Ground-Based LISP for the Aeronautical Telecommunications Network
draft-haindl-lisp-gb-atn-07

Abstract

This document describes the use of the LISP architecture and protocols to address the requirements of the worldwide Aeronautical Telecommunications Network with Internet Protocol Services, as articulated by the International Civil Aviation Organization.

The ground-based LISP overlay provides mobility and multi-homing services to the IPv6 networks hosted on commercial aircrafts, to support Air Traffic Management communications with Air Traffic Controllers and Air Operation Controllers. The proposed architecture doesn’t require support for LISP protocol in the airborne routers, and can be easily deployed over existing ground infrastructures.

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 23 September 2022.

Copyright Notice

Copyright (c) 2022 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. Definition of Terms ........................................... 3
3. Design Overview ............................................. 4
4. Basic Protocol Operation ...................................... 7
   4.1. Endsystem Registration .................................. 7
   4.2. Ground to Airborne Traffic Flow ........................... 8
   4.3. Airborne to Ground Traffic Flow ........................... 8
   4.4. Default forwarding path .................................. 9
   4.5. Traffic symmetry ........................................ 10
5. Multi-Homing and Mobility ..................................... 10
6. Convergence .................................................. 11
   6.1. Use of RLOC-probing ..................................... 12
   6.2. Use of Solicit-Map-Request ............................... 12
   6.3. Use of LISP pub-sub ..................................... 12
7. Multi-domain structure of the ATN/IPS ........................ 13
8. Security Considerations ....................................... 13
   8.1. LISP Basic Security Mechanisms ............................ 13
   8.2. Control Plane overload protection .......................... 14
   8.3. Protecting the LISP control plane from overclaim attacks .................................. 14
   8.4. LISP Reliable Transport .................................. 14
   8.5. Reachability Control ..................................... 16
   8.6. Data Plane Security ...................................... 17
   8.6.1. Segmentation ......................................... 17
   8.6.2. Automated RLOC Filtering ............................... 17

1. Introduction

This document describes the use of the LISP [RFC6830] architecture and protocols to address the requirements of the worldwide Aeronautical Telecommunications Network with Internet Protocol Services (ATN/IPS), as articulated by the International Civil Aviation Organization (ICAO).

ICAO is proposing to replace the existing aeronautical communication services with an IPv6 based infrastructure that supports Air Traffic Management (ATM) between commercial aircrafts, Air Traffic Controllers (ATC) and Air Operation Controllers (AOC).

This document describes how a LISP overlay can be used to offer mobility and multi-homing services to the IPv6 networks hosted on commercial aircrafts without requiring LISP support in the airborne routers. Use of the LISP protocol is limited to the ground-based routers, hence the name "ground-based LISP". The material for this document is derived from [GBL].

2. Definition of Terms

AOC: Airline Operational Control

ATN/IPS: Aeronautical Telecommunications Network with Internet Protocol Services

AC-R: Access Ground Router

A/G-R: Air/Ground Router

G/G-R: Ground/Ground Router

A-R: Airborne Router

A-E: Airborne Endsystem

ATS-E: ATS Endsystem
For definitions of other terms, notably Map-Register, Map-Request, Map-Reply, Routing Locator (RLOC), Solicit-Map-Request (SMR), Ingress Tunnel Router (ITR), Egress Tunnel Router (ETR), xTR (ITR or ETR), Map-Server (MS), and Map-Resolver (MR) please consult the LISP specification [RFC6830].

3. Design Overview

In the ATN/IPS architecture the airborne endsystems hosted on an aircraft are part of an IPV6 network connected to the ground network by one or more Airborne Routers (A-R). A-Rs have multiple radio interfaces that connect them via various radios infrastructures (e.g. SATCOM, LDACS, AeroMACS) to a given radio region, also known as subnetwork, on the ground. Typically an A-R has a corresponding ground based Access Router (AC-R) that terminates the radio protocol with the A-R and provides access services to the ground based portion of the radio network infrastructure. Each radio region is interconnected with the ATN/IPS ground network via an Air-to-Ground router (AG-R).

Similarly, the Air Traffic Controllers and Air Operation Controllers Endsystems (ATS-E and AOC-E) are part of IPv6 networks reachable via one or more Ground-to-Ground Routers (G/G-Rs).

The ATN/IPS ground network infrastructure is the internetworking region located between the A/G routers and the G/G routers.

