Network Working Group   V. Moreno
Internet-Draft   Google LLC
Intended status: Experimental   D. Farinacci
Expires: July 22, 2022   lispers.net
                                                     A. Rodriguez-Natal
                                                   M. Portoles-Comeras
                                                               F. Maino
                                                              S. Hooda
                                                         Cisco Systems
                                                      January 18, 2022

                 Uberlay Interconnection of Multiple LISP overlays
                        draft-moreno-lisp-uberlay-05

Abstract

   This document describes the use of the Locator/ID Separation Protocol
   (LISP) to interconnect multiple disparate and independent network
   overlays by using a transit overlay. The transit overlay is referred
   to as the "uberlay" and provides connectivity and control plane
   abstraction between different overlays. Each network overlay may use
   different control and data plane approaches and may be managed by a
   different organization. Structuring the network into multiple
   network overlays enables the interworking of different overlay
   approaches to data and control plane methods. The different network
   overlays are autonomous from a control and data plane perspective,
   this in turn enables failure survivability across overlay domains.
   This document specifies the mechanisms and procedures for the
   distribution of control plane information across overlay sites and in
   the uberlay as well as the lookup and forwarding procedures for
   unicast and multicast traffic within and across overlays. The
   specification also defines the procedures to support inter-overlay
   mobility of EIDs and their integration with the intra-overlay EID
   mobility procedures defined in [draft-ietf-lisp-eid-mobility].

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 [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 July 22, 2022.

Copyright Notice

Table of Contents

1. Introduction

The main motivation for this specification is to provide a
methodology for the interconnection of LISP domains that may use
disparate control and/or data plane approaches. For instance, one
domain may use native LISP encapsulation for its data plane and a DDT
based mapping system, while another domain may use VXLAN-GPE
encapsulation and a mapping system based on
[I-D.farinacci-lisp-decent]. Furthermore, one domain may use an IPv4
RLOC space and the other domain may use an IPv6 RLOC space and there
may not be connectivity between the domains at the RLOC level. We
propose a method to interconnect and enable interoperability between
these disparate LISP overlay networks by connecting them to a common
transit LISP overlay.

In order to provide interworking across implementations of overlays
that may use different control and data plane approaches, a LISP
network may be structured as a collection of site-overlays
interconnected by a transit area. Each site-overlay is a fully
functional overlay network and has its own set of Map Servers and Map
Resolvers. Site-overlays share a border xTR with a transit area.
Connectivity between site-overlays is provided via the transit area
which we will refer to as "The Uberlay". This specification
describes the Control Plane and Forwarding procedures for the
implementation of an Uberlay connected multi-overlay LISP network.
This approach to the structure of a LISP network may also enable
regional failure survivability and fault isolation.

2. Definition of Terms

LISP related terms, notably Map-Request, Map-Reply, Ingress Tunnel
Router (ITR), Egress Tunnel Router (ETR), Map-Server (MS) and Map-
Resolver (MR) are defined in the LISP specification [RFC6830].

Terms defining interactions with the LISP Mapping System are defined
in [RFC6833].

Terms related to the procedures for signal free multicast are defined
in [RFC8378].

The following terms are here defined to facilitate the descriptions
and discussions within this particular document.

Site-Overlay - Overlay network at a specific area or domain. This
overlay network has a dedicated Mapping System.

Border-xTR - xTR that connects a site-overlay to one or more
uberlays.

xTR - LISP Tunnel Router as defined in [RFC6833]. An xTR connects
end-points to the site-overlay.

Local Mapping System - Mapping system of the site-overlay

Uberlay - Overlay network that interconnects different site-overlays
to each other. The Uberlay has a dedicated Mapping System and
creates an overlay amongst the border xTRs which connect different
site-overlays.

Uberlay Mapping System - Autonomous mapping system dedicated to the
uberlay.

Site-Overlay EIDs - Also referred to as local site-overlay EIDs,
these are the EIDs that are connected to xTRs in a particular site-
overlay and are registered in their own local site-overlay mapping
system. These EIDs will also be registered in the uberlay but not in
the remote site-overlay mapping systems.

Remote site-overlay EIDs - These are EIDs connected and registered in
site-overlays other than the local site-overlay.

Local site-overlay EIDs - These are EIDs connected and registered in

the local site-overlay.

3. Interconnecting multiple LISP site-overlays via the Uberlay

   A LISP network can be structured as a collection of LISP site-
   overlays that are interconnected by one or more LISP Uberlays.

   A LISP site-overlay is an overlay network that has its own set of
   xTRs, its own dedicated Mapping System and it may have a dedicated
   RLOC space, separate from that of other site-overlays or the uberlay.
   A LISP uberlay is also an overlay network with its own set of xTRs,
   its own dedicated Mapping System and it may have its own dedicated
   RLOC space. When the RLOC spaces are dedicated, RLOC routes in the
   local underlay do not leak across to the underlay of other site-
   overlays.

   A site-overlay will have xTRs and Border xTRs. The xTRs provide
   connectivity to the local site-overlay EIDs, which are the EIDs that
   are locally connected to the overlay-site. The Border xTRs are Re-

   encapsulating Tunnel Routers (RTRs) that connect the site-overlays to
   the LISP Uberlay in the transit network. xTRs participate in one
   site-overlay and one site-overlay only. Border xTRs participate in
   the mapping system of the site-overlay it resides in and the mapping
   system of the uberlay it connects the site-overlay to. Border xTRs
   may be shared by more than one site-overlay.

