

Network Working Group
Internet-Draft
Intended status: Experimental
Expires: 5 April 2026

B. Pitta
M. Portoles
Cisco Systems
D. Lewis
ICANN
I. Kouvelas
Arista Networks Inc.
C. Cassar
Rivian Automotive
2 October 2025

LISP Map Server Reliable Transport
draft-ietf-lisp-map-server-reliable-transport-07

Abstract

The communication between LISP ETRs and Map-Servers is based on unreliable UDP message exchange coupled with periodic message transmission in order to maintain soft state. The drawback of periodic messaging is the constant load imposed on both the ETR and the Map-Server. New LISP use cases increase the amount of state that needs to be communicated and challenge the scalability of the system when using the UDP exchange. This document introduces the use of a reliable transport for ETR to Map-Server communications in order to eliminate the periodic messaging overhead, while providing reliability, flow-control and endpoint liveness detection.

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 5 April 2026.

Copyright Notice

Copyright (c) 2025 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	3
2. Requirements Notation	4
3. Definition of terms	4
4. Message Format	4
5. Session Establishment	6
5.1. Signaling to Trigger Session Establishment	7
5.2. Transport Protocol and Port Usage	7
6. Error Notifications	8
7. EID Prefix Registration	9
7.1. Reliable Registration Messages	10
7.1.1. Reliable Mapping Registration Message	10
7.1.2. Registration Acknowledgement Message	11
7.1.3. Registration Rejection Message	12
7.1.4. Registration Refresh Request Message	13
7.1.5. Reliable Mapping Notification Message	15
7.1.6. Note on message format reuse	16
7.2. ETR Behavior	17
7.2.1. Transitions between periodic and reliable transport	18
7.2.2. Transitions between reliable transport states	21
7.3. Map-Server Behavior	22
7.3.1. Map-Server Reject and Withdraw use cases	23
8. Security Considerations	24
9. IANA Considerations	24
9.1. LISP Reliable Transport Message Types	24
9.2. Error Notifications: Error Codes	25
9.3. Registration Rejection Message: Reason Codes	25
9.4. Transport Protocol Port Numbers	26
10. Sample Reliable Transport Experiences	26
10.1. Support of EID mobility at scale	26
10.2. Distributed Overlapping prefix support at scale	27
10.3. Stable interconnect with external networks	27

11. Contributors	28
12. Acknowledgments	28
13. References	28
13.1. Normative References	28
13.2. Informative References	29
Authors' Addresses	29

1. Introduction

The communication channel between LISP ETRs and Map-Servers is based on unreliable UDP message exchange [RFC9301]. Where required, reliability is pursued through periodic retransmissions that maintain soft state on the peer. Map-Register messages are retransmitted every minute by an ETR and the Map-Server times out its state if the state is not refreshed for three successive periods. When registering multiple EID-Prefixes, the ETR includes multiple mapping records in the Map-Register message. Packet size limitations provide an upper bound to the number of mapping records that can be placed in each Map-Register message. When the ETR has more EID-Prefixes to register than can be packed in a single Map-Register message, the mapping records for the EID-Prefixes are split across multiple Map-Register messages.

The drawback of the periodic registration is the constant load that it introduces on both the ETR and the Map-Server. The ETR uses resources to periodically build and transmit the Map-Register messages, and to process the resulting Map-Notify messages issued by the Map-Server. The Map-Server uses resources to process the received Map-Register messages, update the corresponding registration state, and build and transmit the matching Map-Notify messages. When the number of EID-Prefixes to be registered by an ETR is small, the resulting load imposed by periodic registrations may not be significant. The ETR will only transmit a single Map-Register message each period that contains a small number of mapping records.

In some LISP deployments, a large set of EID-Prefixes must be registered by each ETR (e.g. endpoint mobility, database redistribution). Use cases with a large set of EID-Prefixes behind ETRs will result in much higher load in the System. An example is LISP mobility deployments where EID-Prefixes are limited to host entries. ETRs may have thousands of EIDs to register resulting in hundreds of Map-Register and Map-Notify messages per each registration period.

A transport is required for the ETR to Map-Server communication that provides reliability, flow-control and endpoint liveness notifications. This document describes the use of TCP as a LISP reliable transport. The initial application for the LISP reliable

transport session is the support of scalable EID registration. The reliable session mechanism is defined to be extensible so that it can support additional LISP communication requirements as they arise using a single reliable transport session between an ETR and a Map-Server.

The use of a reliable transport session for EID registration is optional and the existing UDP based mechanism remains mandatory as for [RFC9301]. When a Map-Server accepts reliable transport sessions it MUST be able to handle Map-Register messages from multiple ETRs using both UDP as well as reliable transport at the same time.

2. Requirements Notation

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT", "SHOULD", "SHOULD NOT", "RECOMMENDED", "NOT RECOMMENDED", "MAY", and "OPTIONAL" in this document are to be interpreted as described in BCP 14 [RFC2119] [RFC8174] when, and only when, they appear in all capitals, as shown here.

3. Definition of terms

The document uses the terms defined in Section 3 of documents [RFC9300] and [RFC9301].

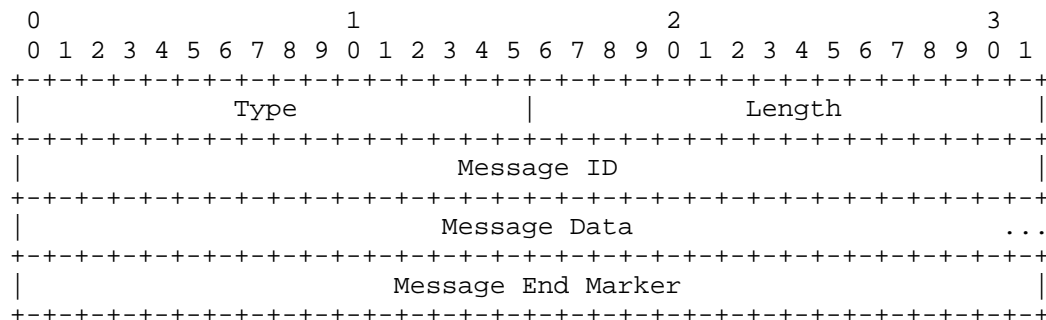
Additionally, this document introduces the following terms:

LISP Application: This is the term used in this document to identify specific use cases that use a LISP transport to exchange messages between Map-Servers and xTRs. .

