Network Working Group   C. Cassar
Internet-Draft   Cisco Systems
Intended status: Experimental   I. Kouvelas
Expires: February 19, 2017   Arista Networks Inc.
                                                         D. Lewis
                                                        J. Arango
                                                         J. Leong
                                                    Cisco Systems
                                                  August 18, 2016

LISP Map Server Reliable Transport
draft-kouvelas-lisp-map-server-reliable-transport-02.txt

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 use cases for LISP have increased the amount of
state that needs to be communicated with requirements that are not
satisfied by the current mechanism. This document introduces the use
of a reliable transport for ETR to Map-Server communication in order
to eliminate the periodic messaging overhead, while providing
reliability, flow-control and endpoint liveness detection.

This document has been renamed to avoid ambiguity. It is an update
to [I-D.kouvelas-lisp-reliable-transport].

Status of This Memo

Copyright Notice

Table of Contents

1. Introduction

   The communication channel between LISP ETRs and Map-Servers is based
   on unreliable UDP message exchange [RFC6833]. 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. mobility, database redistribution). Use
cases with a large set of EID-Prefixes behind an ETR will result in a
much higher load. An example is LISP mobility deployments where EID-
Prefixes are limited to host entries. ETRs may have thousands of
hosts to register resulting in hundreds of Map-Register and Map-
Notify messages per 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 or SCTP as a
LISP reliable transport. The initial application for the LISP
reliable transport session is the support of scalable EID prefix
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 the reliable transport
session for EID prefix registration is an alternative and does not
replace the existing UDP based mechanism.

2. Requirements Notation

The key words "MUST", "MUST NOT", "REQUIRED", "SHALL", "SHALL NOT",
"SHOULD", "SHOULD NOT", "RECOMMENDED", "MAY", and "OPTIONAL" in this
document are to be interpreted as described in [RFC2119].

3. **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.

   To support protocol extensibility when new applications, or
   extensions to existing applications are introduced, the messages are
   based on a TLV format.

```
 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              |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Message ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                         Message Data                      ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Message End Marker                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                  Reliable transport message format

   o Type: 16 bit type field identifying the message type.

   o 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.

   o ID: A 32-bit value that identifies the message. May be used by
      the receiver to identify the message in replies or notification
      messages.

   o Data: Type specific message contents.

   o 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.

   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 relationship with the peer
   down for the feature, or for all LISP features.

[4]. Session Establishment

   To ensure backwards compatibility, the map server and ETR MUST
   communicate via unreliable UDP messages until a TCP session between
   the two is successfully established.

   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 accepts TCP
   connections only from those ETRs that have been authenticated via UDP
   map-register messages.

   The ETR assumes the active role of the TCP session establishment by
   connecting to the map server once it has received a UDP map-notify
   message.

   When a TCP session goes down, UDP authentication must take place
   before a new TCP session is established. The map-server will not
   accept a connection from the ETR until a UDP map-register has been
   received. Similarly, the ETR will not attempt to establish a session
   with the map server until an UDP map-notify message has been
   received.

   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 will only establish a single session with the
   map server.

[5]. 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.