In the ground-based LISP architecture, a LISP overlay is laid over the ATN/IPS internetworking region (that is in the LISP RLOC space) and provides connectivity between endsystems (that are in the LISP EID space) hosted in the aircrafts and in the AOC/ATS regions. The A/G-Rs and the G/G-Rs assume the role of LISP xTRs supported by a LISP mapping system infrastructure.
Figure 1: ATN/IPS and ground-based LISP overlay
Endsystems in the AOC/ATS regions are mapped in the LISP overlay by the G/G-Rs, which are responsible for the registration of the AOC/ATS endsystems to the LISP mapping system. Each G/G-R is basically an xTR which has direct connections only to the terrestrial regions, i.e. no direct connection to the radio regions.

Aircrafts will attach to a specific radio region, via the radio interfaces of the A-Rs. How the radio attachment works is specific to each particular radio infrastructure, and out of the scope of this document, see [GBL].

Typically at the end of the attachment phase, the access router (AC-R) corresponding to the A-R, will announce the reachability of the EID prefixes corresponding to the attached aircraft (the announcement is specific to each particular radio infrastructure, and is out of the scope of this document). A/G-Rs in that particular radio region are responsible to detect those announcements, and, since they act as xTRs, register to the LISP mapping systems the corresponding IPv6 EID prefixes on behalf of the A-R, but with the RLOC of the A/G-R.

The EID prefixes registered by the A/G-Rs are then reachable by any of the AOC/ATS Endsystems that are part of the ground based LISP overlay.

The LISP infrastructure is used to support seamless aircraft mobility from one radio network to another, as well as multi-homing attachment of an aircraft to multiple radio networks with use of LISP weight and priorities to load balance traffic directed toward the aircraft.

The rest of this document provides further details on how ground-based LISP is used to address the requirements of the ATN/IPS use cases. The main design goals are:

* minimize added complexity on the aircraft
  - airborne routers can assume that any ground system is reachable via any A/G router. Static routing policies can be used on board
  - no need for routing/mobility protocols on board. Routing/mobility is managed on the ground ATN/IPS network
  - on-board outgoing link selection can be done with simple static policy
* seamless support for aircraft mobility and multi-homing with minimal traffic overhead on the A/G datalink
* minimize complexity of ground deployment
  - ground-based LISP can be easily deployed over existing ATN/IPS ground infrastructure
  - it is based on COTS solutions
  - can ease IPv4 to IPv6 transition issues

4. Basic Protocol Operation

Figure 1 provides the reference topology for a description of the basic operation. A more detailed description of the basic protocol operation is described in [GBL].

4.1. Endsystem Registration

The following are the steps via which airborne endsystem prefixes are registered with the LISP mapping system:

1. Each Airborne Endsystem (A-E) is assigned an IPv6 address that is the endsystem EID. Each EID includes a Network-ID prefix that comprises (1) an ICAO ID which uniquely identifies the aircraft, and possibly (2) an aircraft network identifier. Airborne devices are grouped in one (and possibly several) IPv6 EID prefixes. As an example an IPv6 EID prefix could be used for all ATC applications located in a safety critical domain of the aircraft network, another IPv6 EID prefix could be used for AOC applications located in a less safety critical domain.

2. After the Airborne Router (A-R) on an aircraft attaches to one radio region, the corresponding Access Router (AC-R) learns the IPv6 EID prefixes belonging to the aircraft. The AC-R also announces reachability of these prefixes in the radio region (subnetwork) e.g. by using an IGP protocol like OSPF. The attachment to a radio includes a preference parameter and a quality parameter, these parameters are used e.g. to calculate the IGP reachability advertisement metric.

3. The Air/Ground Router (A/G-R) in the subnetwork receives the radio region announcements which contain reachability information for the IPv6 EID prefixes corresponding to the Airborne Endsystems. Since each A/G-R is also an xTR, the A/G-R registers the IPv6 EID prefixes with the LISP MS/MR on behalf of the A-R, but with the RLOC of the A/G-R. The included quality parameter (e.g. IGP metric) is converted to a LISP priority, so that a lower quality metric results in a lower LISP priority value.
Ground based endsystems are part of ground subnetworks where the Ground/Ground Router (G/G-R) is an xTR. Each G/G-R therefore registers the prefixes corresponding to the AOC endsystems and ATS endsystems with the LISP mapping system, as specified in [RFC6830].

4.2. Ground to Airborne Traffic Flow

Here is an example of how traffic flows from the ground to the airborne endsystems, when ATS endsystem 1 (ATS-E1) has traffic destined to airborne endsystem 1 (A-E1):

1. The default route in the ATS region takes the traffic to xTR3 which is also a Ground/Ground Router (G/G-R).

2. xTR3 sends a Map-Request message for the address of A-E1 to the LISP mapping system. xTR2 sends a Map-Reply to xTR3 with RLOC set to its address which is reachable from xTR3 via the internetworking region.