   The different site-overlays can be interconnected by an uberlay. The
   uberlay consists of a dedicated Mapping System and the set of Border
   xTRs that connect the participating site-overlays to the Uberlay and
   the Uberlay Mapping System.

   Each site-overlay will have its own set of Map Servers and Map
   Resolvers (MS/MRs) which operate as an autonomous Mapping System.
   The Uberlay Mapping System is also autonomous and includes all
   necessary Map Servers and Map Resolvers. Any of the Mapping Systems,
   in site-overlays or in the Uberlay, follow the control plane
   specification in [RFC6833] and may be structured as a Distributed
   Delegation Tree (DDT) per [RFC8111]for the purposes of horizontal
   scaling or other optimizations within each Mapping System.

   The MS/MRs can be co-located with the border-xTRs of the site-overlay

When a Border xTR services more than one site-overlay, and the MS/MRs
are instantiated on the Border xTR, logical instances of MS/MRs must
be dedicated to each site-overlay.

This specification defines the interaction between the Mapping
Systems of the site-overlays and the uberlay to deliver a multi-
overlay hierarchical network. The forwarding procedures relevant to
the border xTRs are also specified. Figure 1 illustrates the multi-
overlay network.

```
      +------------------------------+
      |  +-----+   +-----+   +-----+  |
      |  | xTR |   | xTR |   | xTR |  |
      |  +-----+   +-----+   +-----+  |
      |                               |
      |                    +-------+  |   RLOC space 1
      |     Site Overlay 1 | MS/MR |  |   (underlay 1)
      |                    +-------+  |
      |                               |
      |                               |
      |     +--------+  +--------+    |
   +-----| Border |--| Border |----+
   +-----|  xTR   |--|  xTR   |----+
      |     +--------+  +--------+    |
      |                               |
```

```
|                              |
|                              |
|                 +-------+ |      Uberlay
|     Uberlay     | MS/MR | |    RLOC Space
|                 +-------+ |   (Transit Underlay)
|                              |
|                              |
|         +----------+         |
+--------|   Border  |---------+
+--------|    xTR    |---------+
|         +----------+         |
|                              |
|                 +-------+ |    RLOC space 2
|   Site Overlay 2| MS/MR | |    (underlay 2)
|                 +-------+ |
|                              |
|  +-----+   +-----+   +-----+  |
|  | xTR |   | xTR |   | xTR |  |
|  +-----+   +-----+   +-----+  |
+------------------------------+
```
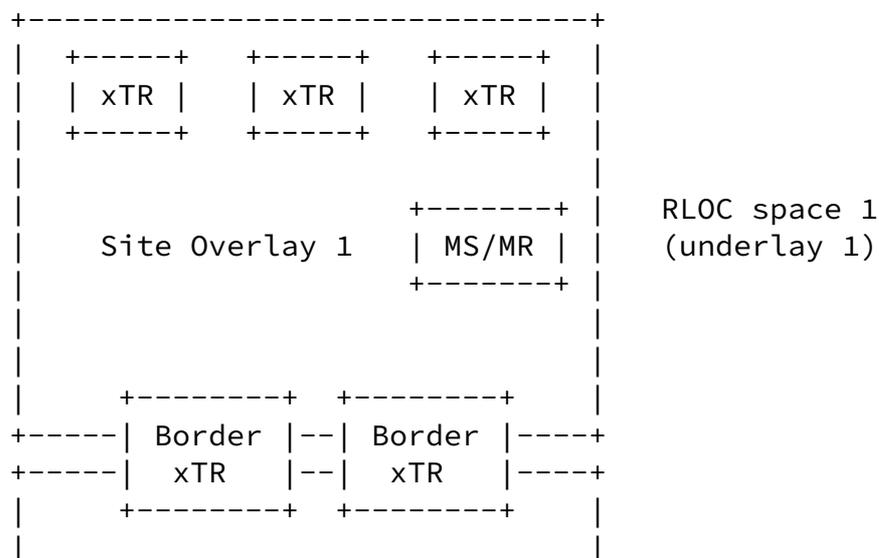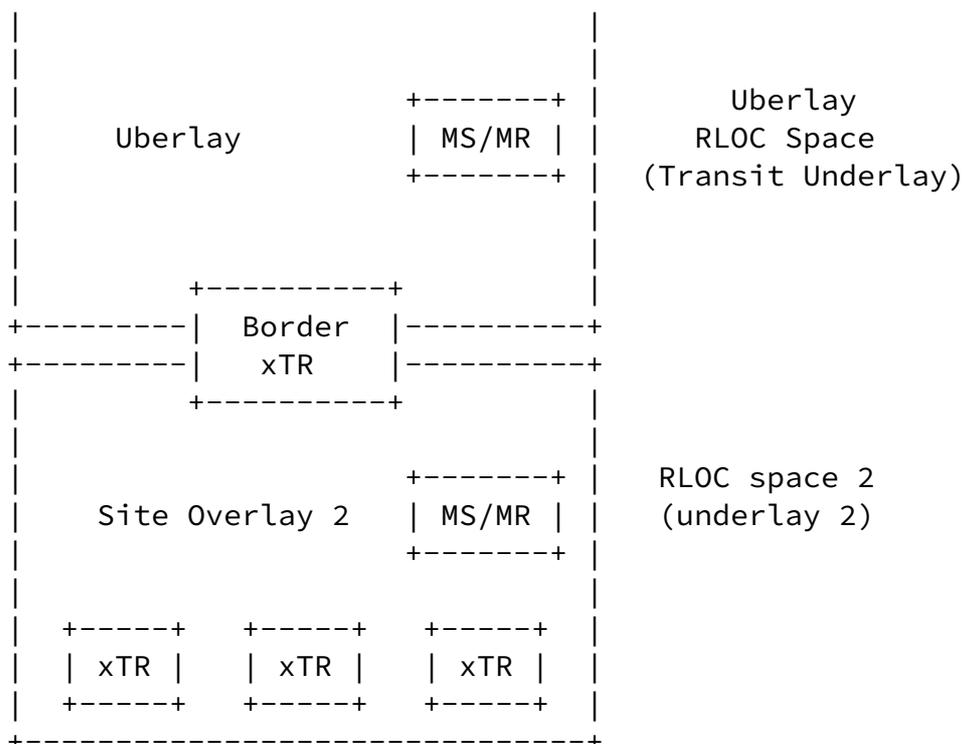
Figure 1. Site-overlays connected via Uberlay

