Internet Engineering Task Force
Internet-Draft
Intended status: Experimental
Expires: November 15, 2018

Y. Filyurin
R. Raszuk
Bloomberg LP
T. Boyes
MLB
D. Farinacci
lispers.net
May 14, 2018

LISP Explicit Locator Path (ELP) Probing
draft-filyurin-lisp-elp-probing-01

Abstract

This document describes a LISP-TE mechanism to probe an Explicit Locator Path (ELP) for reachability and telemetry data. The mechanism is called ELP-Probing.

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 November 15, 2018.

Table of Contents

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

This document describes traffic engineering features of the Locator/Identifier Separation Protocol (LISP), which provides a set of functions for routers to exchange information used to map from non globally routable Endpoint Identifiers (EIDs) to routable Routing Locators (RLOCs). The LISP protocol also defines a mechanism for LISP routers to encapsulate IP packets addressed with EIDs for transmission across the Internet that uses RLOCs for routing and forwarding.

When LISP routers encapsulate packets to other LISP routers, the path stretch is typically 1, meaning the packet travels on a direct path from the encapsulating ITR to the decapsulating ETR at the destination site. The direct underlay path is determined by the underlying routing protocol and metrics it uses to find the shortest path.

This specification will examine how reencapsulating tunnels [RFC6830] [I-D.ietf-lisp-te] can be used so a packet can take an administratively specified path, a congestion avoidance path, a failure recovery path, or multiple load-shared paths, as it travels from ITR to ETR. By using an Explicit Locator Path (ELP) encoding [RFC8060] and the use of ELP-probing described in this document, an ITR can encapsulate a packet on a pre-determined path to a Reencapsulating Tunnel Router (RTR) which decapsulates the packet, then encapsulates it to the next locator in the ELP path.

3. Definition of Terms

Reencapsulating Tunnel Router (RTR): An RTR is a router that acts as an ETR (or PETR) by decapsulating packets where the destination address in the "outer" IP header is one of its own RLOCs. Then acts as an ITR (or PITR) by making a decision where to encapsulate the packet based on the next locator in the ELP towards the final destination ETR. In this document, an RTR and ELP-node are terms used interchangeably.

Explicit Locator Path (ELP): The ELP is an explicit list of RLOCs for each RTR a packet must travel to along its path toward a final destination ETR (or PETR). The list is a strict ordering where each RLOC in the list is visited. However, the path from one RTR to another is determined by the underlying routing protocol and how the infrastructure assigns metrics and policies for the path.

Recursive Tunneling: Recursive tunneling occurs when a packet has more than one LISP IP header. Additional layers of tunneling MAY be employed to implement traffic engineering or other re-routing as needed. When this is done, an additional "outer" LISP header is added and the original RLOCs are preserved in the "inner" header. Any references to tunnels in this specification refers to dynamic encapsulating tunnels and they are never statically configured.

Reencapsulating Tunnels: Reencapsulating tunneling occurs when an ETR removes a LISP header, then acts as an ITR to prepend another LISP header. Doing this allows a packet to be re-routed by the reencapsulating router without adding the overhead of additional tunnel headers. Any references to tunnels in this specification refers to dynamic encapsulating tunnels and they are never statically configured. When using multiple mapping database systems, care must be taken to not create reencapsulation loops through misconfiguration.

RLOC-Probing: An RLOC-probe request is a Map-Request with the probe-bit set that is sent from an encapsulator (an ITR, PITR, or

RTR) to a decapsulator (an ETR, PETR, RTR) to test for reachability among other functions. A RLOC-probe reply is a Map-Reply with the probe-bit set that responds to the ITR-RLOC field of the Map-Request. RLOC-probes are sent between RTRs listed in an ELP list.

ELP-Probing: Is an RLOC-probe that is encapsulated as a LISP data packet sent along the ELP path. Each ELP node of of an ELP path adds telemetry information to the ELP-probe message that has been gathered from RLOC-probing.