4. Message Format

A single LISP reliable transport session may carry information for multiple LISP Applications. One such application is the registration of EID to RLOC mappings that operates over a session between an ETR and a Map-Server. Communication over a session is based on the exchange of messages. This document defines a base set of messages to support session establishment and management. It also defines the messages for the EID to RLOC mapping registration application.

All these messages are based on a TLV format in order to accommodate future extensions.



Reliable transport message format

- * Type: 16 bit type field identifying the message type.
- * Length: 16 bit field that provides the total size of the message in octets including the length, type and end marker fields. The length allows the receiver to locate the next message in the TCP stream. The minimum value of the length field is 8. The purpose of the End Marker is to detect formatting errors.
- * ID: A 32-bit value that identifies the message. May be used by the receiver to identify the message in replies or notification messages.
- * Data: Type specific message contents.
- * End Marker: A 32-bit message end marker that must be set to 0x9FACADE9. The End Marker is used by the receiver to validate that it has correctly parsed or skipped a message and provides a method to detect formatting errors. The combination of the length field and the marker when parsing enables this detection. Note that message data may also contain this marker, and the marker itself is not sufficient for parsing the message.

The base message format does not indicate how the peer should deal with the message in cases where the message type is not supported/understood. This is best dealt with by the application. For example, in case an error notification is returned, or an expected acknowledgement message is not received, the application might choose various courses of action; from simply logging that the feature is not supported, all the way to tearing the session down and triggering a session restart.

5. Session Establishment

The Map-Server and ETR MUST communicate via unreliable UDP messages until a reliable session between the two can be successfully established. This initial UDP exchange is used for both ETR authentication as well as to express intent and negotiate the use of a reliable transport between ETR and Map-Server.

The ETR indicates its intent to use a reliable transport setting a Reliable Transport bit in the UDP Map-Register message. As an acknowledgement, when the Map-Server is ready to accept the establishment of a reliable transport session it also sets the corresponding Reliable Transport bit in the Map-Notify message. Reliable Transport bits are specified below in Section 5.1.

The Map-Server authenticates the ETR with the authentication data contained in the first UDP Map-Register message it receives from the ETR. Once the ETR is authenticated, the Map-Server performs a passive open by listening on TCP port 4342, and does not qualify the remote port. As a security measure, the Map-Server does not create any connection unless a UDP authentication, with a reliable transport bit set, completes first. After that, the Map-Server accepts connections only from those ETRs that have been authenticated via UDP Map-Register messages.

The ETR assumes the role of active connection opener of the session establishment by connecting to the Map-Server once it has received a UDP Map-Notify message with the corresponding Reliable Transport bit set. ETR MUST assume the active connection opener roles and it is always the one attempting the connection.

When a reliable session goes down, UDP authentication must take place before a new reliable session is established. The Map-Server MUST NOT accept a connection from the ETR until a UDP Map-Register with the corresponding Reliable Transport bit set has been received. Similarly, the ETR MUST NOT attempt to establish a session with the Map-Server until a UDP Map-Notify message has been received with the Reliable Transport bit set. Note also that while establishing a reliable transport session is not acknowledged, the ETR and Map-Server SHOULD continue using UDP for transport, as specified in [RFC9301]

A single reliable transport session is established between the Map-Server and the ETR to cover all communication needs. For example, an ETR that has EID prefix registrations for multiple EID instances and EID address families SHOULD only establish a single session with the Map-Server.

Note that the use of TCP in this section and the rest of the document is intentional. The architecture is extensible to support other protocols, as documented in Section 5.2, but this document only specifies the use of TCP.

5.1. Signaling to Trigger Session Establishment

Following the procedures described in this document, the Map-Register and Map-Notify headers are extended with new flags, the Reliable Transport bits, that are used to advertise intent to establish a reliable transport session with a specific protocol (ETR), as well as the capability to accept reliable transport sessions with the same specific protocol (Map-Server).

Following this specification, the Map-Register header, as defined in [RFC9301], is extended with the TCP Reliable Transport bit.

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|Type=3 |P|S|I|           Reserved           |r|E|T|a|R|M| Record Count |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

r: This is the TCP Reliable Transport bit in the Map-Register header. It should be set to 1 when the ETR that sends a Map-Register intends to establish a reliable transport session, using TCP, with the Map-Server that the Map-Register is being sent to.

The Map-Notify header, as defined in [RFC9301], is extended with the TCP Reliable Transport bit.

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|Type=4|           Reserved           |r| Record Count |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

r: This is the TCP Reliable Transport bit in the Map-Notify header. It should be set to 1 when the Map-Server has validated an initial UDP registration and acknowledges the possibility to use TCP to support a reliable transport session with the ETR.