Structuring the LISP network as multiple site-overlays interconnected
by an uberlay delivers the following benefits:

o Enable the interworking of diverse site-overlay implementations in
   which different mapping systems and encapsulations may be used.

o Enhanced resiliency through regional failure survivability.
   Failures in one site-overlay or failures in a portion of the
   underlay should not affect other site-overlays.

o Reduce the state of the site-overlay control plane. The site-
   overlay control plane will only maintain state for EIDs that are
   connected to xTRs within the site-overlay These EIDs are referred
   to as local site-overlay EIDs in this specification. Remote site-

overlay EIDs will not be explicitly registered within the site-
        overlay.

   o   Separate the RLOC space of the different site-overlays as well as
       the uberlay RLOC space. Each site-overlay will only need
       reachability to its own RLOCs, making the RLOCs private to the
       site-overlay Similarly, the uberlay RLOC space does not require
       knowledge of site-overlay specific RLOCs. This simplifies the
       underlay routing protocol structure and reduces the state that
       must be handled and maintained by the underlay routing protocols.

   o   Reduced latency for local site-overlay EID registrations may be
       achieved when xTRs and Map Servers are topologically close.
       Topological proximity is expected when the RLOC spaces for the
       different overlays are kept separate.

   o   Reduced latency for local site-overlay EID lookups may be achieved
       when xTRs, Map Resolvers and Map Servers are topologically close.
       Topological proximity is expected when the RLOC spaces for the
       different overlays are kept separate.

   o   Creates a multicast replication hierarchy where the Border RTRs
       serve as the points of multicast replication for multicast traffic
       that spans multiple site-overlays.

   o   Creates a distributed structure of RTRs that can be leveraged for
       the deployment of NAT traversal in the RLOC space.

   o

3.1. Logical Topology Considerations

   xTRs as defined in RFC6833bis connect a network to the LISP overlay
   and register the EID prefixes from the connected network to the LISP
   mapping system. Border xTRs, as defined in this document, will
   connect site-overlays to the Uberlay and register the EID prefixes
   that originate in a site-overlay in the Mapping System of the
   Uberlay. Conversely, a border xTR may register EID prefixes present
   in the Uberlay Mapping System into the Mapping System of a particular
   site-overlay. Furthermore, border xTRs may connect Uberlays to each

   other and register the EID prefixes from one Uberlay into the other.

There is no provision for the detection of registration loops when concatenating site-overlays and Uberlays, thus any interconnection of overlay domains (site-overlays or Uberlays) must be done in a loop free topology.