4. Overview

LISP-TE functionality [I-D.ietf-lisp-te] describes how reencapsualting LISP routers can be used to traffic engineer a network. By using an overlay approach, much of the underlay topology can be traversed with no special consideration or modification. Coarse grain traffic engineering, versus hop-by-hop traffic engineering, can be accomplished in a simple and unobtrusive manner.

If paths in the network can be constructed out-of-band and stored in the LISP mapping system as ELP RLOC-records, then an encapsulator can solely make a decision which paths an encapsulated packet can take. This approach requires no extra overhead in the data packet. How the encapsulator decides on which paths may be based on the telemetry data returned from ELP-Probing.

When an ITR does a lookup to the LISP mapping system, an EID-to-RLOC mapping is returned. The mapping has a set of RLOC records that can each be encoded as an Explicit Locator Path (ELP). When the best priority of each RLOC-record is the same, the ITR can decide which ELP path to use for forwarding. The ITR sends ELP-Probes on each ELP to gather data to choose either a best path or a policy defined path.

If an EID-to-RLOC mapping has two RLOC-records, each with the ELPs (A, B, C, ETR) and (X, Y, Z, ETR), the ITR would send an ELP-Probe on each ELP path. For the first path, the ITR would encapsulate an ELP-Probe message to RTR A. RTR A would decapsulate the packet, add any telemetry data it has gathered from RLOC-Probes to RTR B, and then encapsulate the ELP-Probe to RTR B. This continues until the ETR receives the ELP-Probe and simply sends an ELP-Probe reply back to the ITR. The ITR follows the same procedures for the second ELP path that starts with RTR X.