5.2. Transport Protocol and Port Usage

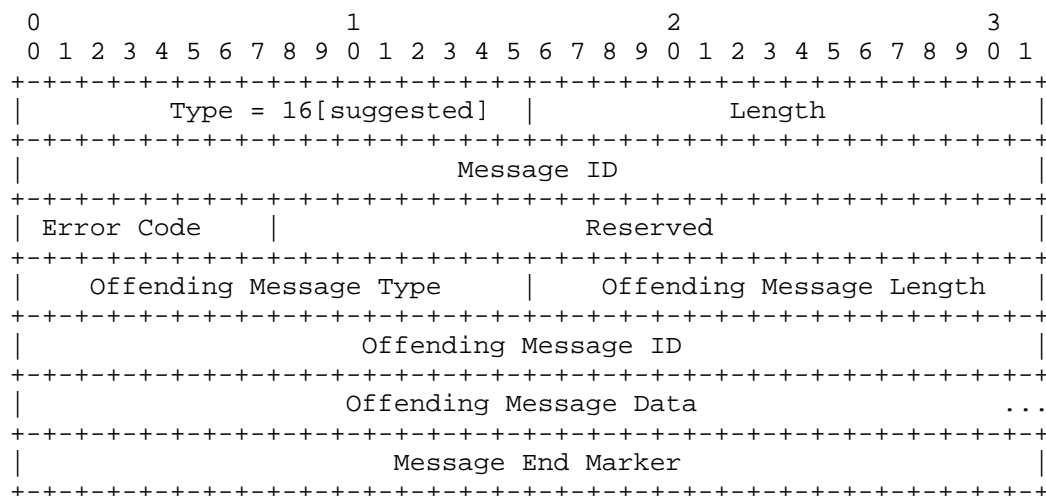
This document establishes the following transport protocols that can be used to sustain reliable transport sessions between ETRs and the Map-Servers:

TCP [RFC0793]: The port 4342 is used for this purpose (see Section 9.4). TCP sessions are long-lived and maintained unless one of the endpoints closes or resets the session, or the communication times out and needs to be re-established following the procedures described above in this section.

Alternative protocols may be added in the future. New protocols proposed going forward SHOULD reserve a new bit in the Map-Register and Map-Notify headers. The use of individualized bits for each protocol allows the ETR and Map-Server to advertise and negotiate the protocols supported.

6. Error Notifications

The error notification message is used to communicate base reliable transport session communication errors. LISP applications making use of the reliable transport session and having to communicate application specific errors must define their own messages to do so. An error notification is issued when the receiver of a message does not recognize the message type or cannot parse the message contents. The notification includes the offending message type and ID and as much of the offending message data as the notification sender wishes to.



Error Notification message format

- * Error Code: An 8 bit field identifying the type of error that occurred. Defined errors are:

- Unrecognized message type.
- Message format error.
- * Reserved: Set to zero by the sender and ignored by the receiver.
- * Offending Message Type: 16 bit type field identifying the message type of the offending message that triggered this error notification. This is copied from the Type field of the offending message.
- * Offending Message Length: 16 bit field that provides the total size of the offending message in octets. This is copied from the Length field of the offending message.
- * Offending Message ID: A 32-bit field that is set to the Message ID field of the offending message.
- * Offending Message Data: The Data from the offending message that triggered this error notification. The sender of the notification may include as much of the original data as is deemed necessary. The length of the Offending Message Data field is not provided by the Offending Message Length field and is determined by subtracting the size of the other fields in the message from the Length field. It is valid to not include any of the offending message data when sending an error notification.
- * End Marker: A 32-bit message end marker that must be set to 0x9FACADE9. The End Marker is used by the receiver to validate that it has correctly parsed or skipped a message and provides a method to detect formatting errors. Note that message data may also contain this marker, and that the marker itself is not sufficient for parsing the message.

An error notification cannot be the offending message in another error notification and MUST NOT trigger such a message.

7. EID Prefix Registration

EID prefix registration uses the reliable transport session between an ETR and a Map-Server to communicate the ETR local EID database EID to RLOC mappings to the Map-Server. In contrast to the UDP based periodic registration refresh, mapping information over the reliable transport session is only sent when there is new information available for the Map-Server. The Map-Server does not maintain a timer to expire registrations communicated over the reliable transport session. Instead the ETR MUST send an explicit de-registration (a registration carrying a zero TTL) when it intends to

delete the registration state maintained by the Map-Server.

The key used to identify registration mapping records in the ETR to Map-Server communication is the EID prefix. The prefix may be specified using an LCAF encoding that includes an EID instance ID.

When the reliable transport session goes down, registration mappings learned by the Map-Server are treated as periodic UDP registrations and a timer is used to expire them after 3 minutes. During this period UDP based registrations or the re-establishment of the reliable transport session and subsequent communication of a new mapping can update the EID prefix mapping state.

7.1. Reliable Registration Messages

This section defines the LISP reliable transport session message used to communicate local EID database registrations between the ETR and the Map-Server.

7.1.1. Reliable Mapping Registration Message

The reliable transport registration message is used to communicate EID to RLOC mapping registrations from the ETR to the Map-Server. The "Message Data" field uses the same format as the UDP Map-Register message but without the IP and UDP headers. This message format reuse allows the ETR to take advantage of all registration options and flags specified as part of [RFC9301] without redefinition.

A reliable registration message **MUST** contain a single mapping-record. The Record Count field is always set to 1. The Map-Server **MUST** discard any reliable registration message that contains more than one mapping record. This approach simplifies individual EID record Acknowledgment to sustain the state machine described below in this document

An ETR **MUST NOT** set the T bit in a Map-Register message sent over a reliable transport session. If the T-bit is set, the Map-Server **MAY** ignore the Map-Register message.

The reliable transport session is authenticated by means of the session establishment procedure. Thus, although the Map-Register **MUST** carry the authentication data, it is up to the Map-Server to determine if each individual reliable registration message should be authenticated.

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|           Type = 17[suggested] |           Length           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Message ID                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Map-Register message            |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               ...                               |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Map-Register message            |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Message End Marker              |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Registration message format

7.1.2. Registration Acknowledgement Message

The Acknowledgement message is sent from the Map-Server to the ETR to confirm successful registration of an EID prefix previously communicated through a Registration message, through a reliable registration message.

Note that the objective of this message is to ensure proper state synchronization between Map-Server and ETR. In the cases when the Map-Server view of the mapping is needed, the Reliable Mapping Notification Message is used, as described below in section Section 7.1.5

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|           Type = 18[suggested] |           Length           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Message ID                     |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Prefix-Length |           EID-Prefix-AFI           | EID-Prefix ... |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                               Message End Marker              |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Registration Acknowledgement message format

- * Prefix-Length: Mask length for the EID prefix.
- * EID-Prefix AFI: Address family identifier for the EID prefix in the following field.

