

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: April 18, 2018

M. Boucadair
C. Jacquenet
Orange
October 15, 2017

Improving ITR Resiliency in Locator/ID Separation Protocol (LISP)
Networks
draft-boucadair-lisp-itr-failure-05

Abstract

This document defines an extension to the Locator/ID Separation Protocol (LISP) to minimize LISP service disruption during Ingress Tunnel Routers (ITRs) failure events.

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](#)].

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 April 18, 2018.

Copyright Notice

Copyright (c) 2017 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

Internet-Draft

ITR Resiliency

October 2017

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.	Introduction	2
2.	Procedure	3
3.	Map-Solicit-Request: Message Format & Behavior	5
4.	Map-Solicit-Reply: Message Format & Behavior	7
5.	Security Considerations	10
6.	IANA Considerations	10
7.	Acknowledgments	10
8.	References	10
8.1.	Normative references	10
8.2.	Informative References	11
	Authors' Addresses	11

[1.](#) Introduction

Locator/ID Separation Protocol (LISP, [[RFC6830](#)]) operation relies upon a mapping mechanism that is used by Ingress/Egress Tunnel Routers (xTR) to forward traffic over the LISP network.

A reboot of an ITR may dramatically affect the LISP-based forwarding service for hosts connected to the LISP network. Because of the purge of the mapping cache maintained by the rebooting ITR, the absence of a matching entry for packets to be forwarded over the LISP network will simply cause the dropping of such packets, even though other ITRs of the LISP domain may be "ready-to-serve".

An ITR that loses its local mapping table for some reason is very likely to drop incoming packets whose forwarding decision relies upon the entries of the local mapping table. This type of ITR failure may rarely occur, but when it does, it is likely to provoke severe service degradation.

This document proposes a solution to enhance the robustness of LISP networks during such ITR failure events. This document assumes that several ITRs are available within the LISP network. The solution allows for an automatic discovery of the available ITRs of a given

LISP domain.

The approach exclusively focuses on engineering tweaks that can be implemented within a LISP-enabled network without soliciting the help of the LISP Mapping System. [[I-D.boucadair-lisp-subscribe](#)] is a

companion document that specifies a procedure that is meant to rapidly populate a local mapping cache upon restart or whenever failures affect ITR operation.

[2.](#) Procedure

The overall procedure is as follows:

1. A dedicated IPv4 and/or IPv6 multicast address is reserved for ITR resiliency (called @MCAST in this document). An address can be reserved by an administrator for this purpose.
2. A list of unicast addresses of available ITRs in a given domain is maintained by the requesting ITR (ITR-PEER-LIST).
3. When an ITR loses its mapping table for some reason (power failure, software issue, etc.), but can still forward packets, it checks whether it maintains a list of peer ITRs. If the peer ITR list is empty, it sends a message, denoted Map-Solicit-Request ([Section 3](#)), to @MCAST. If a list is available, the ITR follows Steps (5, 6, and 7).

Note that the same IP address (@MCAST) is used to announce the availability of an ITR within a LISP domain on a regular basis.

4. Once this message is received by another ITR reachable in the LISP domain, it replies with a Map-Solicit-Reply ([Section 4](#)) using its unicast address as the source IP address. The Map-Solicit-Reply includes the following information:
 - * Database Status (including cache status). A status set to 'Null' indicates this ITR does not maintain any cache because, e.g., it is a new ITR, it lost its mappings, etc.
 - * The content of local ITR-PEER-LIST: This is to accelerate the process of discovering other ITRs within a LISP domain without

waiting for responses from other ITRs.

- * Synchronisation reachability information (address, port number, protocol, etc.)
- 5. Bulk mapping requests (e.g., [[I-D.boucadair-lisp-bulk](#)] or [[I-D.boucadair-lisp-multiple-records](#)]) are then sent to peer ITRs to retrieve a copy of their map cache. One or several ITRs can be solicited simultaneously.
- 6. In the meantime, cache synchronisation is in progress, packets that do not match a mapping entry are redirected to another ITR

in the domain that has its database 'ready-to-serve'. These packets are encapsulated in a LISP header using the unicast address discovered in the previous steps.

7. A peer ITR decapsulates the packet, encapsulates it according to the matching mapping entry, and forwards the encapsulated packet towards the next hop. Moreover, it sends an unsolicited Map-Reply to the original ITR so that it can handle locally subsequent packets that belong to this flow.

The 'nonce' of the unsolicited Map-Reply must echo the one included in the encapsulated packet received from the first ITR.