A loop free topology is hereby defined for reference. This is a general concept and is not encoded into any of the protocol messages in LISP. A loop free topology limits the peerings between Uberlays and/or overlays to a strict hierarchy. At the top of the hierarchy is a single central Uberlay or Core Uberlay. The loop free topology is defined by two simple rules: Uberlays must only connect to Uberlays in the next consecutive level of hierarchy (no level skipping) and uberlays within the same level of hierarchy must not connect to each other. The loop-free topology hierarchy is illustrated in Figure 2.

```
    +---------------+                  +---------------+
    | site-overlay 1 |                 | site-overlay 2 |
    |    (Level 2)   |                 |    (Level 2)   |
    +---------------+                  +---------------+
            +---+                              +---+
            |RTR|                              |RTR|
            +---+                              +---+
        +----------+                      +----------+
        | Uberlay 1 |                     | Uberlay 2 |
        | (Level 1) |                     | (Level 1) |
        +----------+                      +----------+
             +---+        +---+
             |RTR+--------+RTR|
             +--++        ++--+
               |   Core     |
               |  Uberlay   |
               | (Level 0)  |
             +--++        ++--+
             |RTR+--------+RTR|
             +---+        +---+
        +----------+                      +----------+
        | Uberlay 3 |                     | Uberlay 4 |
        | (Level 1) |                     | (Level 1) |
        +----------+                      +----------+
            +---+                              +---+
            |RTR|                              |RTR|
            +---+                              +---+
    +---------------+                  +---------------+
    | site-overlay 3 |                 | site-overlay 4 |
    |    (Level 2)   |                 |    (Level 2)   |
    +---------------+                  +---------------+
```
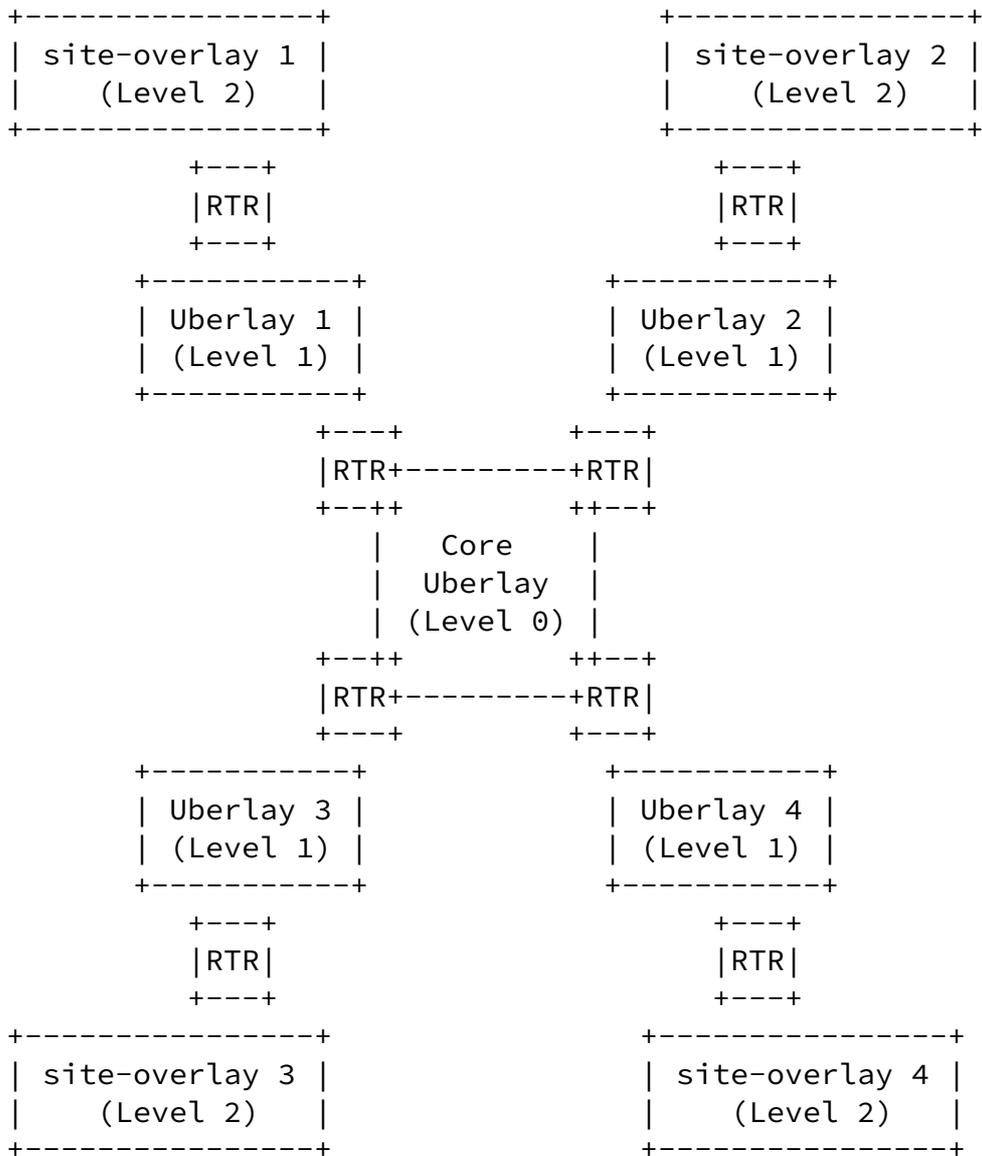
          Figure 2. Loop-free topology hierarchy

4. General Procedures

   A site-overlay maintains state only for its local site-overlay EIDs
   and RLOCs. Tunnels never cross site-overlay or uberlay boundaries.
   Remote site-overlay EIDs are reachable at the source site-overlay via
   a default mapping which will steer all traffic destined to remote
   site-overlay EIDs to the border xTRs where it can be handed off to
   the uberlay. Traffic will be decapsulated at the border xTRs and a
   lookup in the uberlay mapping system will determine the site-overlay
   to which traffic is to be re-encapsulated. The uberlay maintains
   state for the EIDs of all interconnected site-overlays and will steer

traffic from the source site-overlay to the destination site-overlay
by encapsulating the traffic from the source site-overlay border xTR

to the destination site-overlay border xTR. At the border xTR of the
destination site-overlay, traffic will be de-capsulated, a lookup in
the local destination site-overlay Mapping System will take place and
traffic will be re-encapsulated to the xTR that connects to the
destination EID. Thus, forwarding is achieved by concatenating
overlays and doing Re-encapsulation at the border xTRs to forward the
traffic from the Ingress site-overlay to the Egress site-overlay via
the Uberlay.

Traffic for non-LISP sites, or for EIDs not registered in any site-
overlay, will also be forwarded to the border xTR where it will be
forwarded or dropped as appropriate.

4.1. Control Plane Procedures

Local EIDs must be registered by the xTRs into the local Mapping
System of the site-overlay. Intra-site communication follows the
standard procedures of registration, resolution, caching and
encapsulation defined in [I-D.ietf-lisp-rfc6830bis] and
[I-D.ietf-lisp-rfc6833bis] amongst the xTRs within the local site-
overlay.

The border xTRs at a site-overlay should have a local site-overlay
RLOC-set and will also have an uberlay RLOC-set. The local site-
overlay RLOC-set is in the private site-overlay RLOC space and is
used by the border xTRs as the RLOC set for any mappings it may
register with the site-overlay Mapping System. The uberlay RLOC-set
for the border-xTRs of a particular site-overlay are the RLOCs to
reach the site-overlay in the uberlay RLOC space. The border xTR
will use the uberlay RLOC-set in any mappings it may register with
the uberlay Mapping System. It is possible for a deployment to
connect the RLOC spaces of the site-overlays and the uberlay, it is
also possible in the scenario of a common RLOC space for the uberlay
and local site-overlay RLOC sets to be one and the same. Any
implementation of this specification should support disjoint RLOC
spaces or joint RLOC spaces.