* EID-Prefix: The EID prefix from the received Registration.

7.1.3. Registration Rejection Message

The Registration Rejection Message is sent by the Map-Server to the ETR to indicate that the registration of a specific EID prefix is being rejected or withdrawn.

A rejection refers to a registration that the Map-Server does not accept when the registration is received, without recording it. A withdrawal refers to a registration that was previously accepted, but for which the conditions may have changed and the Map-Server cannot longer process it.

Note that rejection or withdrawal of registration does not confer authoritative status to the Map-Server over the Mappings, but they are used to reflect status of the Map-Server to allows the ETR to keep actual registration status of each of the mappings. Section Section 7.3.1 exemplifies conditions to reject or withdraw a mapping from the Map-Server.

```

      0               1               2               3
    0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Type = 19[suggested] |                                     Length |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                     Message ID |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Reason | Reserved | Prefix-Length |
+-----+-----+-----+-----+-----+-----+-----+-----+
| EID-Prefix-AFI | EID-Prefix | ... |
+-----+-----+-----+-----+-----+-----+-----+-----+
| Message End Marker |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Registration Rejection message format

* Reason: Code identifying the reason for which the Map-Server rejected or withdrew the registration.

- 1 - Not a valid site EID prefix.
- 2 - Authentication failure.
- 3 - Locator set not allowed.
- 4 - Used to cover reason that's not defined.

- * **Reserved:** This field is reserved for future use. Set to zero by the sender and ignored by the receiver.
- * **Prefix-Length:** Mask length for the EID prefix.
- * **EID-Prefix-AFI:** Address family identifier for the EID prefix in the following field.
- * **EID-Prefix:** The EID prefix being rejected or withdrawn.

7.1.4. Registration Refresh Request Message

The Registration Refresh Request message is sent by the Map-Server to the ETR to request the (re-)transmission of EID prefix database mapping Registration messages.

There are multiple cases when this message is used. An example is when the reliable session is initially established. The Map-Server uses this message to request the ETR to send all registrations and achieve proper state synchronization with that ETR. Other examples are when there are configuration changes (e.g. a new instance-Id supported) on the Map-Server, that affect EID registrations that may have been rejected in the past.