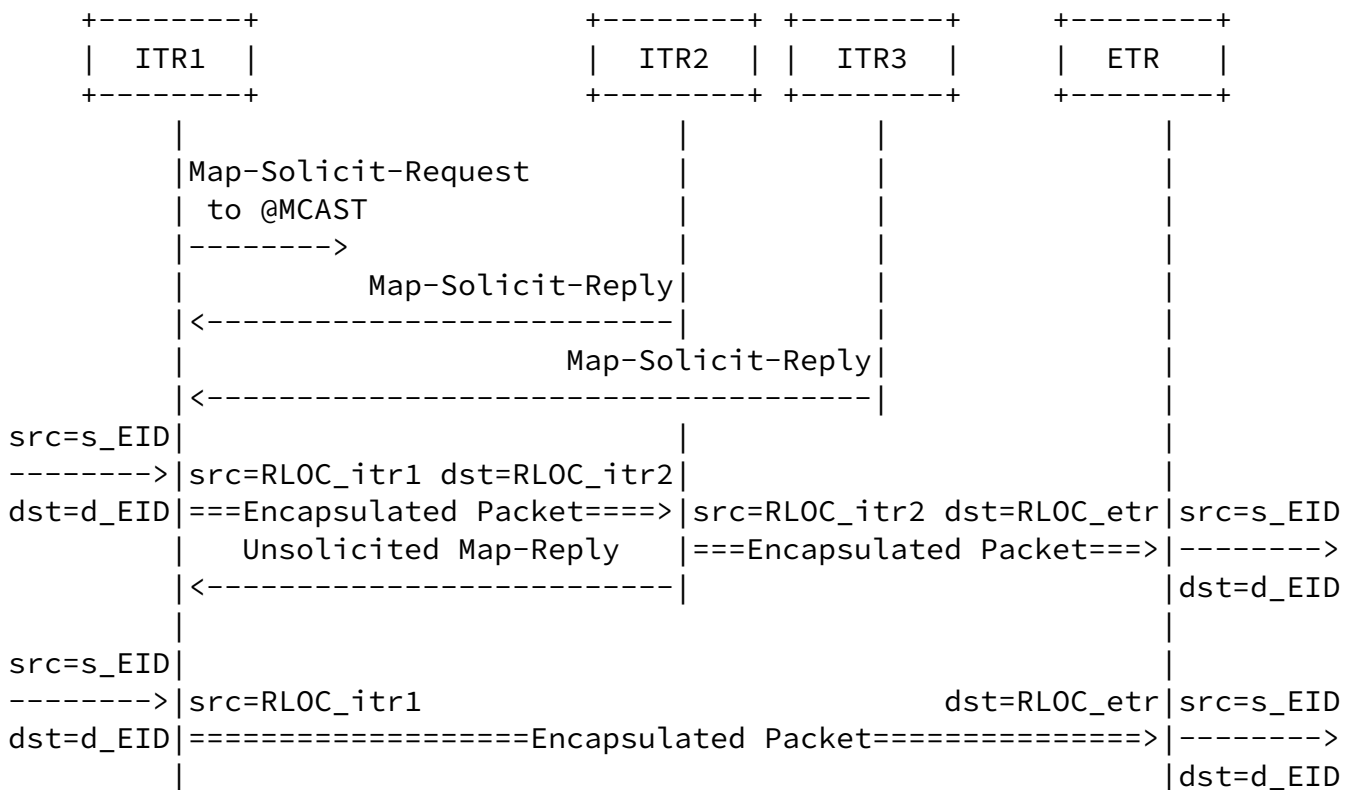
An indication to disable data gleaning may be included by the relay ITR (e.g., using the extension defined in Section 3 of [[I-D.boucadair-lisp-ms-assisted-forwarding](#)]).

Figure 1 illustrates an example of an ITR (ITR1) which encounters a loss of its mapping cache. As a result, it generates a Map-Solicit-Request that it sends to the multicast address @MCAST. Upon receipt of that request by ITR2 and ITR3, they each reply with a Map-Solicit-Reply message. The first reply is used by ITR1 to decide to which peer ITR it will redirect packets during the failure event (ITR2). These packets are encapsulated with a LISP header and forwarded to ITR2. Once received by ITR2, these packets are forwarded to their ultimate ETR. In the meantime, ITR2 generates an unsolicited Map-Reply to inform ITR1 with the mapping entries related to the destination EID. Subsequent packets that belong to this flow are therefore handled locally by ITR1 without soliciting ITR2.

