-AFI: EID address family for which the membership is being requested.

- EID IID: The EID instance ID identifying the VPN for which the membership is being requested [I-D.lewis-lisp-vpns]. Although the IID is only 24 bits in size in the data encapsulation, it is being defined as a 32 bit field for consistency with the LCAF Instance ID header [I-D.ietf-lisp-lcaf].

4.2. Membership Subscribe ACK

The Membership-Subscribe-ACK message is sent by the Map-Server to the xTR to acknowledge acceptance of a Membership-Request. This message indicates that the Map-Server will be providing the requested membership to the xTR.

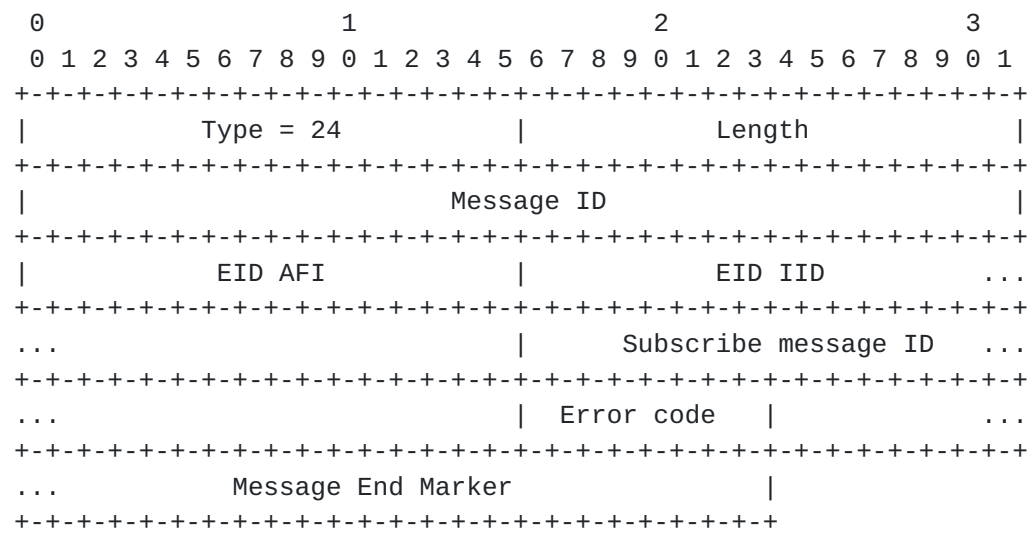


Membership-Subscribe-ACK message format

- Subscribe message ID: The message ID carried over from the membership subscribe message.

4.3. Membership Subscribe NACK

The Membership-Subscribe-NACK message is sent by the Map-Server to the xTR to reject a membership request. This message indicates that the Map-Server will not be providing the requested membership to the xTR. The membership subscribe NACK message can be sent at any point following the receipt of a Membership-Subscribe message. The Map-Server may initially acknowledge a subscription with a Membership Subscribe ACK and later when conditions change cancel the subscription by issuing a membership subscribe NACK message.



Membership subscribe NACK message format

- Subscribe message ID: The message ID carried over from the membership subscribe message.

- Error code: The error code provides a reason for which the registration was rejected by the Map-Server. Defined values are:

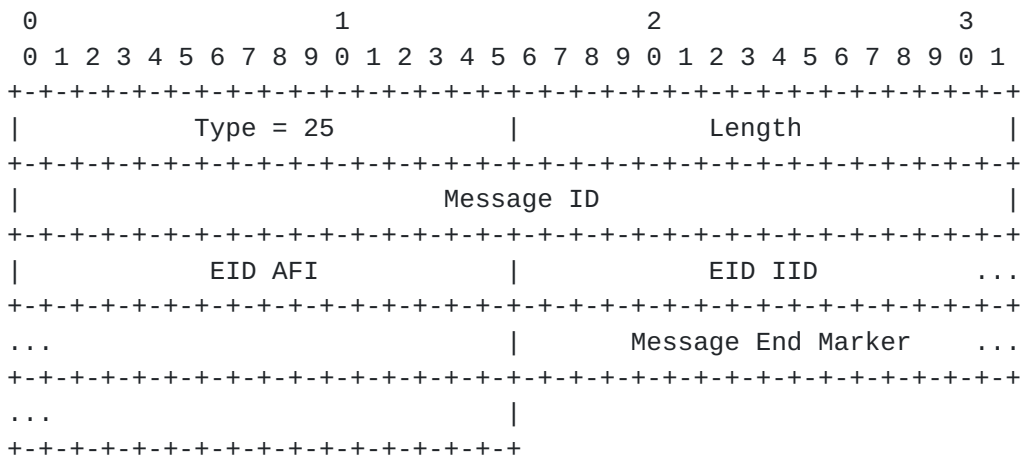
  1 - Not found: The EID instance and address family do not match the Map-Server configuration.

  2 - Not enabled: The Map-Server is not configured to allow membership distribution for the requested EID instance and address family.

  3 - Not authorized: The xTR that sent the request does not have a valid registration under the EID instance and address family.

4.4. Membership Unsubscribe

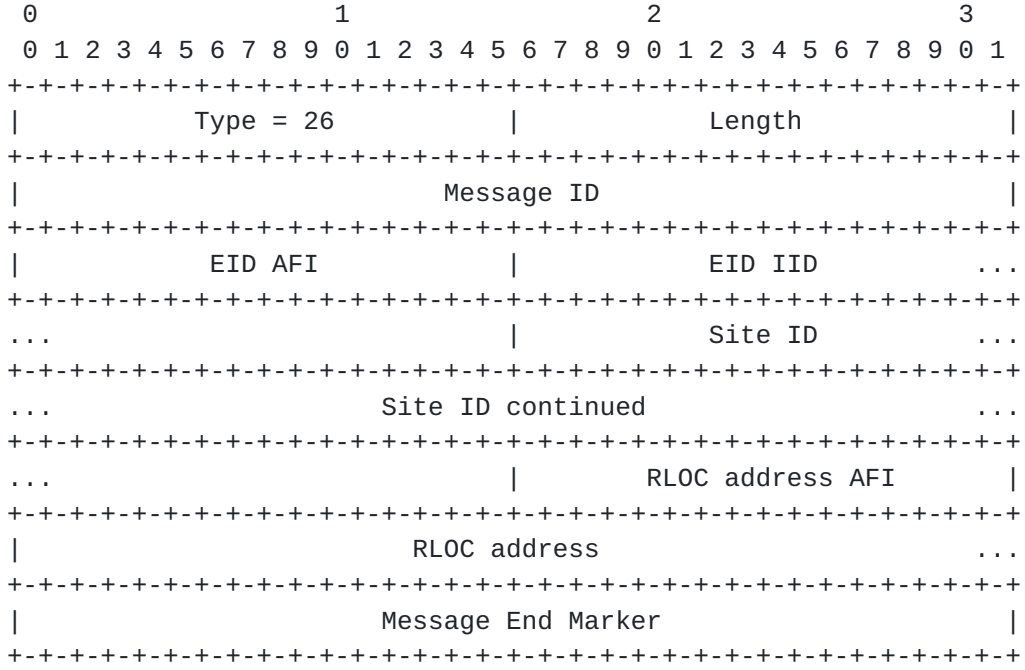
The Membership-Unsubscribe message is sent by the xTR to the Map-Server to terminate RLOC membership set distribution for a specific EID AFI and instance ID.



Membership-Unsubscribe message format

4.5. Membership Element Add

The Membership-Element-Add message is sent by the Map-Server to the xTR to communicate a single RLOC that is a member of the set for the specified EID instance and address family.