```
 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 = 16          |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Message ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Error Code    |                  Reserved                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|    Offending Message Type     |   Offending Message Length    |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Offending Message ID                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Offending Message Data            ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Message End Marker                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```
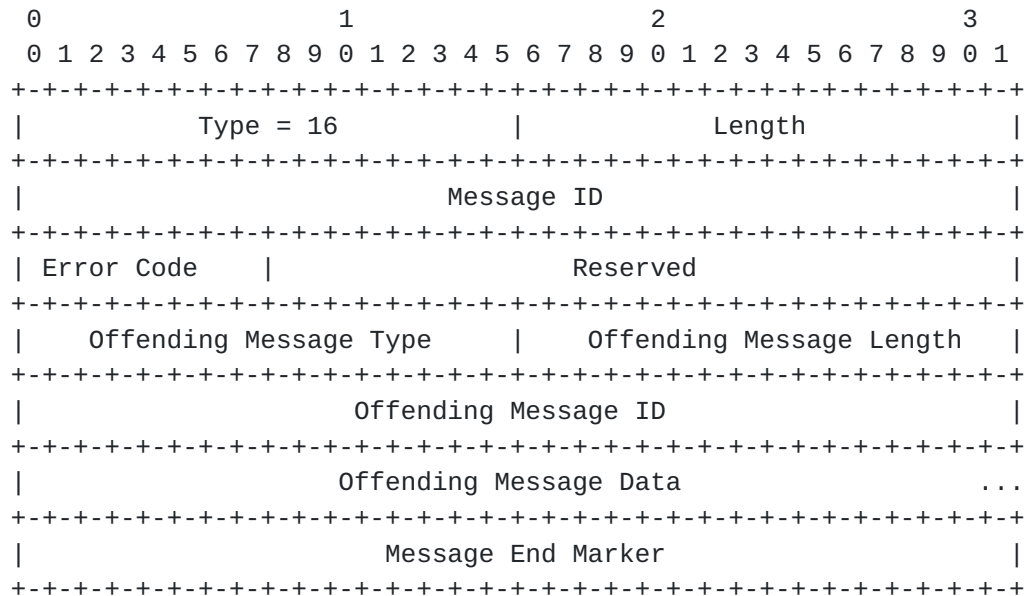
                   Error Notification message format

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

   * Unrecognized message type.

   * Message format error.

o Reserved: Set to zero by the sender and ignored by the receiver.

o 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.

o 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.

o Offending Message ID: A 32-bit field that is set to the Message ID
   field of the offending message.

o 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.

o 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.

6. 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, 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 an explicit de-registration (a
registration carrying a zero TTL) is needed to delete the 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.

6.1. Reliable Mapping Registration Messages

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

6.1.1. Registration Message

   The reliable transport registration message is used to communicate
   EID to RLOC mapping registrations from the ETR to the Map-Server. To
   increase code reuse, the "Message Data" field uses the same format as
   the UDP Map-Registers but without the IP and UDP headers. A reliable
   registration message MUST contain a single mapping-record. The map
   server MUST discard any reliable registration message that contains
   more than one mapping record.

   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           |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Message ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Map-Register message                  ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
...                   Map-Register message                     |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Message End Marker                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                      Registration message format

6.1.2. 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 by a reliable transport session Registration message.
   The Registration Acknowledgement message does not carry a mapping
   record (the map servers view of the mapping). This is accomplished
   by the LISP reliable transport Map Notification message.

```
  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           |            Length             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                          Message ID                           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 | Prefix-Length |        EID-Prefix-AFI         | EID-Prefix  ...
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                      Message End Marker                        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

             Registration Acknowledgement message format

   o Prefix-Length: Mask length for the EID prefix.

   o EID-Prefix AFI: Address family identifier for the EID prefix in
      the following field.

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

6.1.3. **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 recently-sent
   registration that is being immediately rejected. A withdrawal refers
   to a previously accepted registration that is no longer acceptable,
   perhaps due to a configuration change in the map-server. The ETR
   must keep track of rejected EID prefixes and be prepared to re-
   register their mappings when requested through a registration refresh
   message.

```
  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           |            Length             |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                          Message ID                           |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |     Reason    |           Reserved            | Prefix-Length |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |         EID-Prefix-AFI        |          EID-Prefix      ...
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
 |                      Message End Marker                        |
 +-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

             Registration Rejection message format

   o 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.

   o Reserved: This field is reserved for future use. Set to zero by
      the sender and ignored by the receiver.

   o Prefix-Length: Mask length for the EID prefix.

   o EID-Prefix-AFI: Address family identifier for the EID prefix in
      the following field.

   o EID-Prefix: The EID prefix being rejected or withdrawn.

