

Workgroup: Internet Engineering Task Force  
Internet-Draft:  
draft-vda-lisp-underlay-multicast-trees-00.xml

Published: 11 March 2023

Intended Status: Experimental

Expires: 12 September 2023

Authors: V. Govindan, Ed.    D. Farinacci    A. Kuppusami  
         Cisco                    lispers.net    Cisco

## **Enhancements to Signal-Free Locator/ID Separation Protocol (LISP) Multicast**

### **Abstract**

This draft augments the signal-free LISP Multicast RFC 8378 and RFC6831 to support multicast underlays between LISP sites. This draft defines the many-to-1, 1-to-many, and many-to-many relationship between multicast EIDs and the Replication List Entries (RLEs) RLOC records they map to. The mechanisms in this draft allow a multicast LISP overlay to run over a mixed underlay of unicast and/or multicast functionality.

### **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 <https://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 12 September 2023.

### **Copyright Notice**

Copyright (c) 2023 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 (<https://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 Revised BSD License text as described in Section 4.e of the Trust Legal Provisions and are provided without warranty as described in the Revised BSD License.

## Table of Contents

- [1. Introduction](#)
  - [1.1. Requirements Language](#)
- [2. Definition of Terms](#)
- [3. Additions to General Procedures](#)
  - [3.1. Receiver site procedures](#)
  - [3.2. Consolidation of the Replication list](#)
  - [3.3. Source-site Procedures](#)
    - [3.3.1. Multicast Tree-building at source site procedures](#)
  - [3.4. Combinations for the RLE-set](#)
  - [3.5. Sequence diagram](#)
- [4. Acknowledgements](#)
- [5. Contributors](#)
- [6. IANA Considerations](#)
- [7. Security Considerations](#)
- [8. Normative References](#)
- [Authors' Addresses](#)

## 1. Introduction

When a multicast capable underlay connects multiple LISP sites, we can take advantage of the multicast capabilities and perform replication more efficiently than using head-end replication. This draft addresses the problem of selecting the underlay multicast group(s), to transport a given overlay multicast flow. There are 4 different scenarios possible:

1. A 1:1 mapping of a overlay multicast flow, either a (Source, Group) or (\*, Group) to an underlay group.
2. A many:1 mapping where many overlay multicast flows share the same underlay group.
3. A 1: many mapping where a single overlay group is transported over different underlay groups. Note in this case, the "many" means mapping one EID to multiple RLOCs in a RLE-set where the set can be a mix of underlay groups and underlay unicast addresses. It is really important that this algorithm compute unicast RLOCs as well as multicast RLOCs.
4. A special (actually realistic) case is the combination of 2 and 3 where m:n mapping is achieved. Typically we expect  $m \gg n$ .

There are two methods proposed to derive such mappings described above:

\*Computation based underlay group assignment using hash functions: Here all ETRs will use the same multicast underlay group, when they are all on the same underlay e.g we hash (S-EID, G) Multicast EID, and append either 24 bits in terms of IPv4 or 120 bits in terms of IPv6. If we have disjoint native multicast underlays, we may have to discuss more complicated schemes e.g. different address ranges for IPv4 v/s IPv6 or different ranges inside either of these as well. Since this is a very simple and powerful method, we want to include this design in the current draft.

\*Lookup based group assignment: The LISP signal free mechanism can be extended to have the co-ordinated assignment of underlay group(s) for a given overlay multicast flow. To implement this, we need a signaling mechanism that is independent of any multicast routing protocol that runs in the LISP underlay (e.g PIM)

The scope of this draft covers underlays based on IPv4 and IPv6 only. It does not cover other transport mechanisms like BIER or MPLS or Layer-2 underlays.

Note: All terminology used in this document are based on the definitions of [RFC8378](#) [[RFC8378](#)].