Membership-Element-Add message format

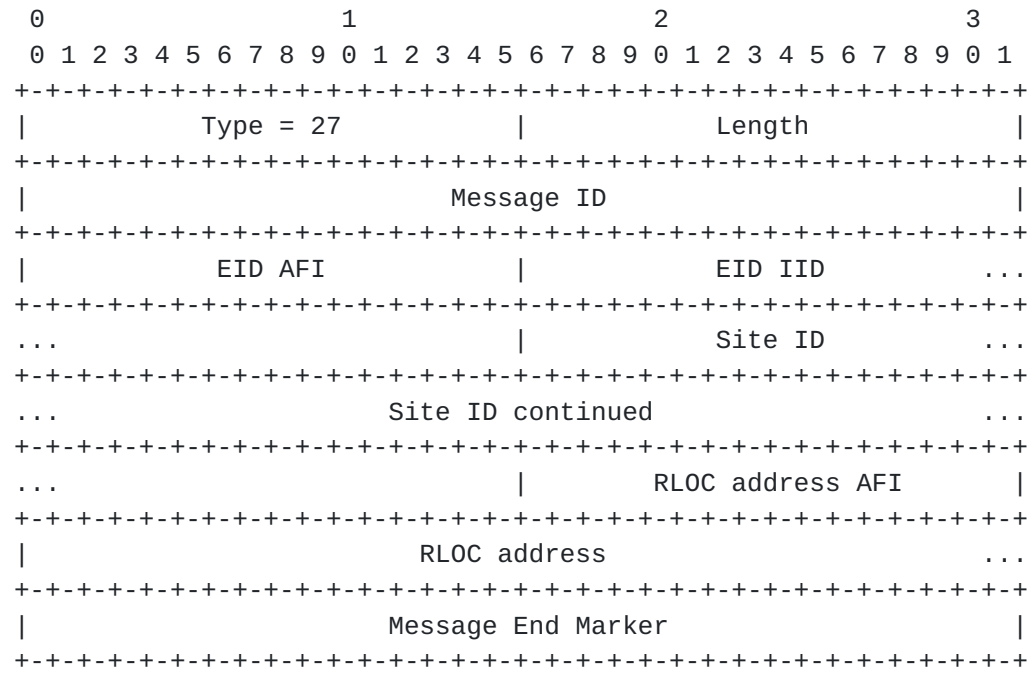
- Site ID: The 64 bit site ID value from the mapping registration that contributed this RLOC to the membership list. The site ID can be used by the receiving xTR to derive information about the grouping of member RLOCs to remote sites.

- RLOC address AFI: Address family identifier for the RLOC address in the following field.

- RLOC address: The actual RLOC membership set element address being communicated. Note that the length of this field depends on the RLOC address AFI in the preceding field.

4.6. Membership Element Delete

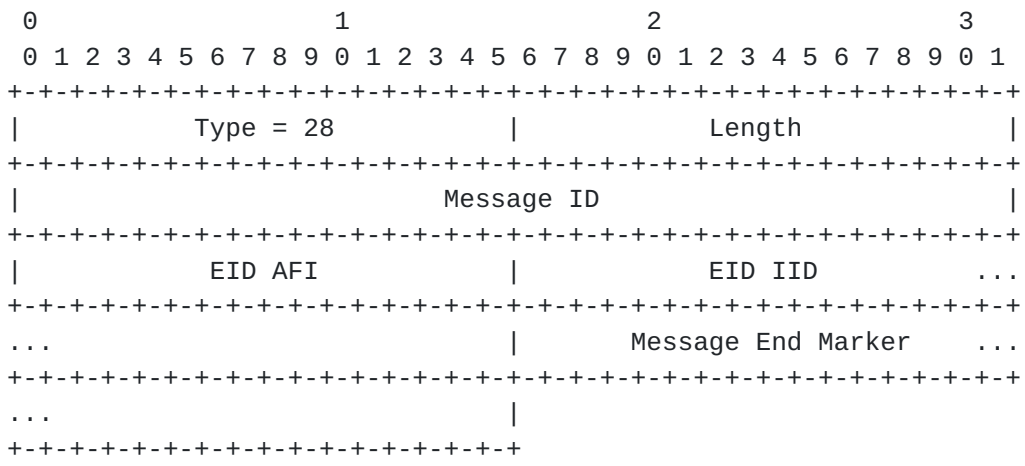
The Membership-Element-Delete message is sent by the Map-Server to the xTR to communicate a single RLOC that is no longer a member of the set for the specified EID instance and address family.