Internet-Draft

ITR Resiliency

October 2017



• • • •

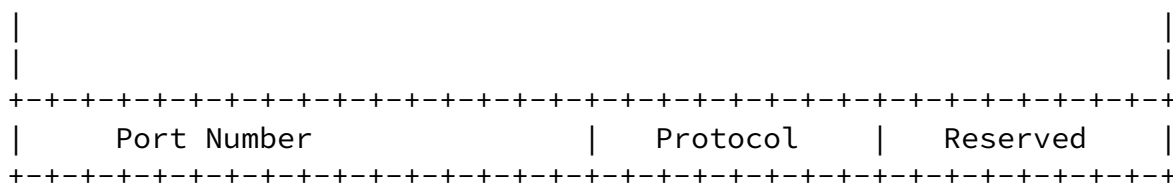


Figure 2: Map-Solicit-Request Message Format

The description of the fields is as follows:

- o Type MUST be set to 15 [[RFC8113](#)].
- o sub-type: MUST be set to 1026.
- o R: MUST be set to 0 for Map-Solicit-Request messages.
- o S: when set, this flag indicates that the originating ITR supports a mechanism for state synchronisation of the mapping cache between ITRs. When this flag is set, the message MUST carry the port number, protocol, and IP Address used for synchronisation purposes. This specification allows to indicate a distinct IP address for state synchronisation purposes.
- o D: This flag indicates the status of the mapping cache table. It is RECOMMENDED to set this flag to 1 when the ITR is up and running for at least one hour and has a non-empty mapping cache. An ITR that lost its state MUST set this flag to 0.
- o Nonce, Key ID, Authentication Data Length, and Authentication Data are similar to those of a LISP Map-Register message ([\[RFC6830\]](#)).

- o IP Address: If S-bit is set, this field indicates the IP address used to receive state synchronisation messages. If S-bit is unset, this field MUST be set to zero at the originating ITR and MUST be ignored at receipt. The length of this field is 128 bits. IPv4 addresses are encoded as IPv4-mapped IPv6 addresses [[RFC4291](#)] (::ffff:0:0/96).
- o Port Number: If the S-bit is set, this field indicates the port

number used to receive state synchronisation messages. If unset, this field MUST be set to zero at the originating ITR and MUST be ignored at receipt.

- o Protocol: If the S-bit is set, this field indicates the protocol used to transport state synchronisation messages. If unset, this field MUST be set to zero at the originating ITR and MUST be ignored upon receipt.

An ITR that issues this message MUST use one of its unicast IP addresses as the source address. The destination IP address MUST be set to the @MCAST multicast address introduced in [Section 2](#). An ITR that loses its cache MUST issue this message with a D-bit set to 0.

[4](#). Map-Solicit-Reply: Message Format & Behavior

All ITRs of a LISP domain MUST subscribe to the multicast group defined by the aforementioned @MCAST multicast address.

Upon receipt of the Map-Solicit-Request message by an ITR within the domain, it replies (unicast) with a Map-Solicit-Reply. It is the responsibility of the first ITR to initiate a state synchronisation with that peer if the D-bit and S-bit are unset and if it supports the synchronisation protocol indicated in the Map-Solicit-Reply.

ITRs of a LISP domain MUST send Map-Solicit-Reply in a regular interval (that is configured by an administrator) or upon major change in the ITR stats (e.g., loss of the mapping cache, change of the IP address). This message MUST use one of the ITR unicast IP addresses as the source address while the destination IP address MUST be set to the @MCAST.

The format of the Map-Solicit-Reply message is shown in Figure 3.