6.1.4. Registration Refresh Message

   Sent by the Map-Server to the ETR to request the (re-)transmission of
   EID prefix database mapping Registration messages.

```
 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           |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                           Message ID                          |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|     Scope     |R|        Reserved         | Prefix-Length |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|       EID-Prefix-AFI          |          EID-Prefix       ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                       Message End Marker                      |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                  Registration Refresh message format

   o 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.

o R: Request from 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.

o Reserved: This field is reserved for future use. Set to zero by
  the sender and ignored by the receiver.

o Prefix-Length: Mask length for the EID prefix. Refer to scope for
  more details.

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

o EID-Prefix: The EID prefix being refreshed. Refer to scope for
  more details.

6.1.5. **Mapping Notification Message**

   Mapping Notification messages communicate the Map-Server view of the
   mapping for an EID prefix and no longer serve as a registration
   acknowledgement. Mapping Notifications do not need message level
   authentication as they are received over a reliable transport session
   to a known Map-Server. Note that reliable transport Mapping
   Notification messages do not reuse the UDP Map-Notify message format.

```
 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 = 21            |             Length            |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                          Message ID                           |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        xTR-ID 128 bits                     ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        xTR-ID 128 bits                     ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        xTR-ID 128 bits                     ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        xTR-ID 128 bits                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        site-ID 64 bits                     ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        site-ID 64 bits                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                        Mapping Record                      ...
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
...                       Mapping Record                        |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
|                      Message End Marker                       |
+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

                  Mapping Notification message format

   o xTR-ID: xTR-ID taken from the last valid registration the map-
      server received for the EID-prefix conveyed in the mapping record.

   o site-ID: site-ID taken from the last valid registration the map-
      server received for the EID-prefix conveyed in the mapping record.

   o Mapping Record: Mapping record of the EID-prefix the map-server is
      conveying to the ETR.

6.2. ETR Behavior

   The ETR operates the following per EID prefix, per MS state machine
   that defines the reliable transport EID prefix registration behavior.

   There are five states:

   o No state: The local EID database prefix does not exist.

   o Periodic: The local EID database prefix is being periodically
      registered through UDP Map-Register messages as specified in [].

   o 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.

   o 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.

   o 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:

   o DB creation: The local EID database entry for the EID prefix is
      created.

   o DB deletion: The local EID database entry for the EID prefix is
      deleted.

   o DB change: The mapping contents or authentication information for
      the local EID database entry changes.

   o Session up: The reliable transport session to the Map-Server is
      established.

   o Session down: The reliable transport session the Map-Server goes
      down.

   o Recv Refresh: A Registration refresh message is received from the
      Map-Server.

o Recv ACK: A Registration Acknowledge message is received from the
   Map-Server.

o Recv Rejected: A Registration Rejected message is received from
   the Map-Server.

o Periodic timer: The timer that drives generation of periodic UDP
   Map-Register messages fires.

The state machine is:

Event	Prev State	
	No state	Periodic
DB creation [session down]	-> Periodic A1	N/A
DB creation [session up]	-> AckWait A2	N/A
DB deletion	N/A	-> No state A3
DB change	N/A	- A1
Session up	-	-> Stable A4
Session down	-	N/A
Recv Refresh	-	N/A
Recv Refresh [rejected]	-	N/A
Recv ACK	-	N/A
Recv Rejection	-	N/A
Timer	N/A	- A5

                    xTR per EID prefix per MS state machine

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
Session up	N/A	N/A	N/A
Session down	-> Periodic A7	-> Periodic A7	-> Periodic A7
Recv Refresh	-> AckWait A2	- A2	-> AckWait A2
Recv Refresh [rejected]	-	- A2	-> AckWait A2
Recv ACK	-	-> Stable	-> AckWait A2
Recv Rejection	-> Rejected	-> Rejected	-
Timer	N/A	N/A	N/A

              xTR per EID prefix per MS state machine

Action descriptions:

o A1: Start periodic registration timer with zero delay.

o A2: Send Registration over reliable transport session.

o A3: Send UDP registration with zero TTL.

o A4: Stop periodic registration timer.

   o A7: Send UDP registration and start periodic registration timer
      with registration period.

   o A6: Send Registration with TTL zero over reliable transport
      session.

   o A7: Start periodic registration timer with registration period.