Membership-Element-Delete message format

4.7. Membership Refresh Request

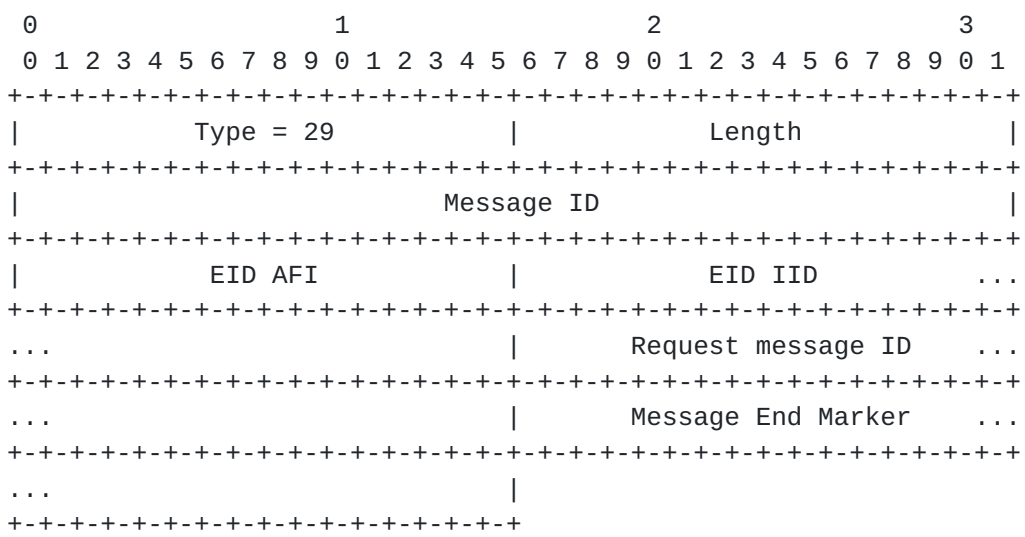
The Membership-Refresh-Request message is sent by the xTR to the Map-Server to request that the Map-Server send the complete RLOC membership set contents for the specified instance ID and AFI.



Membership-Refresh-Request message format

4.8. Membership Refresh Begin

The Membership-Refresh-Begin message is sent by the Map-Server to the xTR to acknowledge an earlier Membership-Refresh-Request message and to indicate that the following membership updates are part of the refresh.

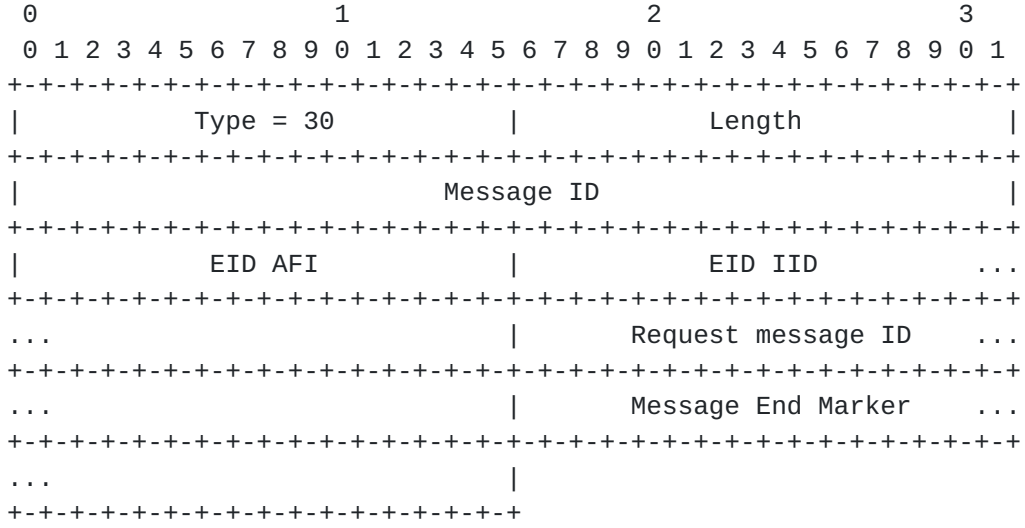


Membership-Refresh-Begin message format

- Request message ID: The message ID carried over from the membership request message.