### 1.1. 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. Definition of Terms

1. (S-EID,G) State: refers to multicast state in multicast source and receiver sites where S-EID is the IP address of the multicast source host (its EID). An S-EID can appear in an IGMPv3 report, an MSDP SA message or a PIM Join/Prune message that travels inside of a site.
2. (S-RLOC,G) State: refers to multicast state in the core where S is a source locator (the IP address of a multicast ITR) of a site with a multicast source. The (S-RLOC,G) is mapped from the (S-EID,G) entry by doing a mapping database lookup for the EID- Prefix that S-EID maps to. An S-RLOC can appear in a PIM Join/ Prune message when it travels from an ETR to an ITR over the Internet core.

### 3. Additions to General Procedures

A set of receivers spread across multiple LISP sites having interest on a given (S-EID, G-EID) can be generically represented by the mapping below:

`*(S-EID,G-EID) -> RLE [(S-RLOC1, G-RLOCa), (S-RLOC1, G-RLOCb), (S-RLOC1, U-RLOCc), (S-RLOC2, U-RLOCd)]`

The following additions are made to the General Procedures defined in [RFC 8378](#) [[RFC8378](#)]:

#### 3.1. Receiver site procedures

The receiver site procedures require the following extensions to that outlined in [RFC 8378](#) [[RFC8378](#)]:

The Map-Register messages [RFC 8060](#) [[RFC8060](#)] sent by the Receiver-ETRs MUST have the following fields set as specified here:

1. The RLOC in the Map-Register message MUST be encoded using the RLE LCAF Type defined in [RFC 8060](#) [[RFC8060](#)]. The address encoded indicates that the RLOC is requesting the flow to be mapped to an underlay multicast group.
2. want-map-notify bit (M) set to 1. Unlike the use-case of the unicast underlay, the Receiver-ETR MUST be map-notified about the initial assignment of the underlay multicast group(s) and subsequent changes if any to the replication list.

\*Comment from Dino: We need the mapping system to tell ITRs when the RLE-set changes so it can update its map-cache. This was part of the LISP architecture before we had PubSub, but PubSub uses the same technique. It is just implicit for multicast where with PubSub you send subscribe-requests. So all you have to specify is that each ETR that registers mappings sets the M-bit, but now that I am writing this the M-bit is used to acknowledge the Map-Register. So you don't need to set it for updates to mapping entries, because when you register an (S,G) EID encoding, the map-server knows to send RLOC/RLE changes. It knows where to send them because the S-EID is registered in the mapping system, so it knows the RLOCs for that site. But the Merge-bit is crucial or else, from above example, RLOCc and RLOCd could not both be added to the RLE set. So the map-server knows to merge in U-RLOCc from RLOCc and U-RLOCd from RLOCd when they come separately from the ETRs. Ditto, for the entity that registers G-RLOCa and G-RLOCb (could be ETRs or could be a controller somewhere).

3. The consolidated replication list of underlay group(s) assigned to transport the overlay (S-EID, G) is constructed as a RLE by the mapping system and sent as Map-notification to the ETR.
4. After the above step, the mapping system sends a Map-notify message containing a list of underlay group addresses encoded as RLE LCAF [RFC 8060](#) [[RFC8060](#)] elements.
5. The ETR can then join these underlay group addresses to receive the traffic for the multicast entry EIDs.

Another method of explaining the sequence is as below:

1. The ETR receives an IGMPv3 (S,G) report. That should be spec'ed as (S-EID, G-EID).
2. The ETR will hash that 2-tuple to get a G-RLOC.
3. The ETR sends a Map-Register for (S-EID,G-EID) with G-RLOC.
4. Other ETRs will do the same so only when the last member leaves the G-EID, then there will be no (S-EID,G-EID) registered.
5. If one ETR (call it x) doesnt think it can get packets natively via G-RLOC, then it will register (U-RLOCx). That is, its IP address that can get it multicast packes on the unicast underlay.