3. xTR3 encapsulates the traffic to xTR2 using the RLOC information in the Map-Reply message.

4. xTR2 decapsulates the traffic coming from xTR3. The destination address of the inner packet belongs to A-E1 which has been advertised by the AC-R in the same region. The traffic is therefore forwarded to AC-R2.

5. AC-R2 sends the traffic to the Airborne Router of the aircraft and the A-R sends it to the endsystem.

4.3. Airborne to Ground Traffic Flow

Here is an example of how traffic flows from the airborne endsystems to the ground when airborne endsystem 2 (A-E2) has traffic destined to ATS endsystem 2 (ATS-E2):

1. The default route in the aircraft points to the Airborne Router (A-R). The latter forwards the traffic over the radio link to AC-R2.

2. The default route on AC-R2 points to xTR2 (also an A/G-R), so the traffic is sent from AC-R2 to xTR2.

3. xTR2 sends a Map-Request message for the address of ATS-E2 to the LISP mapping system. xTR3 sends a Map-Reply to xTR2 with RLOC set to its address which is reachable from xTR2 via the internetworking region.
4. xTR2 encapsulates the traffic to xTR3 using the RLOC information in the Map-Reply message.

5. xTR3 decapsulates the traffic coming from xTR2, and forwards it to ATS-E2.

4.4. Default forwarding path

When an xTR is waiting for a Map-Reply for an EID, the xTR does not know how to forward the packets destined to that EID. This means that the first packets for ground-to-air traffic would get dropped until the Map-Reply is received and a map-cache entry is created. However if a device acting as RTR, see [I-D.ermagan-lisp-nat-traversal], has mappings for all EIDs, the xTR could use the RTR as default path for packets which have to be encapsulated. How the RTR gets all the mappings is outside the scope of this document but one example is the use of LISP pub-sub as specified in [I-D.ietf-lisp-pubsub]. Note that the RTR does not have to be a new device, the device which has the MS/MR role can also act as RTR. It is only the RTR which needs to subscribe to all the aircraft EIDs, the XTRs (i.e. the A/G-Rs and G/G-Rs) do not need to subscribe.

RTRs stitch two legs of a communication flow by acting as an ETR for the purposes of the first leg and as an ITR for the purposes of the second leg. As an ITR (second leg), the RTR will follow all standard procedures of an ITR (issue requests, cache mappings, subscribe to EIDs, etc). In the specific case of the first packet drop scenario, the RTR will subscribe to the entire EID space registered in the Mapping System and maintain a complete cache of all relevant destinations. Any changes to the registration state will be published promptly to the RTR using the pub/sub mechanisms. This ITR role can be made redundant by simply having each RTR in the redundancy group subscribe to the Mapping System. From an ETR perspective, the RTR will also follow all standard procedures for an ETR, but rather than registering specific prefixes, the RTRs will (optionally) register themselves as the "First Packet Handlers". The ITRs sending traffic requiring first packet handling will be configured to forward traffic to the First Packet Handlers if there isn't a mapping already cached for the destination.

The ITRs will know who the first packet handlers are by one of two mechanisms:

1. Configuration of the RLOCs of the first packet handlers on the ITR. This configuration would be done by a network management system.
2. Subscription of the ITR to the "First Packet Handler" EID. As First Packet Handler RLOCs are added or removed the subscribing ITRs are updated.

In both cases the resiliency mechanisms for the RLOCs are the same as for any other RLOC: Routing table reachability combined with optional data plane probes can be leveraged to accelerate failover. In the case in which subscriptions to the "First Packet Handler" EID are used, the RTR will also benefit from the updates in the publication to trigger failover processes.

4.5. Traffic symmetry

The requirements for traffic symmetry are still TBD.

5. Multi-Homing and Mobility

Multi-homing support builds on the procedures described in Section 4:

1. The Airborne Router (A-R) on an aircraft attaches to multiple radio regions. As an example, and referring to Figure 1, the A-R attaches to the LDACS and SATCOM regions, via AC-R2 and AC-R1 respectively.

2. Through the preference parameter sent to each region, the A-R has control over which path (i.e. radio region) ground to air traffic flows. For example, A-R would indicate preference of the LDACS region by choosing a better preference value for the LDACS region compared to the preference value sent to the SATCOM region.

3. Both xTR1 and xTR2 register the IPv6 EID prefixes with the LISP mapping system using merge semantic, as specified in section 4.6 of [I-D.ietf-lisp-rfc6833bis]. Since the priority used in the LISP registrations is derived from the preference and quality parameters, xTR2 would use a lower priority value than xTR1. In this way the LISP mapping system will favour xTR2 (A/G-R for the LDACS region) over xTR1 (A/G-R for the SATCOM region), as specified by the preference and quality parameters.