```

0           1           2           3
0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1
+-----+-----+-----+-----+-----+-----+-----+-----+
|           Type = 20[suggested]           |           Length           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                           Message ID                                           |
+-----+-----+-----+-----+-----+-----+-----+-----+
|      Scope      |R|           Reserved           | Prefix-Length |
+-----+-----+-----+-----+-----+-----+-----+-----+
|           EID-Prefix-AFI           |           EID-Prefix           ...
+-----+-----+-----+-----+-----+-----+-----+-----+
|                                           Message End Marker                                           |
+-----+-----+-----+-----+-----+-----+-----+-----+

```

Registration Refresh Request message format

- * **Scope:** Determines the set of registrations being refreshed.

- 0 - All prefixes under all address families under all EID instances are being refreshed. When using this scope the Prefix-Length, EID-Prefix-AFI, and EID-Prefix fields MUST be omitted. That is, the Message End Marker follows immediately after the Reserved field. The total length of the message MUST be 15 bytes.
- 1 - All prefixes under all address families under a single EID instance are being refreshed. The Prefix-Length MUST be set to zero, EID-Prefix-AFI MUST be set to LCAF type, the EID-Prefix encodes the LCAF Instance ID, the LCAF address AFI MUST be set to UNSPECIFIED. The total length of the message MUST be 30 bytes.
- 2 - All prefixes under a single address family under a single EID instance are being refreshed. The Prefix-Length MUST be set to zero, the EID-Prefix-AFI MUST be set to LCAF type and the EID-Prefix MUST encode the Instance ID. The LCAF address AFI MUST specify the address family to refresh, the actual address SHOULD be set to zero.
- 3 - All prefixes covered by a specific EID prefix in a single EID instance is being refreshed. The Prefix-Length, EID-Prefix-AFI and EID prefix MUST be encoded accordingly.
- 4 - A specific EID prefix in a single EID instance is being refreshed. The Prefix-Length, EID-Prefix-AFI and EID prefix MUST be encoded accordingly.

The Map-Server has the flexibility to control the granularity of the refresh by issuing refresh with different scopes. It can send a single refresh with a coarse scope or send individual refreshes with narrower scope. The ETR MUST be able to process all scopes to ensure the Map-Server registration states are synchronized with the ETR.

- * R: Request the ETR to only refresh registrations that have been previously rejected by the Map-Server. If the R bit is set then the scope cannot have a value of 3 and the EID-Prefix and Prefix-Length fields must be omitted.
- * Reserved: This field is reserved for future use. Set to zero by the sender and ignored by the receiver.
- * Prefix-Length: Mask length for the EID prefix. Refer to scope for more details.

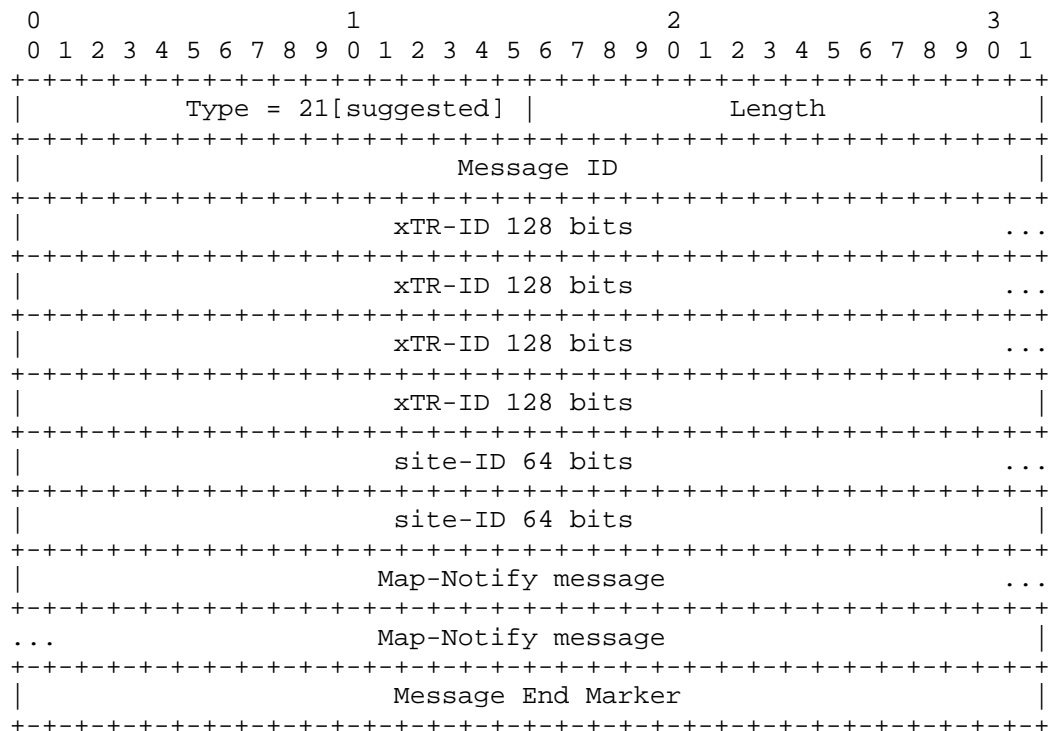
- * EID-Prefix-AFI: Address family identifier for the EID prefix in the following field. Refer to scope for more details.
- * EID-Prefix: The EID prefix being refreshed. Refer to scope for more details.

7.1.5. Reliable Mapping Notification Message

Reliable Mapping Notification messages communicate the Map-Server view of the mapping for an EID prefix and no longer serve as a registration acknowledgement.

The "Message Data" field includes the UDP Map-Notify message as illustrated in the figure but without the IP and UDP headers. This approach allows reusing semantics of unsolicited Map-Notifies as described in [RFC9437]. Note that, as in the Map-Register case, the Record Count is always set to 1.

The reliable transport session is authenticated by means of the session establishment procedure. Thus, although the Map-Notify MUST carry the authentication data, it is up to the ETR to determine if each individual reliable Mapping Notification needs to be validated.



Reliable Mapping Notification message format

- * xTR-ID: xTR-ID taken from the last valid registration the Map-Server received for the EID-prefix conveyed in the mapping record.
- * site-ID: site-ID taken from the last valid registration the Map-Server received for the EID-prefix conveyed in the mapping record.
- * Mapping-Notify message: Map-Notify message formatted following [RFC9301] specifications.

7.1.6. Note on message format reuse

The messages that this section defines are used to support EID prefix registrations over a reliable transport. The use of the messages is contextualized in the following sections with the introduction of the state machine that each ETR maintains.

Since the purpose and objective of some of the messages does not change from what is specified in [RFC9301], this document proposes embedding the same packet format as used in the UDP based specification. In particular the two messages reusing packet formats are:

- * **Reliable Mapping Registration:** This message embeds the same message as the Map-Register and it is used with the same objective. A Map-Server MUST be able to distinguish between Map-Register received from UDP from the ones received over a reliable transport. All the fields in the Map-Register are reused with the exception of the T-bit (MUST NOT be set) and the Record count (MUST be 1).
- * **Reliable Mapping Notification:** this message embeds the same message as the Map-Notify, and it is used in a reduced number of cases. In particular Notifications are not required as confirmations of reception. This role is taken by the Acceptance/Reject messages described above. The Notification is used by the Map-Server to share relevant information of a Mapping (e.g. EID mobility or RLOC merging). The Record Count MUST be set to 1.
- * All other messages defined here are new and do not embed any other message defined in [RFC9301].

7.2. ETR Behavior

The ETR operates the following state machine for each EID prefix and Map-Server pair. This defines the reliable transport EID prefix registration behavior.

There are five states:

- * **No state:** The local EID database prefix does not exist.