TODO: Merge the above two sequences later

### 3.2. Consolidation of the Replication list

1. After the mapping system merges all Receiver-ETR or delivery-group RLOCs to build the comprehensive replication list, it allocates one or more underlay group addresses to enable underlay multicast transport for the overlay flow. The allocated group(s) are then notified to both ITR and ETR using the procedures listed in the respective sections.
2. The mapping system MUST NOT merge any duplicate RLOC requests for the unicast and multicast level value fields.
3. With a mixed RLE-set, a generic representation of the mapping is of the form below:

\*(S-EID,G-EID) -> RLE [(S-RLOC1, G-RLOCa), (S-RLOC1, G-RLOCb), (S-RLOC1, U-RLOCc), (S-RLOC2, U-RLOCd)]

4.

The above notation is really important because for each replication that is encapsulated, the packet could go out a different interface which means the source address in the outer header is different (e.g. this ITR has RLOC1 and RLOC2 on each interface). There may be different pockets of native multicast connectivity in the underlay, so we may have to replicate to different (G-RLOCi) and hence the notation above of G-RLOCa and G-RLOCb. The unicast RLOCs are included above because RLOCc and RLOCd do not have native multicast connectivity to them so they need to use the unicast underlay.

### **3.3. Source-site Procedures**

#### **3.3.1. Multicast Tree-building at source site procedures**

The following points are added:

1. The source site must include both underlay multicast groups allocated for and unicast RLOCs in the format TBD.
2. When an ITR receives a packet with header addresses (S-EID, G-EID), it does an (S-EID, G-EID) lookup in the mapping system by sending a Map-Request.
3. When a Map-Reply is returned, there will be 1 RLOC-record in the RLOC-set. The RLOC-record is encoded as an RLE LCAF per [RFC 8060](#) [[RFC8060](#)].
4. The ITR will replicate a packet for each entry in the RLE-set and encapsulate the replicated packet to each G-RLOC and U-RLOC in the RLOC-set.

#### **3.4. Combinations for the RLE-set**

Placeholder text from Dino: We need a section that just focuses on all the combination of the RLE-set. And we have to recommend how to setup the RLE-set in an ETR when it doesn't know how any ITR will get multicast packets to it (via G-RLOC or U-RLOC).

#### **3.5. Sequence diagram**

TBD

### **4. Acknowledgements**

TBD

## 5. Contributors

Stig Venaas  
Cisco

Email: [svenaas@cisco.com](mailto:svenaas@cisco.com)

## 6. IANA Considerations

No new requests to IANA

## 7. Security Considerations

TBD

## 8. Normative References

- [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>>.
- [RFC6831] Farinacci, D., Meyer, D., Zwiebel, J., and S. Venaas, "The Locator/ID Separation Protocol (LISP) for Multicast Environments", RFC 6831, DOI 10.17487/RFC6831, January 2013, <<https://www.rfc-editor.org/info/rfc6831>>.
- [RFC8059] Arango, J., Venaas, S., Kouvelas, I., and D. Farinacci, "PIM Join Attributes for Locator/ID Separation Protocol (LISP) Environments", RFC 8059, DOI 10.17487/RFC8059, January 2017, <<https://www.rfc-editor.org/info/rfc8059>>.
- [RFC8060] Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", RFC 8060, DOI 10.17487/RFC8060, February 2017, <<https://www.rfc-editor.org/info/rfc8060>>.
- [RFC8378] Moreno, V. and D. Farinacci, "Signal-Free Locator/ID Separation Protocol (LISP) Multicast", RFC 8378, DOI 10.17487/RFC8378, May 2018, <<https://www.rfc-editor.org/info/rfc8378>>.

## Authors' Addresses

Vengada Prasad Govindan (editor)  
Cisco

Email: [venggovi@cisco.com](mailto:venggovi@cisco.com)

Dino Farinacci

lispers.net

Email: [farinacci@gmail.com](mailto:farinacci@gmail.com)

Aswin Kuppusami

Cisco

Email: [aswkuppu@cisco.com](mailto:aswkuppu@cisco.com)