The border xTRs must register a default EID-prefix as specified in
Section 4.3 with the local site-overlay Mapping System. Remote EIDs

will be generally reachable by xTRs in a site-overlay using the
default EID mapping registered by the border xTRs. This is expected
to be the mapping used for most communications to remote site-overlay
EIDs. Remote site-overlay EIDs may be registered with the local
site-overlay Mapping System for the purposes of supporting inter-
overlay EID mobility as specified in Section 6, these mappings will
be preferred over the default EID mapping whenever present.

   Local EIDs registered with the site-overlay mapping system must also
   be registered with the Uberlay Mapping System. The registration of
   the local site-overlay EIDs with the uberlay Mapping System is
   originated by the Border xTRs. The local site-overlay EIDs SHOULD be
   aggregated into the shortest covering prefix possible before being
   registered with the uberlay Mapping System. How this aggregation is
   achieved is implementation specific.

   In order to be able to register the local site-overlay EIDs with the
   uberlay Mapping System, the border xTRs must subscribe to all EIDs
   registered in their local site-overlay Mapping System. This is a
   subscription to 0.0.0.0/0 (or 0::/0) with the N-bit set as specified
   in [I-D.ietf-lisp-pubsub]. The subscription populates all local
   site-overlay EID mappings in the map-cache of the border xTRs.

   Once received through the subscription, the local site-overlay EIDs
   in the map-cache at the border xTRs must be registered by the border
   xTRs with the uberlay Mapping System. The local site-overlay EIDs
   will be registered using the 'uberlay' RLOC-set for the registering
   border xTR.

   Following [I-D.ietf-lisp-eid-mobility], the border xTRs will also
   subscribe to any EID prefixes it registers with the uberlay Mapping
   System. This allows the border xTRs to get Map Notify messages from
   the uberlay Mapping System for EID prefixes that may move from their
   local site-overlay to a remote site-overlay.

4.1.1. Split-horizon at the Border xTRs

   Remote site-overlay EIDs may be learnt at a border xTR due to
   resolution of a remote destination EID or due to a mobility event as
   specified in Section 6. Remote site-overlay EIDs learnt from the

uberlay will be installed in the map-cache of the border xTR with the corresponding remote uberlay RLOC-set for the remote border xTR. When these remote site-overlay EIDs are learnt as a consequence of the map-notify messages defined in the Inter-overlay mobility procedures in Section 6, the EIDs will also be registered with the local site-overlay mapping system using the local site-overlay RLOC-set for the border-xTR. The remote site-overlay EIDs registered with the local site-overlay mapping system will be learnt back at the border xTR because of the border xTR's subscription to all local site-overlay EIDs. This can cause the mapping for the remote EID that is installed in the border xTR map-cache to flip flop between the uberlay RLOC-set and the local site-overlay RLOC-set.

In order to avoid this flip flopping a split horizon procedure must be implemented. When a mapping received at the border xTR (as part of its subscription to all local site-overlay EID prefixes) has the

local site-overlay RLOC-set for the border xTR, the mapping received in the subscription corresponds to a remote site-overlay EID and should be ignored by the border xTR. The mapping should not be installed in the map-cache of the border xTR and the EIDs in the mapping should not be advertised to the uberlay. More robust split horizon mechanisms can be proposed in future revisions of this specification.

4.1.2. Border-xTR Resiliency

Redundancy at the border xTRs requires that border xTRs be logically grouped so that the redundant array doesn't create a registration loop. As border xTRs interconnect overlay domains, the border xTRs will register the EID prefixes from one domain into the neighboring domain. From the perspective of the border xTR, the EID prefixes to be registered in one domain are learnt from a neighbor domain which we will refer to as the "site-of-origin". The site-of-origin may be an overlay-site, an Uberlay or an IP network.

Border xTRs should be logically grouped in Border Sets. A border set is a group of border xTRs that register EID prefixes from the same site-of-origin. Members of a border set will register the EIDs from a particular site-of-origin into the neighboring overlay (site-overlay or uberlay) using a common site-id. The use of the site-ID namespace is locally significant to each overlay domain (site-overlay

or Uberlay) and does not require cross-domain synchronization or
dispersion. A border-xTR may be a member of multiple border sets to
allow different site-of-origin domains to be serviced by the border-
xTR. Note that not all site-of-origin domains will connect to the
same combination of border-xTRs.

EID Mappings will be tagged with a site-ID according to their site-
of-origin when they are registered by the border-xTR. The site-ID
must be maintained in the Mapping System as part of the registration
record. EID Mappings published and received at the border xTR must
include the site-ID for the EID Mapping. If the border-xTR receives
a mapping for an EID with a site-ID that matches the site-ID for one
of its border sets (site-of-origin), the Border xTR will not register
that information to the site-of-origin associated with that site-ID
and thus prevent any registration loops from occurring.

4.2. Resolution and Forwarding Procedures

Intra-site communication follows the standard procedures of
registration, resolution, caching and encapsulation defined in
[I-D.ietf-lisp-rfc6830bis] and [I-D.ietf-lisp-rfc6833bis] amongst the
xTRs within the local site-overlay.