- * **Periodic:** The local EID database prefix is being periodically registered through UDP Map-Register messages as specified in [RFC9301].
- * **Stable:** From the ETR's perspective, no registrations are due to be sent to the peer. The session to the peer is up, and the peer has either acknowledged the registration, or is expected to request a refresh in the future.
- * **AckWait:** A Registration message for the prefix has been transmitted to the Map-Server and the ETR is waiting for either a Registration Acknowledge or Registration Rejected reply from the Map-Server.

- * **Reject:** The reliable transport registration for the local EID database prefix was rejected by the Map-Server. From the ETR's perspective, no registration is due to the peer AND the peer is known to have rejected the registration.

The following events drive the state transitions:

- * **DB creation:** The local EID database entry for the EID prefix is created.
- * **DB deletion:** The local EID database entry for the EID prefix is deleted.
- * **DB change:** The mapping contents or authentication information for the local EID database entry changes.
- * **Session up:** The reliable transport session to the Map-Server is established.
- * **Session down:** The reliable transport session the Map-Server goes down.
- * **Recv Refresh:** A Registration refresh message is received from the Map-Server.
- * **Recv ACK:** A Registration Acknowledge message is received from the Map-Server.
- * **Recv Rejected:** A Registration Rejected message is received from the Map-Server.
- * **Periodic timer:** The timer that drives generation of periodic UDP Map-Register messages fires.

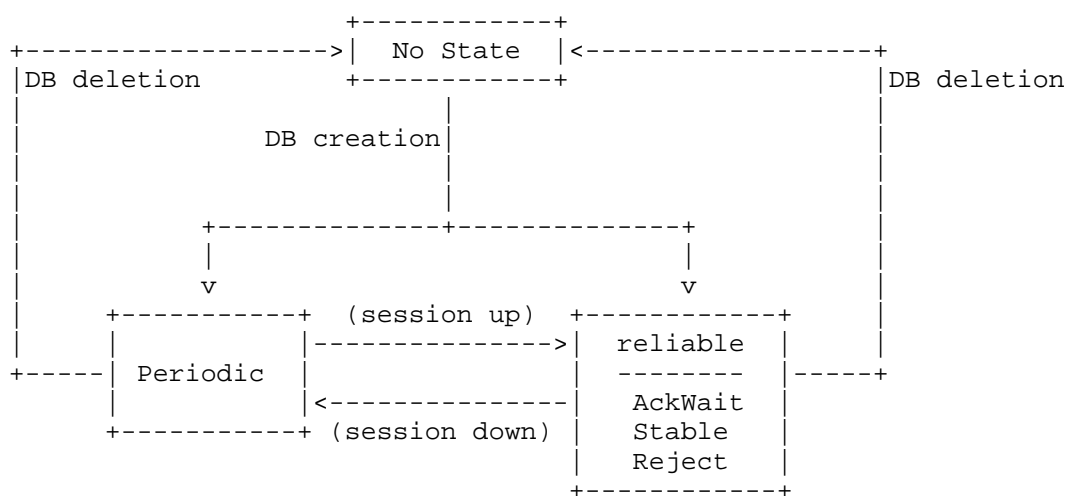
To simplify the description, the rest of the section splits the state machine description into two separate parts: The first one describes the switching between the periodic (UDP) and the reliable transport states. The second part describes transitions within reliable registration states, when the reliable session is already established.

7.2.1. Transitions between periodic and reliable transport

Transitions between periodic and reliable transports states are driven by the status of the session. EID prefixes transition between periodic and reliable transport states following session events (session going up or down).

Additionally, when a database is created, the status of the session matters to determine the initial state of the EID prefix in the state machine. If the session is down (not established yet) the EID prefix starts in periodic state, following [RFC9301]. If the session is already up the EID prefix directly starts following reliable transport states.

All this is illustrated in the following figure. Note how DB deletion events remove move the EID prefix to "No State" regardless of session state.



State machine in periodic/reliable transitions

To illustrate this in more detail the following table represents the same state transitions as in the figure above. However in this case each state transition is complemented with an action that the ETR performs during the transition. The actions are indicated in the table (A1 to A5) and listed below.

Event	Prev State		
	No state	Periodic	Reliable
DB creation [session down]	-> Periodic A1	N/A	N/A
DB creation [session up]	-> Reliable A2	N/A	N/A
DB deletion	N/A	-> No state A3	-> No state A6
DB change	N/A	- A1	- A2
Session up	-	-> Reliable A4	N/A
Session down	-	N/A	-> Periodic A5
Timer	N/A	- A1	N/A

State machine per EID prefix, Map-Server pair

Action descriptions:

- * A1: Send UDP registration and start periodic registration timer with registration period.
- * A2: Send Registration over reliable transport session.
- * A3: Send UDP registration with zero TTL.
- * A4: Stop periodic registration timer.
- * A5: Start periodic registration timer with registration period.
- * A6: Send Registration with TTL zero over reliable transport session.

All timer start actions must be jittered.

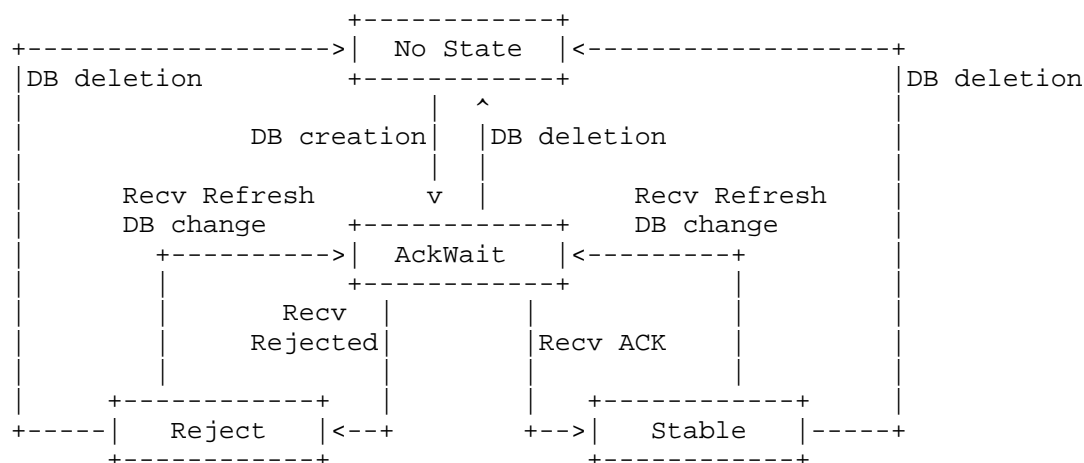
7.2.2. Transitions between reliable transport states

Once the reliable session has been established between the ETR and the Map-Server, state transitions are oriented to keep, at the ETR, an accurate representation of the registration status of each EID prefix.

When a DB Mapping is created on the ETR, a registration is sent to the Map-Server and it always transitions to the AckWait state. In this state the ETR is waiting for a confirmation about whether the Map-Server can consume the registration (Stable) or it cannot (Reject).

Most triggers in this case move the EID prefixes back to AckWait state and repeat the cycle.

This is illustrated in the following figure that depicts which particular event moves the EID prefix between the three reliable transport states:



The following table describes these transitions in more detail. As in the previous section the table also associates particular actions to each transition (A2 and A6). In this case the actions are reduced to either sending the registration or removing it (registration with TTL 0)

Event	Prev State		
	Stable	AckWait	Rejected
DB creation	N/A	N/A	N/A
DB deletion	-> No state A6	-> No state A6	-> No state
DB change	-> AckWait A2	- A2	-> AckWait A2
Recv Refresh	-> AckWait A2	- A2	-> AckWait A2
Recv Refresh [rejected]	-	- A2	-> AckWait A2
Recv ACK	-	-> Stable	-> AckWait A2
Recv Rejection	-> Rejected	-> Rejected	-

State machine per EID prefix, Map-Server pair

Action descriptions:

- * A2: Send Registration over reliable transport session.
- * A6: Send Registration with TTL zero over reliable transport session.

As described before in this document, when session comes up the Map-Server actively requests a registration refresh using the message described in Section 7.1.4. Until this message is received the ETR MUST continue using periodic UDP registrations following the procedures of [RFC9301].

7.3. Map-Server Behavior

Received registrations create/update or delete mapping state following [RFC9301]. When the Registration is received over a reliable transport, the Map-Server MUST keep the entry until the ETR indicates otherwise.