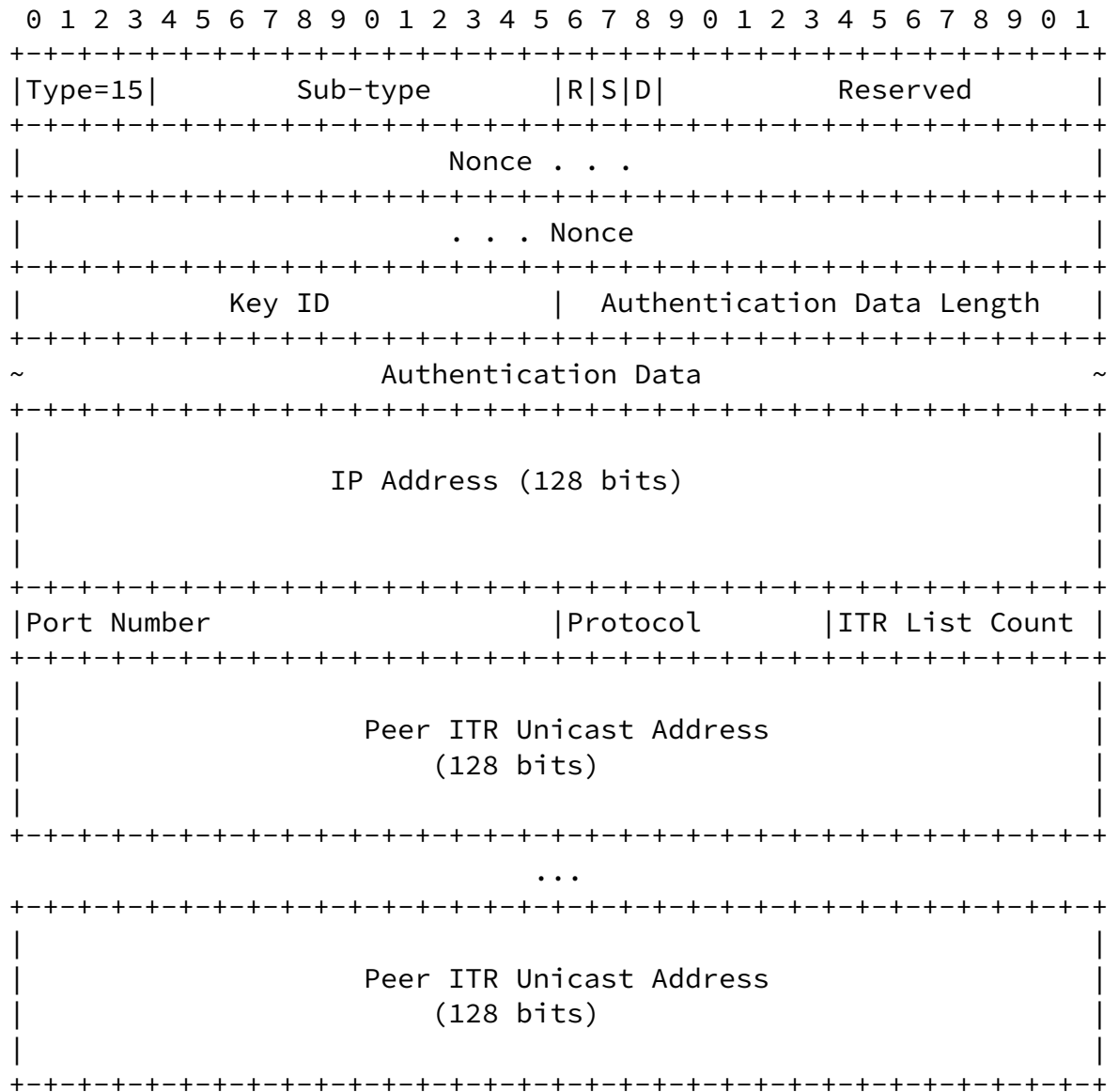


Figure 3: Map-Solicit-Reply Message Format

The description of the fields is as follows:

- o Type MUST be set to 15 [[RFC8113](#)].
- o sub-type: MUST be set to 1026.
- o R: MUST be set to 1.
- o S: when set, this flag indicates that the originating ITR supports a mechanism for state synchronisation of the mapping caches between ITRs. When set, the message MUST carry the port number, protocol, and IP Address used for synchronisation purposes. This

specification allows to indicate a distinct IP address for state synchronisation purposes.

- o D: This flag indicates the status of the mapping cache table. It is RECOMMENDED to set this flag when the ITR is up and running for at least one hour and has a non-empty mapping cache.
- o Nonce: The 'Nonce' field of multicast Map-Solicit-Reply MUST be set to 0 while it MUST echo the one included in a Map-Solicit-Request when replying to a multicast Map-Solicit-Request.
- o Key ID, Authentication Data Length, and Authentication Data are similar to those of a LISP Map-Register message ([RFC6830]).
- o IP Address: If the S-bit is set, this field indicates the IP address used to receive state synchronisation messages. If unset, this field MUST be set to zero at the originating ITR and MUST be ignored upon receipt. The length of this field is 128 bits. IPv4 addresses are encoded as IPv4-mapped IPv6 addresses [RFC4291] (::ffff:0:0/96).
- o Port Number: If the S-bit is set, this field indicates the port number used to receive state synchronisation messages. If unset, this field MUST be set to zero at the originating ITR and MUST be ignored upon receipt.
- o Protocol: If the S-bit is set, this field indicates the protocol used to transport state synchronisation messages. If unset, this field MUST be set to zero at the originating ITR and MUST be ignored upon receipt.
- o ITR List Count: This field indicates whether peer ITR addresses are also included. When this field is set to 0, it indicates that no peers other than the solicited peer ITR are known to the originating ITR.
- o Peer ITR Unicast Address: one or multiple IP addresses that belong to other ITRs in the domain as known to the originating ITR. The length of each "Peer ITR Unicast Address" is 128 bits. IPv4 addresses are encoded as IPv4-mapped IPv6 addresses (::ffff:0:0/96).