4.9. Membership Refresh End

The Membership-Refresh-End message is sent by the Map-Server to the xTR to indicate that the communication of the full membership refresh for the specified EID instance ID and AF is now complete.

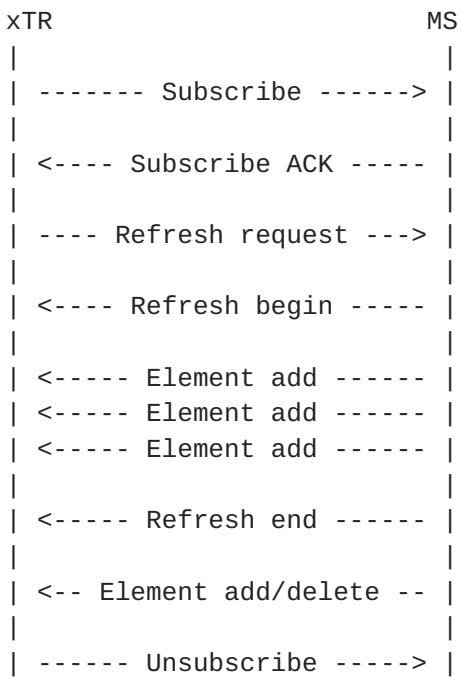


Membership-Refresh-End message format

- Request message ID: The message ID carried over from the membership request message.

5. Membership Distribution Message Exchange

Following the reliable transport session establishment, the EID membership communication relies on the exchange of the membership messages defined in the previous section. The description in this section presents the exchange from the perspective of a single xTR and Map-Server.



Typical membership distribution message exchange

The xTR starts the exchange by issuing a Membership-Subscribe-Request message to the Map-Server for a specific EID instance. Assuming the Map-Server is configured to allow membership distribution and the requesting router is authorized to receive the membership of the EID instance, the MS will reply with a Membership-Subscribe-ACK. After sending the ACK, the MS will start sending to the xTR Membership-Element-Add and Membership-Element-Delete messages corresponding to changes of the EID instance membership.

On receipt of the Membership-Subscribe-ACK message, the xTR issues a Membership-Refresh-Request message in order to receive the complete contents of the EID instance membership held by the MS. The MS responds to the Membership-Refresh-Request by issuing a Membership-Refresh-Begin message, followed by a Membership-Element-Add message for each member of the EID instance and finally completes the refresh by sending a Membership-Refresh-End message.

On receipt of Membership-Element-Add and Membership-Element-Delete messages, the xTR updates its membership database for the EID instance ID and address family by adding or deleting the entry corresponding to the communicated RLOC address. Note that the membership state on the xTR is Map-Server specific and the xTR has to maintain separate RLOC membership entries received from each Map-Server it subscribes with.

When the xTR receives the Membership-Refresh-End message it purges all the stale membership entries it may have obtained during a previous session instantiation that were not updated during the refresh.

The MS may issue Membership-Element-Add and Membership-Element-Delete messages corresponding to membership changes at any point after issuing the Membership-ACK message, even during a refresh.

The xTR may request additional full refreshes of the complete membership set at any point after having received a Membership-Subscribe-ACK message by issuing a new Membership-Refresh-Request.

When the Map-Server determines that an xTR is no longer eligible to receive membership updates, for example the EID instance and address family registration state of the xTR becomes invalid, then the Map-Server SHOULD send it a Membership-NACK message to indicate the termination of the membership communication.

6. Implementation Status

[Note to RFC Editor: Please remove this section and the reference to [RFC6982] before publication.]