Upon receiving a registration over a reliable transport the Map-Server MUST accept (Registration Acknowledgement) or reject (Registration Rejection) the mapping so that the ETR can keep track of the registration status. If the Map-Server fails to generate these messages the ETR SHOULD reset the reliable session and continue with UDP registrations procedures again.

The Map-Server MUST send a Registration Refresh Request with global scope when a session between an ETR and the Map-Server is first established. This allows the Map-Server to obtain the complete database contents from the ETR. It also ensures the ETR completes a registration cycle for each of the EID prefixes that are going to be registered through the reliable transport.

When the Map-Server has previously rejected registrations, but the conditions change and it can start accepting them again, it SHOULD send a Refresh Request for rejected registrations (R bit set). An example of conditions changing is when the Map-Server configuration changes.

When authentication information changes on the Map-Server the session MUST be reset. The ETR and the Map-Server then follow the establishment sequence with an initial UDP based authentication to validate the use of the new authentication data.

7.3.1. Map-Server Reject and Withdraw use cases

When an ETR and a Map-Server use a reliable transport, EID registrations MUST be individually accepted by the Map-Server so that the ETR can keep track of the status of each one of them.

However, there is a possibility that the Map-Server is not in a state to accept the registration and it MAY reject or withdraw the registration. The objective of this is not to confer authoritative status to the Map-Server over the Mapping, but rather reflect operational conditions that do not make the registration possible.

The following is a short list of example use-cases when the Map-Server may withdraw or reject a registration:

- * Map-Server configuration not ready (Reject): An example of this is the case when the Map-Server has not been set up yet to accept registrations with a specific Instance-ID. When an ETR attempts to register an EID prefix with that Instance-ID, the Map-Server rejects the registration.

- * Map-Server configuration changes (Withdraw): Building on the previous example, when the configuration changes on the Map-Server to not support a specific Instance-ID anymore, all the existing registrations on the Map-Server are withdrawn
- * The reliable session is down (Withdraw): When a reliable session is interrupted and goes down, the Map-Server withdraws all registrations from the corresponding ETR as it cannot guarantee state synchronization anymore.

In general rejecting or withdrawing registrations reflect a case where ETR and Map-Server status diverge and proper state synchronization cannot be guaranteed.

8. Security Considerations

The LISP reliable transport session MUST be authenticated. On controlled RLOC networks that can guarantee that the source RLOC address of data packets cannot be spoofed, the authentication check can be a source address validation on the reliable transport packets. When the RLOC network does not provide such guarantees, reliable transport authentication MUST be used. Implementations MUST support the TCP Authentication Option (TCP-AO) [RFC5925].

When an ETR is LISP-SEC capable it can set the S-bit in the Reliable Registration packet. This is possible since the Reliable Registration uses the same format as the Map-Register described in [RFC9301], and can follow the same indications in [RFC9303] to decide when to set the bit.

Note also that, since the Map-Notify is embedded in the Reliable Notification Message, in those cases where the Map-Server needs to send an unsolicited Map-Notify (e.g., following [RFC9437]), the Map-Server MAY make use of the reliable transport to send the Map-Notify as a Reliable Mapping Notification. In such a case the message can be secured by leveraging the same security association procedures used for the Map-Notify defined in [RFC9303].

9. IANA Considerations

9.1. LISP Reliable Transport Message Types

IANA is asked to create a registry group name "LISP Reliable Transport Message Types" under the "Locator/ID Separation Protocol (LISP) Parameters" registry.

Such registry group should be populated with a one-octet registry named "Message Type" and used to identify the type of message used. The registry is populated as shown in Table 1.

Type	Name	Reference
0-15	Reserved	This document
16	Error Notification	This document
17	Registration Message	This document
18	Registration Acknowledgement Message	This document
19	Registration Rejected Message	This document
20	Registration Refresh Message	This document
21	Mapping Notification Message	This document
22-64999	Unassigned	
65000-65535	Reserved for Experimental Use	

Table 1

9.2. Error Notifications: Error Codes

This document asks IANA to create a registry of error codes used in Error Notifications over LISP reliable transport sessions. The initial error code allocation is as follows,

Code	Name	Reference
0	None	This document
1	Unrecognized message type	This document
2	Message format error	This document

Table 2

Error codes are in the range of 0 to 255. Values are assigned on a First Come First Served basis.

9.3. Registration Rejection Message: Reason Codes

This document asks IANA to create a registry of reason codes used in Registration Rejection messages over LISP reliable transport sessions. The initial reason code allocation is as follows,

Code	Name	Reference
0	None	This document
1	Not a valid site EID prefix	This document
2	Authentication failure	This document
3	Locator set not allowed	This document
4	Not defined	This document

Table 3

Error codes are in the range of 0 to 255. Values are assigned on a First Come First Served basis.

9.4. Transport Protocol Port Numbers

Following the guidelines of [RFC8126], the authors request IANA is to assign a TCP port (4342 is suggested) to sustain reliable transport over TCP.

Service Name	Port Number	Transport Protocol	Description	Reference
lisp-tcp	4342 (suggested)	tcp	LISP TCP Reliable Control Plane	TBD

Table 4

10. Sample Reliable Transport Experiences