Inter-site communication is achieved by encapsulating traffic
destined to remote site-overlay EIDs from the xTRs to the border
xTRs. Traffic will be decapsulated at the border xTRs and a lookup
in the uberlay mapping system will determine the site-overlay to
which traffic is to be re-encapsulated. The lookup should return the
uberlay RLOCs for the border xTRs of the site-overlay where the
destination EID is located. At the border xTR of the destination
overlay-site, traffic will be de-capsulated, and re-encapsulated to
the destination xTR, just like an RTR does. The border xTR already
has the destination EID in its cache per its subscription to all
local site-overlay EIDs.

When receiving encapsulated traffic, a border xTR will de-capsulate
the traffic and will do a lookup for the destination EID in its map
cache. If the destination EID is present in the map cache, the
traffic is forwarded and no lookup takes place. If the destination
EID is not present in the cache, the destination EID is not in any
local site-overlay connected to the border xTR, in which case the

border xTR will issue a map-request to all Uberlay Mapping Systems it is connected to. The criteria to determine which Mapping Systems are Uberlay Mapping Systems is simply to select those Mapping Systems with which the border xTR doesn't hold a subscription to 0.0.0.0/0 (or 0::/0).

4.2.1. Multi-overlay requests at border xTR

A Border xTR may query all Mapping Systems in all uberlays it participates in. The border xTR will then chose based on longest prefix match the more specific EID mapping provided by any of the Mapping Systems. This procedure could also include site-overlay Mapping Systems, however those are not expected to be queried as the border xTR subscribes to all EIDs in the site-overlays and the presence of the mappings in the cache will prevent any lookups. The processing of Map Requests following the multi-domain request logic works as follows:

1. The Border xTR sends a map request for the prefix that it intends to resolve to each of the uberlay Mapping Systems it participates in.

2. The Border xTR receives Map Replies from each of the different uberlay Mapping Systems it sent requests to. The Border xTR will treat the replies differently depending on their contents:

    * Negative Map Replies (NMR) are ignored and discarded unless all Map Replies received are Negative, then the border xTR follows the procedures specified in [RFC6833] for Negative Map Replies.

    * Map Replies with RLOCs that belong to the requesting border xTR are ignored.

    * Map Replies with EID prefixes that are not already in the map cache of the border xTR are accepted and cached.

    * If the EID prefix received in the Map-Reply already exists in the cache/routing table, but the Map-Reply contains a different RLOC-set than the one cached, the mappings are merged so that the RLOCs received in the Map-Reply are added to the RLOC-set previously cached for the EID prefix.

> > * If the EID prefix received in the Map-Reply is more specific
> >   or less specific than an EID prefix already cached, the
> >   mapping received MUST be cached.

It is expected that a deployment of the uberlay would include the
dynamic registration of default EIDs. It is also recommended that an
implementation adopts mechanisms for the dynamic resolution of
default EIDs. In an environment leveraging the dynamic registration
and resolution of default EIDs, the border xTR should not receive
Negative Map-Replies, but all replies (including those in response to
requests for destinations that are external to the EID space) will be
Map-replies with a non-zero locator count. Nevertheless, an
implementation could opt to not use dynamic default-EID handling. In
these cases, the border xTR will receive NMRs. The implementation of
the Border xTR should defer the decision on caching an NMR until all
relevant Map-replies are received. To this effect, the
implementation should implement mechanisms to ensure that sufficient
replies are received before programming the map-cache. The
mechanisms by which this is achieved are an implementation specific
matter and therefore not specified in this document.

When following these rules to process multi-domain requests, the
Border xTR guarantees proper discovery and use of destination
prefixes that will be associated with their corresponding overlay-
site. By ignoring the negative replies the procedure works
regardless of whether the Mapping Systems of multiple uberlays have
consistent configurations or operate individually without being aware
of the whole addressing space in the overlay fabric.

4.3. Default EID registration and treatment

Border xTRs will register a mapping to be used as a default mapping
to handle the forwarding of traffic destined to any EIDs that are not
explicitly registered. These mappings will be registered in the
local site-overlay Mapping System of each site-overlay. The RLOCs
for the mappings will be the site-overlay RLOCs of the border xTR.

This registration is intended to instruct the Mapping System to
follow the procedures in [RFC6833] for Negative Map Replies and
calculate the broadest non-registered EID prefix that includes the
requested destination EID and issue a map-reply with the calculated

EID and the RLOCs registered by the border xTRs. The map-reply for
this default mapping will have a shorter TTL to accommodate any
changes in the registrations.

The instruction to the Mapping System can be encoded as the
registration of an agreed upon distinguished name such as "Default".
The registration will contain the RLOC set desired for the default
handling.

5. Multicast Specific Procedures

This specification will focus on the procedures necessary to extend
signal-free multicast [RFC8378] across multiple site-overlays
interconnected with an uberlay. The specification will focus on the
extensions of the Sender and Receiver site procedures

5.1. Inter-site-overlay Control Plane Procedures for Signal-free
   multicast

   1. At the listener sites, xTRs with multicast listeners will follow
       the receiver site procedures described in [RFC8378]. A
       replication list will be built and registered on the site-overlay
       Mapping System for the multicast channel being joined by the
       listeners.

   2. The Mapping System for the listener site-overlay will send Map-
       Notify messages towards the multicast source or RP per [RFC8378].
       The multicast source or RP is reachable via the border-xTRs of
       the listener site-overlay via the default EID mapping registered
       in the listener site-overlay.