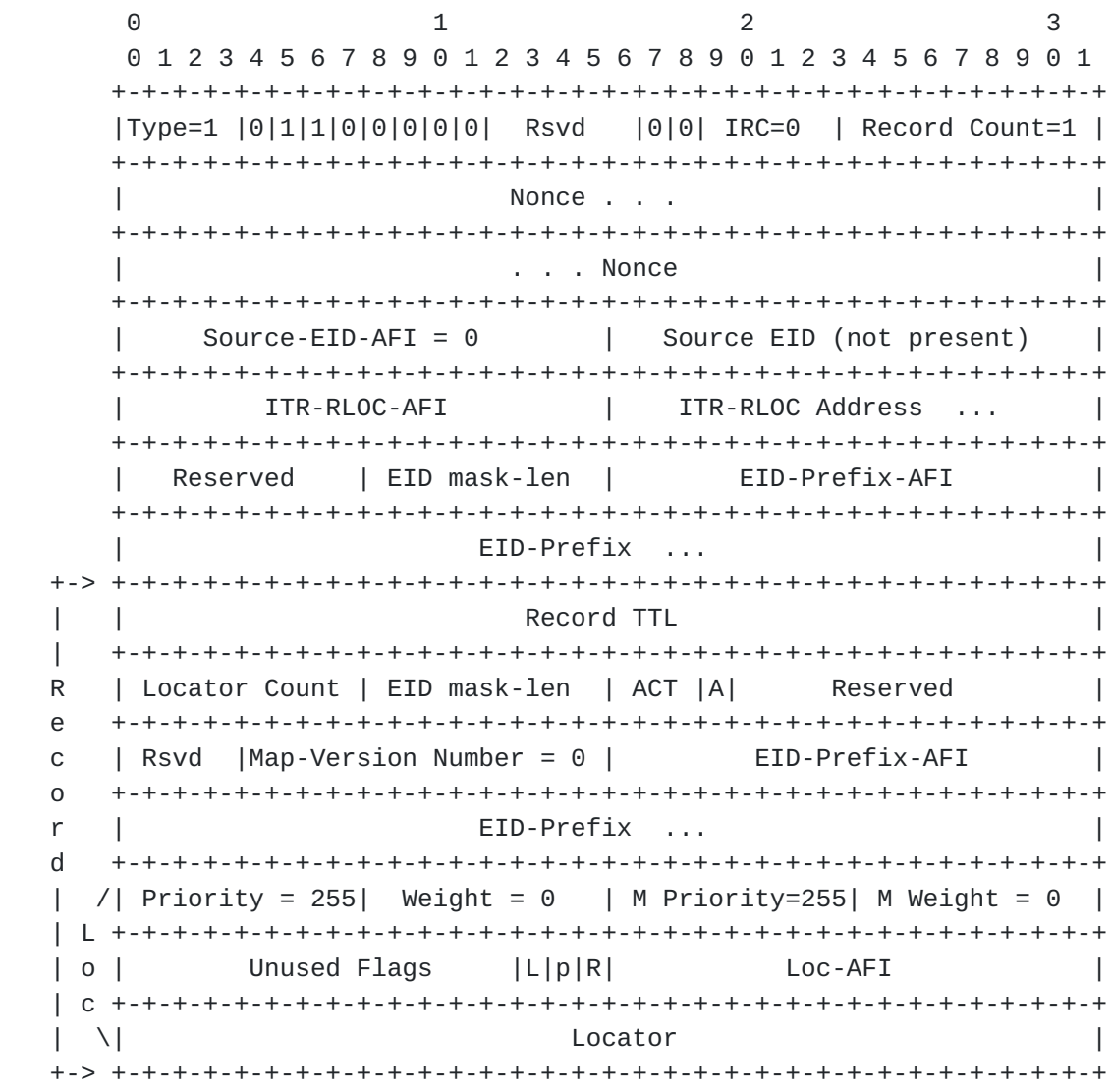
5. RLOC-Probing

The general procedure for RLOC-probing is described in [I-D.ietf-lisp-rfc6830bis]. RLOC-probes are sent between RTRs in an ELP when the P-bit is set in the ELP-node entry of the ELP list [RFC8060]. ELP-Probing depends on an RTR sending RLOC-probes to the next RTR in the ELP list. To get full telemetry data from each ELP-node hop, this specification recommends that the P-bit is set in each ELP-node listed in an ELP.

The ELP-nodes do RLOC-probing asynchronously to gather reachability and RTT data from the next ELP hop. So that when an ELP-Probe is received, the ELP-node has some measured data to add to the ELP-Probe message.

6. ELP-Probing

See [I-D.ietf-lisp-rfc6833bis] for the general format of an RLOC-probe Map-Request. An ELP-Probe message has the following format:



A ELP-Probe message is an RLOC-Probe Map-Request encapsulated with a LISP data-plane header to port 4341 [RFC6830]. The TTL in the outer header must be set to 255. The Instance-ID in the LISP data-plane header must be 0xffffff. The specific field settings for an ELP-Probe in a Map-Request message are:

M-bit: Is set specifying there is an EID-record after the requesting EID-prefix.

P-bit: Is set specifying this Map-Request is an RLOC-probe message being used for ELP-probing.

Source EID: Is not specified by setting the Source-EID-AFI to 0.

EID-Prefix: Is the EID prefix stored in the mapping system that corresponds to a RLOC-set with ELPs imbedded.

EID-Record (Record): Inserted by the originator of an ELP-Probe message. The Locator Count is 0 and the EID-prefix is the same as the EID-prefix earlier in the message.

Record TTL, ACT, A: Not used therefore sent as 0 and ignored on receipt.

RLOC-Record (Loc): Each ELP-node will append an RLOC-record that holds its telemetry data. The Loc-AFI will be the AFI of a LISP Canonical Address Format (LCAF) [RFC8060].

L, p, R bits: All set to 0 and ignore on reception.