LISP reliable transport has been in use on multiple production networks for some time. The following sections describe some of the practical lessons gathered from these deployments.

10.1. Support of EID mobility at scale

A common deployment example of LISP in production networks supports EID mobility (see [I-D.ietf-lisp-eid-mobility]) at scale. Large campuses support large numbers of endpoints that are constantly roaming from xTR to xTR in the network. In practice this implies that a large number of EIDs are registered with the Mapping System and the registration is constantly updated as they move from location to location.

The use of a reliable transport to support registrations in these environments reduces the total amount of signaling that the Map-Server needs to handle. Since periodic registrations are not used, the Map-Server reduces activity to mostly serving Map-Resolutions and handling EID updates on mobility events. In practice this has allowed deploying systems with a much larger scale of xTRs and Endpoints without the need to fine-tune specific parameters related to UDP operation.

10.2. Distributed Overlapping prefix support at scale

The support of overlapping prefixes in practical LISP deployments and at scale has revealed challenging. Following [RFC9301] (section 5.5) when an iTR resolves an EID prefix that covers other EID prefixes, it will also receive all the more specific EID prefixes as part of the Map-Reply. This becomes a problem in networks with large disaggregation of prefixes (e.g. with EID mobility).

The practical approach to solve this problem is to distribute the task of overlapping prefix handling. Following the idea of the away table introduced in [I-D.ietf-lisp-eid-mobility], the Map-Server notifies ETRs of more specific EID prefixes that their registrations cover. These ETRs use this information to build an away table with all these EID prefixes. The away table assists in generating SMRs to help iTRs in the LISP network to discover and resolve more specific EID prefixes, but only when needed.

Experience has revealed that this distributed strategy for overlapping EID prefix support does not scale well under a periodic UDP registration regime, since it increments the signaling load that the Map-Server needs to support to keep the away table of all eTRs up to date.

The use of reliable transport to support registrations and notifications solve this distributed state maintenance at scale with minimal signaling and updates in the system. Practical deployments are successfully using reliable transport to support large covering subnets and across a large number of xTRs.

10.3. Stable interconnect with external networks

Some LISP networks deployments have complex interconnects with external networks, using redistribution to and from other protocols at routers that lie at the border between the LISP network and other domains. In some of these deployments the LISP network is used as a scalable transit to interconnect external networks.

In this type of deployments the use of periodic UDP registrations and notifications subjects the routers acting as interconnects to a constant task of reevaluation of EID prefixes, to discover changes.

The use of a reliable transports largely simplifies this task. LISP routers only register changes and the Map-Servers only need to notify about these changes. Experience reveals that this highly reduces the amount of work on these border interconnects and it also highly reduces the amount of signaling observed in the network.

11. Contributors

Jesus Arango
Microsoft
United States of America
Email: jearango@microsoft.com

Johnson Leong
Uber Technologies
Market St.
San Francisco, CA 94103
United States of America

12. Acknowledgments

The authors would like to thank Noel Chiappa, Dino Farinacci, Jesper Skriver, Andre Pelletier, Les Ginsberg and Alberto Rodriguez-Natal for their contributions to this document.

13. References

13.1. 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>>.
- [RFC5925] Touch, J., Mankin, A., and R. Bonica, "The TCP Authentication Option", RFC 5925, DOI 10.17487/RFC5925, June 2010, <<https://www.rfc-editor.org/info/rfc5925>>.
- [RFC8174] Leiba, B., "Ambiguity of Uppercase vs Lowercase in RFC 2119 Key Words", BCP 14, RFC 8174, DOI 10.17487/RFC8174, May 2017, <<https://www.rfc-editor.org/info/rfc8174>>.

- [RFC9300] Farinacci, D., Fuller, V., Meyer, D., Lewis, D., and A. Cabellos, Ed., "The Locator/ID Separation Protocol (LISP)", RFC 9300, DOI 10.17487/RFC9300, October 2022, <<https://www.rfc-editor.org/info/rfc9300>>.
- [RFC9301] Farinacci, D., Maino, F., Fuller, V., and A. Cabellos, Ed., "Locator/ID Separation Protocol (LISP) Control Plane", RFC 9301, DOI 10.17487/RFC9301, October 2022, <<https://www.rfc-editor.org/info/rfc9301>>.
- [RFC9303] Maino, F., Ermagan, V., Cabellos, A., and D. Saucez, "Locator/ID Separation Protocol Security (LISP-SEC)", RFC 9303, DOI 10.17487/RFC9303, October 2022, <<https://www.rfc-editor.org/info/rfc9303>>.
- [RFC9437] Rodriguez-Natal, A., Ermagan, V., Cabellos, A., Barkai, S., and M. Boucadair, "Publish/Subscribe Functionality for the Locator/ID Separation Protocol (LISP)", RFC 9437, DOI 10.17487/RFC9437, August 2023, <<https://www.rfc-editor.org/info/rfc9437>>.

13.2. Informative References

- [I-D.ietf-lisp-eid-mobility]
Portoles-Comeras, M., Ashtaputre, V., Maino, F., Moreno, V., and D. Farinacci, "LISP L2/L3 EID Mobility Using a Unified Control Plane", Work in Progress, Internet-Draft, draft-ietf-lisp-eid-mobility-16, 8 May 2025, <<https://datatracker.ietf.org/doc/html/draft-ietf-lisp-eid-mobility-16>>.

Authors' Addresses

Balaji Pitta
Cisco Systems
Tasman Drive
San Jose, CA 95134
United States of America
Email: bvenkata@cisco.com

Marc Portoles
Cisco Systems
Tasman Drive
San Jose, CA 95134
United States of America
Email: mportole@cisco.com

Darrel Lewis
ICANN
Los Angeles, CA 90292
United States of America
Email: darrel.lewis@icann.org

Isidor Kouvelas
Arista Networks Inc.
5453 Great America Parkway
Santa Clara, CA 95054
United States of America
Email: isidor@kouvelas.net

Chris Cassar
Rivian Automotive
United Kingdom
Email: christiancassar@acm.org