A Map-Solicit-Reply can be generated by an ITR to advertise its availability to the other ITRs of the LISP domain, as per normal LISP operation.

When an ITR receives a LISP-encapsulated packet from an ITR that is present in its list of peer ITRs, it may generate an unsolicited Map-

Reply that conveys the mapping entry that was used to process the encapsulated packet.

Upon failure or reboot that lead to lose the contents of its mapping cache, an ITR uses the list of peers ITRs it discovered by means of the Map-Solicit-Request message sent to @MCAST to redirect packets that do not match any entry of its local cache (which is likely to be empty).

In order to minimize the risk of overloading some ITRs, a mechanism to distribute the load among all the peer ITRs or part of them is deemed useful. Of course, other traffic load distribution policies may be enforced. The exact set of policies to be enforced are implementation- and deployment-specific.

[5.](#) Security Considerations

LISP security considerations are discussed in [[RFC6830](#)].

This document specifies a mechanism that enhances the serviceability of LISP networks by redirecting traffic that do not match a local mapping entry to other ITRs of the domain. These ITRs are assumed to belong to the same administrative domain. Means to ensure that only trusted ITRs are maintained in a peer list MUST be enabled.

[6.](#) IANA Considerations

IANA has assigned the following code from the LISP Shared Extension Message Type Sub-types ([[RFC8113](#)]):

Message	Sub-type	Reference
=====	=====	=====
Map-Solicit-Request/Map-Solicit-Reply	1026	[This document]

[7.](#) Acknowledgments

This work is partly funded by ANR LISP-Lab project #ANR-13-INFR-009-X.

Many thanks to Chi Dung Phung for the review.

8. References

8.1. Normative references

Boucadair & Jacquenet Expires April 18, 2018 [Page 10]

Internet-Draft ITR Resiliency October 2017

- [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>>.
- [RFC4291] Hinden, R. and S. Deering, "IP Version 6 Addressing Architecture", [RFC 4291](#), DOI 10.17487/RFC4291, February 2006, <<https://www.rfc-editor.org/info/rfc4291>>.
- [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>>.
- [RFC8113] Boucadair, M. and C. Jacquenet, "Locator/ID Separation Protocol (LISP): Shared Extension Message & IANA Registry for Packet Type Allocations", [RFC 8113](#), DOI 10.17487/RFC8113, March 2017, <<https://www.rfc-editor.org/info/rfc8113>>.

8.2. Informative References

- [I-D.boucadair-lisp-bulk] Boucadair, M. and C. Jacquenet, "LISP Mapping Bulk Retrieval", [draft-boucadair-lisp-bulk-05](#) (work in progress), April 2017.
- [I-D.boucadair-lisp-ms-assisted-forwarding] Boucadair, M. and C. Jacquenet, "Mapping System-Assisted Forwarding for Inter-Domain LISP Deployments", [draft-boucadair-lisp-ms-assisted-forwarding-00](#) (work in progress), September 2015.

[I-D.boucadair-lisp-multiple-records]

Boucadair, M. and C. Jacquenet, "Retrieving Multiple LISP Records", [draft-boucadair-lisp-multiple-records-00](#) (work in progress), October 2017.

[I-D.boucadair-lisp-subscribe]

Boucadair, M. and C. Jacquenet, "LISP Subscription", [draft-boucadair-lisp-subscribe-05](#) (work in progress), April 2017.

Authors' Addresses

Boucadair & Jacquenet Expires April 18, 2018

[Page 11]

Internet-Draft

ITR Resiliency

October 2017

Mohamed Boucadair
Orange
Rennes 35000
France

EMail: mohamed.boucadair@orange.com

Christian Jacquenet
Orange
Rennes 35000
France

EMail: christian.jacquenet@orange.com