As the ELP-Probe moves from RTR to RTR, each RTR adds an RLOC-record to the EID-record in the Map-Request. The RLOC-record will use a LCAF JSON Type [RFC8060] format. Each RTR constructs the following JSON string:

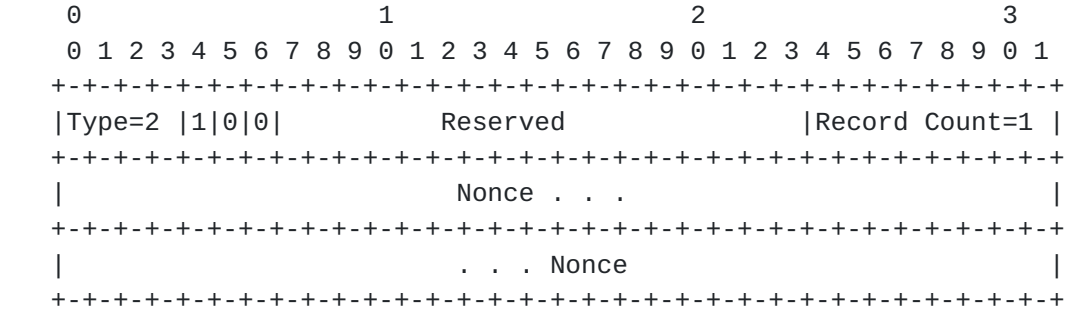
```
{ "ELP-node" : "<rloc>", "HOPs" : "<hc>", "RTTs": ["<rtt1>", ..., "<rttn>"] }
```

ELP-node: Contains the same RLOC address as listed in the ELP.

HOPs: Is the number of underlay hops to this ELP-node from either the last ELP-node or the originator of the ELP-Probe. The value is computed as 255 minus the arrival TTL value in the outer header of the ELP-Probe message.

RTTs: A list of round-trip-times to the next ELP-node. Ordered from recent to less recent.

A ELP-Probe Map-Reply message has the following format. The EID-record is copied from the ELP-Probe Map-Request after the following header:



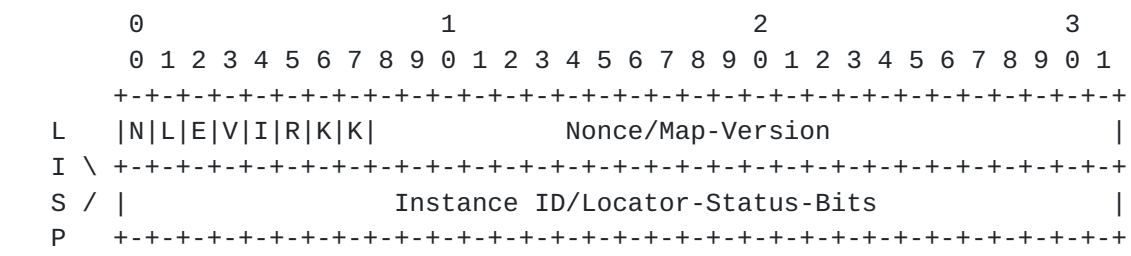
P-bit: This bit is set indicating this is a RLOC-probe message used for ELP-probing.

Nonce: Copied from the ELP-Probe Map-Request nonce.

The last ELP-node in an ELP sends the ELP-Probe Map-Reply to the ITR-RLOC address from the ELP-Probe Map-Request with a source port 4342 and destination port equal to the ephemeral port from the source port of the ELP-Probe Map-Request. Optionally, the ELP-Probe Map-Reply can be data encapsulated to destination port 4341 but if the ELP-Probe originator is behind a NAT device, the source port must be 4341 and the destination port is the translated ephemeral port from the source port of ELP-Probe Map-Request.

7. Data-Plane Operation

When an ITR/PITR/RTR select an ELP from one of many ELP-encoded RLOC-records, the downstream RTRs need to use the same path or forwarding loops can occur. The use of the Locator-Status-Bits in the LISP header will serve the encapsulator to instruct the downstream ELP-nodes which RLOC-record to use.



The encapsulator sets the L-bit in the LISP header which allows it to select up to 256 RLOC-records by specifying values 0 to 255 in the low-order 8-bits of the "Instance ID/Locator-Status-Bits" field above.