   3. Upon reception of the Map-Notify in the previous step, the
       listener site-overlay border-xTR will register the multicast EID
       with the uberlay Mapping System using the uberlay RLOCs for its
       site-overlay as the RLOC set for the mapping being registered.
       Only one of the RLOCs in the set should be active in the
       registration per the procedures in [RFC8378]. A replication tree
       is built in the uberlay as specified in [RFC8378].

   4. After the listener site-overlay border-xTR registers the
       multicast EID with the uberlay Mapping system, the uberlay MS
       will send a Map-Notify toward the multicast source per [RFC8378]

   5. Upon reception of the Map-Notify in the previous step, the border
       xTR at the source site-overlay registers its interest in the
       multicast EID with the source site-overlay Mapping System
       following the procedures described in [RFC8378].

5.2. Border xTR Resolution and Forwarding procedures for Signal-free
      multicast

   The mapping resolution procedures for multicast EIDs at border xTRs
   fall within the scope of the mechanisms specified in Section 4. The
   Map-replies obtained from the lookup will follow the behavior
   specified in [RFC8378] for signal-free multicast.

   Forwarding will also follow the General Procedures specified in
   Section 4 without alteration. It is worth noting that the
   concatenation of overlays between listener sites, uberlay and sender
   site-overlays creates a convenient replication structure where the
   border xTRs act as the replication points to form an optimized end-
   to-end multi-level replication tree.

6. Inter site-overlay Mobility Procedures

   The receiver and sender site procedures defined in
   [I-D.ietf-lisp-eid-mobility] apply without change to each site-
   overlay and to the uberlay. Border xTRs are connected to two or more
   overlay networks which are following the mobility procedures. An
   away table is defined at the border xTR for each overlay network it
   participates in. In order to illustrate the procedures required,
   this specification describes a scenario where a border xTR has one
   local site-overlay away table and one uberlay facing away table. The
   procedures for mobility described in this section are extensible to
   border xTRs participating in more than two overlays.

   When a map notify for an EID is received at an xTR, an away entry is
   created on the receiving side table. Any away entries for the
   specific EID in other tables on the same LISP node (xTR or RTR) must
   be removed. This general rule addresses convergence necessary for a
   first move as well as any subsequent moves (moves that take place
   after the away tables are already populated with entries for the
   moving EID due to previous moves).

   The following set of procedures highlights any additions to the
   mobility procedures defined in [I-D.ietf-lisp-eid-mobility]:

   1. Detect the roaming EID per the mechanisms described in
       [I-D.ietf-lisp-eid-mobility] and register the EID with the site-
       overlay Mapping System at the landing site-overlay

2.   The site-overlay Mapping System at the landing site-overlay must send a Map-Notify to the last registrant xTR (if it is local to the site-overlay) and to the border xTR as the border xTR subscribes to all EIDs in the site-overlay.

3.   The border xTR will install an entry for the moved host in the local away table of the border xTR.

4.   The border xTR from the landing site-overlay will register the roaming EID with the uberlay Mapping System using the uberlay RLOC-set for the landing site-overlay

5.   The Uberlay Map Server will send Map-Notify messages to the border xTRs at the departure site-overlay as specified in [I-D.ietf-lisp-eid-mobility] (border xTR with the previously registered RLOCs).

6.   Upon reception of the Map-Notify, the border xTR must check if the Map-Notify is for an EID-prefix that is covered by a broader or equal EID-prefix that is locally registered. Local registration is determined by the presence of the broader or equal EID prefix in the map-cache of the border xTR.

7.   If the roaming EID-prefix received in the Map-Notify is not covered under a previously registered EID-prefix in the local site-overlay, the EID-prefix is a newly registered prefix and no further action is required.

8.   If the roaming EID-prefix received in the Map-Notify is covered under a registered EID-prefix, the Map-Notify is due to a move event. In this case, the site-overlay border xTR must register the roaming EID prefix in the site-overlay mapping system using the site-overlay facing RLOC-set of the border-xTRs. The roaming EID-prefix must also be installed in the uberlay facing away table of the border xTR at the departure site-overlay.

9.   The departure site-overlay Map-Server will send Map-Notify messages to the xTRs at the departure site-overlay as specified in [I-D.ietf-lisp-eid-mobility] (edge xTRs with the previously registered RLOCs).

10. When the site-overlay xTR at the departure site-overlay receives
     the Map-Notify from the border xTR, it will include the EID
     prefix received in the Map-Notify in its away table per the
     procedures described in [I-D.ietf-lisp-eid-mobility].

11. Data triggered Solicit Map Requests (SMRs) will be initiated in
     the different site-overlays and the uberlay as traffic matches

     the different away tables. As specified in
     [I-D.ietf-lisp-eid-mobility], these SMRs notify the different
     ITRs involved in communications with the roaming EID that they
     must issue a new Map-Request to the mapping system to renew
     their mappings for the roaming EID.

7. Virtual Private Network (VPN) Considerations

   When supporting multiple Instance IDs as specified in
   [I-D.ietf-lisp-vpn] the Instance IDs range is divided in two sets. A
   reuse-set that can be used in each site-overlay and a global-set used
   across site-overlays and the uberlay.

   Instance-IDs that are local to a site-overlay should only provide
   intra-overlay connectivity and are in the site-overlay mapping system
   only for VPN use for the xTRs in the site-overlay. When the VPN
   reaches across site-overlays, then the global-set instance-IDs are in
   the uberlay mapping system as well as each site-overlay mapping
   system where the VPN members exist.