   All timer start actions must be jittered.

   When the reliable transport session is established the ETR moves the
   state machine into the Stable state without first registering the EID
   prefix over the reliable transport session. The map server will send
   a refresh message with a scope of 0 that will trigger the
   registration message to be sent. Because other applications may be
   using the reliable session, the refresh message signals the ETR that
   the map server supports reliable map registration messages. This
   model will also allow future optimizations where the Map-Server may
   retain registration state from a previous instantiation of the
   reliable transport session with the ETR and only request the refresh
   of EID prefix state beyond some negotiated session progress marker.

   Aa Map-Server authentication key change is treated as a DB change
   event and will result in triggering a new Registration message to be
   transmitted.

6.3. Map-Server Behavior

   Received registrations create/update or delete mapping state.

   A refresh with global scope is sent when a session between the ETR
   and map-server is first established so the map-server can obtain the
   complete database contents from the ETR. This refresh is also
   serving as a capability signaling from the map-server to the ETR that
   it can support reliable registration.

   Refresh for rejected registrations sent (R bit set) when a new EID
   prefix is configured on the Map-Server.

   Refresh is sent whenever authentication key is changed or EID prefix
   is deconfigured. Upon reception of the registration map-server can
   decide whether to acknowledge the registration or issue rejection.

   Mapping Notification message sent whenever the mapping for a
   registered or more specific prefix for which notifications are
   requested changes. ETR acknowledgement or rejection messaging for
   Mapping Notification is not required because the ETR decides how to
   process the message based on the registered mapping information. If

the mapping information changes the resulting registration will
trigger a new Mapping Notification message from the Map-Server.

7. Security Considerations

The LISP reliable transport session SHOULD 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 SHOULD be used. Implementations SHOULD
support the TCP Authentication Option (TCP-AO) [RFC5925] and SCTP
Authenticated Chunks [RFC4895].

8. IANA Considerations

8.1. LISP Reliable Transport Message Types

Assignment of new LISP reliable transport message types is done
according to the "IETF Review" model defined in [RFC5266].

The initial content of the registry should be as follows.

```
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
22-30        Reserved for EID membership distribution  TBD
31-64999     Unassigned
65000-65535  Reserved for Experimental Use
```

8.2. Transport Protocol Port Numbers

TCP port 4342 already reserved for LISP CONS that is now obsolete.
Repurpose for reliable transport over TCP. Reserve an SCTP port.

9. Acknowledgments

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

10. Normative References

   [I-D.ietf-lisp-lcaf]
              Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical
              Address Format (LCAF)", draft-ietf-lisp-lcaf-14 (work in
              progress), July 2016.

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

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

   [RFC5266] Devarapalli, V. and P. Eronen, "Secure Connectivity and
              Mobility Using Mobile IPv4 and IKEv2 Mobility and
              Multihoming (MOBIKE)", BCP 136, RFC 5266,
              DOI 10.17487/RFC5266, June 2008,
              <http://www.rfc-editor.org/info/rfc5266>.

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

   [RFC6833] Fuller, V. and D. Farinacci, "Locator/ID Separation
              Protocol (LISP) Map-Server Interface", RFC 6833,
              DOI 10.17487/RFC6833, January 2013,
              <http://www.rfc-editor.org/info/rfc6833>.

Authors' Addresses

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

   Email: ccassar@cisco.com

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

Email: kouvelas@cisco.com

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

Email: darlewis@cisco.com

Jesus Arango
Cisco Systems
Tasman Drive
San Jose, CA 95134
USA

Email: jearango@cisco.com

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

Email: joleong@cisco.com