The following example is constructed to explain the operation:

```
EID-prefix:   10.0.0.0/8
Locator-set:  (A, B, C, ETR-A): priority 1, weight 50  (path 1)
              (A, I, J, ETR-A): priority 1, weight 50  (path 2)
              (S, T, U, ETR-A): priority 2, weight 50  (path 3)
              (X, T, Z, ETR-A): priority 2, weight 50  (path 4)
```

When an ITR receives a packet from a source destined to EID 10.1.1.1, the above mapping will be returned from the mapping system. The ITR will decide which of the two priority 1 RLOC-records to use. When it determines the lower delay path is path 1, it sets the LSB field to 0. When A gets the encapsulated data packet, it is instructed to use path 1 and encapsulate to RTR B (and not RTR I from path 2).

When the ITR determines path 1 and path 2 are both down, it can use the priority 2 RLOC-records. If the path from T to U is down, the ITR would select path 4 and set the LSB field to 3 so T does not try to encapsulate to the down path U but use the path to Z.

Using this head-end mechanism allows one node in the network to switch to reachable, quality, and non-looping paths very quickly without any explicit control-plane signaling to any other nodes.

8. Security Considerations

RLOC-record ELPs stored in the mapping system use the authentication mechanisms described [I-D.ietf-lisp-rfc6833bis] and [I-D.farinacci-lisp-ecdsa-auth]. The ELP-Probe Map-Reply messages can be signed using [I-D.ietf-lisp-sec].

Since the ELP-Probe message is encapsulated as a LISP data packet, telemetry data can be kept private by the use of [RFC8061]. ELP-Probe Map-Reply messages could also be data encapsulated to make use of payload encryption.

9. IANA Considerations

At this time there are no requests for IANA.

10. References

10.1. 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>.

[RFC6830] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "The Locator/ID Separation Protocol (LISP)", RFC 6830, DOI 10.17487/RFC6830, January 2013, <https://www.rfc-editor.org/info/rfc6830>.

[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>.

[RFC8061] Farinacci, D. and B. Weis, "Locator/ID Separation Protocol (LISP) Data-Plane Confidentiality", RFC 8061, DOI 10.17487/RFC8061, February 2017, <https://www.rfc-editor.org/info/rfc8061>.

10.2. Informative References

[I-D.farinacci-lisp-ecdsa-auth] Farinacci, D. and E. Nordmark, "LISP Control-Plane ECDSA Authentication and Authorization", draft-farinacci-lisp-ecdsa-auth-02 (work in progress), April 2018.

[I-D.ietf-lisp-rfc6830bis]
Farinacci, D., Fuller, V., Meyer, D., Lewis, D., and A. Cabellos-Aparicio, "The Locator/ID Separation Protocol (LISP)", draft-ietf-lisp-rfc6830bis-12 (work in progress), March 2018.

[I-D.ietf-lisp-rfc6833bis]
Fuller, V., Farinacci, D., and A. Cabellos-Aparicio, "Locator/ID Separation Protocol (LISP) Control-Plane", draft-ietf-lisp-rfc6833bis-10 (work in progress), March 2018.

[I-D.ietf-lisp-sec]
Maino, F., Ermagan, V., Cabellos-Aparicio, A., and D. Saucez, "LISP-Security (LISP-SEC)", draft-ietf-lisp-sec-15 (work in progress), April 2018.

[I-D.ietf-lisp-te]
Farinacci, D., Kowal, M., and P. Lahiri, "LISP Traffic Engineering Use-Cases", draft-ietf-lisp-te-02 (work in progress), April 2018.

Appendix A. Acknowledgments

The authors would like to thank the LISP working group for their contributions and commentary.

Appendix B. Document Change Log

B.1. Changes to draft-filyurin-lisp-elp-probing-01.txt

- Posted May 2018.
- Update document timer.

B.2. Changes to draft-filyurin-lisp-elp-probing-00.txt

- Initial draft posted November 2017.

Authors' Addresses

Yan Filyurin
Bloomberg LP
731 Lexington Ave
New York, NY
USA

Email: yfilyurin@bloomberg.net

Robert Raszuk
Bloomberg LP
731 Lexington Ave
New York, NY
USA

Email: rraszuk@bloomberg.net

Truman Boyes
MLB
75 9th Ave
New York, NY
USA

Email: truman.boyes@mlb.com

Dino Farinacci
lispers.net
San Jose, California
USA

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