8. IANA Considerations

   This document has no IANA implications

9. Acknowledgements

   The authors want to thank Kedar Karamarkar, Prakash Jain and Vina
   Ermagan for their insightful contribution to shaping the ideas in
   this document. We would also like to acknowledge the valuable input
   from the workgroup chairs Joel Halpern and Luigi Iannone in refining
   the objectives of the document.

10. References

10.1. Normative References

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

   [RFC3618] Fenner, B., Ed. and D. Meyer, Ed., "Multicast Source
              Discovery Protocol (MSDP)", RFC 3618,
              DOI 10.17487/RFC3618, October 2003,
              <https://www.rfc-editor.org/info/rfc3618>.

   [RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas,
              "Protocol Independent Multicast - Sparse Mode (PIM-SM):
              Protocol Specification (Revised)", RFC 4601,
              DOI 10.17487/RFC4601, August 2006,
              <https://www.rfc-editor.org/info/rfc4601>.

   [RFC4607] Holbrook, H. and B. Cain, "Source-Specific Multicast for
              IP", RFC 4607, DOI 10.17487/RFC4607, August 2006,
              <https://www.rfc-editor.org/info/rfc4607>.

10.2. Informative References

   [I-D.farinacci-lisp-decent]
              Farinacci, D. and C. Cantrell, "A Decent LISP Mapping
              System (LISP-Decent)", draft-farinacci-lisp-decent-08
              (work in progress), August 2021.

   [I-D.ietf-lisp-eid-mobility]
              Comeras, M. P., Ashtaputre, V., Moreno, V., Maino, F., and
              D. Farinacci, "LISP L2/L3 EID Mobility Using a Unified
              Control Plane", draft-ietf-lisp-eid-mobility-08 (work in
              progress), July 2021.

   [I-D.ietf-lisp-pubsub]
              Rodriguez-Natal, A., Ermagan, V., Cabellos, A., Barkai,
              S., and M. Boucadair, "Publish/Subscribe Functionality for

             LISP", draft-ietf-lisp-pubsub-09 (work in progress), June
             2021.

   [I-D.ietf-lisp-rfc6830bis]
             Farinacci, D., Fuller, V., Meyer, D., Lewis, D., and A.
             Cabellos, "The Locator/ID Separation Protocol (LISP)",
             draft-ietf-lisp-rfc6830bis-36 (work in progress), November
             2020.

   [I-D.ietf-lisp-rfc6833bis]
             Farinacci, D., Maino, F., Fuller, V., and A. Cabellos,
             "Locator/ID Separation Protocol (LISP) Control-Plane",
             draft-ietf-lisp-rfc6833bis-30 (work in progress), November
             2020.

   [I-D.ietf-lisp-vpn]
             Moreno, V. and D. Farinacci, "LISP Virtual Private
             Networks (VPNs)", draft-ietf-lisp-vpn-08 (work in
             progress), January 2022.

   [RFC6407] Weis, B., Rowles, S., and T. Hardjono, "The Group Domain
              of Interpretation", RFC 6407, DOI 10.17487/RFC6407,
              October 2011, <https://www.rfc-editor.org/info/rfc6407>.

   [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>.

   [RFC6831] Farinacci, D., Meyer, D., Zwiebel, J., and S. Venaas, "The
              Locator/ID Separation Protocol (LISP) for Multicast
              Environments", RFC 6831, DOI 10.17487/RFC6831, January
              2013, <https://www.rfc-editor.org/info/rfc6831>.

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

   [RFC7348] Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger,
              L., Sridhar, T., Bursell, M., and C. Wright, "Virtual
              eXtensible Local Area Network (VXLAN): A Framework for
              Overlaying Virtualized Layer 2 Networks over Layer 3
              Networks", RFC 7348, DOI 10.17487/RFC7348, August 2014,
              <https://www.rfc-editor.org/info/rfc7348>.

   [RFC8060] Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical
              Address Format (LCAF)", RFC 8060, DOI 10.17487/RFC8060,
              February 2017, <https://www.rfc-editor.org/info/rfc8060>.

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

   [RFC8111] Fuller, V., Lewis, D., Ermagan, V., Jain, A., and A.
              Smirnov, "Locator/ID Separation Protocol Delegated
              Database Tree (LISP-DDT)", RFC 8111, DOI 10.17487/RFC8111,
              May 2017, <https://www.rfc-editor.org/info/rfc8111>.

   [RFC8378] Moreno, V. and D. Farinacci, "Signal-Free Locator/ID
              Separation Protocol (LISP) Multicast", RFC 8378,
              DOI 10.17487/RFC8378, May 2018,
              <https://www.rfc-editor.org/info/rfc8378>.

Authors' Addresses

   Victor Moreno
   Google LLC
   1600 Amphitheater Parkway
   Mountain View, CA 94043
   USA

   Email: vimoreno@google.com

   Dino Farinacci

         lispers.net
         San Jose, CA 95120
         USA

         Email: farinacci@gmail.com

         Alberto Rodriguez-Natal
         Cisco Systems
         170 Tasman Drive
         San Jose, California 95134
         USA

         Email: natal@cisco.com

         Marc Portoles-Comeras
         Cisco Systems
         170 Tasman Drive
         San Jose, California 95134
         USA

         Email: mportole@cisco.com

         Fabio Maino
         Cisco Systems
         170 Tasman Drive
         San Jose, California 95134
         USA

         Email: fmaino@cisco.com

         Sanjay Hooda
         Cisco Systems
         170 Tasman Drive
         San Jose, California 95134
         USA

Email: shooda@cisco.com