This section records the status of known implementations of the LISP RLOC Membership Distribution at the time of posting of this Internet-Draft, and is based on a proposal described in [RFC6982].

The description of implementations in this section is intended to assist the IETF in its decision processes in progressing drafts to RFCs.

Cisco has a production implementation of the RLOC membership distribution mechanism described in this draft on IOS, IOS-XE and IOS-XR. The RLOC membership information is used to implement the data plane security functionality described in: LISP Data Plane Security [1]. For additional information please contact lisp-support@cisco.com.

7. Security Considerations

The RLOC membership distribution message communication takes place over a LISP reliable transport connection. The security mechanisms of the reliable transport apply to this solution.

8. IANA Considerations

The following message types must be assigned out of the space defined in [I-D.kouvelas-lisp-reliable-transport].

Type	Name	Reference
22	Membership Subscribe	This document
23	Membership Subscribe ACK	This document
24	Membership Subscribe NACK	This document
25	Membership Unsubscribe	This document
26	Membership Element Add	This document
27	Membership Element Delete	This document
28	Membership Refresh Request	This document
29	Membership Refresh Begin	This document
30	Membership Refresh End	This document

9. Acknowledgments

The authors would like to thank Michiel Blokzijl, Selina Heimlich, Vasileios Lakafosis, Fabio Maino, Andre Pelletier, Jesper Skriver and Chao Yu, for their contributions to this specification.

10. References

10.1. Normative References

[I-D.ietf-lisp-rfc6833bis]
Farinacci, D., Maino, F., Fuller, V., and A. Cabellos-Aparicio, "Locator/ID Separation Protocol (LISP) Control-Plane", draft-ietf-lisp-rfc6833bis-25 (work in progress), June 2019.

[I-D.kouvelas-lisp-reliable-transport]
Cassar, C., Kouvelas, I., and D. Lewis, "LISP Reliable Transport", draft-kouvelas-lisp-reliable-transport-02 (work in progress), March 2015.

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, DOI 10.17487/RFC2119, March 1997, <https://www.rfc-editor.org/info/rfc2119>.

10.2. Informative References

[I-D.ietf-lisp-ddt]
Fuller, V., Lewis, D., Ermagan, V., Jain, A., and A. Smirnov, "LISP Delegated Database Tree", draft-ietf-lisp-ddt-09 (work in progress), January 2017.

[I-D.ietf-lisp-lcaf]
Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", draft-ietf-lisp-lcaf-22 (work in progress), November 2016.

[I-D.lewis-lisp-vpns]
Lewis, D. and G. Schudel, "LISP Virtual Private Networks (VPNs)", draft-lewis-lisp-vpns-00 (work in progress), February 2014.

[RFC6982] Sheffer, Y. and A. Farrel, "Improving Awareness of Running Code: The Implementation Status Section", RFC 6982, DOI 10.17487/RFC6982, July 2013, <https://www.rfc-editor.org/info/rfc6982>.

10.3. URIs

[1] http://www.cisco.com/c/en/us/td/docs/ios-xml/ios/iproute_lisp/configuration/xe-3s/irl-xe-3s-book/irl-data-plane-sec.html

Authors' Addresses

Johnson Leong
Cisco Systems
Tasman Drive
San Jose, CA 95134
USA

Email: joleong@cisco.com

Darrel Lewis
Cisco Systems
Tasman Drive
San Jose, CA 95134
USA

Email: darlewis@cisco.com

Gregg Schudel
Cisco Systems
Tasman Drive
San Jose, CA 95134
USA

Email: gschudel@cisco.com

Anton Smirnov
Cisco Systems
Tasman Drive
San Jose, CA 95134
USA

Email: asmirnov@cisco.com

Chris Cassar
Tesla
10 New Square Park
Bedfont Lakes, Feltham TW14 8HA
United Kingdom

Email: christiancassar@acm.org

Isidor Kouvelas
Arista Networks Inc.
5453 Great America Parkway
Santa Clara, CA 95054
USA

Email: isidor@kouvelas.net