4. Upon registration the LISP MS/MR will send Map-Notify messages to both xTR1 and xTR2, to inform that they have reachability to the aircraft’s IPv6 EID prefixes. Both xTRs are notified because they have both set the merge-request and want-map-notify bits in their respective Map-Register message.

5. Upstream and downstream traffic flows on the same path, i.e. both use the LDACS region.
With mobility, the aircraft could want to switch traffic from one radio link to another. For example while transiting from an area covered by LDACS to an area covered by SATCOM, the aircraft could desire to switch all traffic from LDACS to SATCOM. For air-to-ground traffic, the A-R has complete control over which radio link to use, and will simply select the SATCOM outgoing interface. For ground-to-air traffic:

1. The A-R sends a radio advertisement to AC-R1 indicating a better preference for the SATCOM link.

2. This leads to AC-R1 lowering its quality parameter (e.g. IGP metric) for the IPv6 EID prefixes.

3. Upon receiving the better preference value, xTR1 registers the IPv6 EID prefixes with the MS/MR, using a lower priority value than what xTR2 had used. Both xTR1 and xTR2 receives Map-Notify messages signaling to xTR2 that xTR1 is now the preferred path toward the aircraft.

4. xTR3 has a map-cache which still points to xTR2, therefore xTR3 still sends traffic via xTR2. xTR2 sends Solicit-Map-Request (SMR) to xTR3 who queries the LISP mapping system again. This results in updating the map-cache on xTR3 which now points to XTR1 so ground-to-air traffic now flows on the SATCOM radio link.

The procedure for mobility is derived from [I-D.ietf-lisp-eid-mobility].

6. Convergence

When traffic is flowing on a radio link and that link goes down, the network has to converge rapidly on the other link available for that aircraft.

For air-to-ground traffic, once the A-R detects the failure it can switch immediately to the other radio link.

For ground-to-air traffic, when a radio link fails, the corresponding AC-R sends a reachability update that the IPv6 EID prefixes are not reachable anymore. This leads to the A/G-R (also an xTR) in that region to unregister the IPv6 EID prefixes with the MS/MR. This indicates that the xTR in question has no reachability to the EID prefixes. The notification of the failure should reach all relevant xTRs as soon as possible. For example, if the LDACS radio link fails, xTR3 and xTR4 need to learn about the failure so that they stop sending traffic via xTR2 and use xTR1 instead.
In the sub-sections below, we the use of RLOC-probing, Solicit-Map-Request, and LISP pub-sub as alternative mechanisms for link failure notification.

6.1. Use of RLOC-probing

RLOC-probing is described in section 6.3.2 of [RFC6830].

At regular intervals xTR3 sends Map-Request to xTR2 for the aircraft’s EID prefixes. When xTR3 detects via RLOC-probing that it can not use xTR2 anymore, it sends a Map-Request for the aircraft’s EID prefixes. The corresponding Map-Reply indicates that xTR1 should now be used. The map-cache on xTR3 is updated and air-to-ground traffic now goes through xTR1 to use the SATCOM radio link to the aircraft.

The disadvantage of RLOC-probing is that fast detection becomes more difficult when the number of EID prefixes is large.

6.2. Use of Solicit-Map-Request

Solicit-Map-Request is used as described in Section 5:

1. xTR3 is still sending traffic to xTR2 since its map-cache has not been updated yet.

2. Upon detecting that the link is down, and receiving data plane traffic from the ground network, xTR2 sends an SMR to xTR3 that sends a Map-Request to update its map-cache. The corresponding Map-Reply indicates that xTR1 should now be used.

The disadvantage of this approach is that the traffic is delayed pending control-plane resolution. This method also depends on data traffic being continuous, in many cases data traffic may be sporadic, leading to very slow convergence.

6.3. Use of LISP pub-sub

As specified in [I-D.ietf-lisp-pubsub], ITRs can subscribe to changes in the LISP mapping system. So if all ITRs subscribe to the EID prefixes for which they have traffic, the ITRs will be notified when there is mapping change.

In the example where the LDACS radio link fails, when xTR2 unregisters the EID prefixes with the MS/MR, xTR3 would be notified via LISP pub-sub (assuming xTR3 has a map-cache entry for these EID prefixes).
This mechanism provides the fastest convergence at the cost of more state in the LISP mapping system.

7. Multi-domain structure of the ATN/IPS

The overlay on the ATN/IPS can be structured as a collection of independent administrative domains following the model defined in [I-D.moreno-lisp-uberlay]. In this model, the different administrative domains are interconnected by a transit area referred to as an uberlay. Each administrative domain is independent from the perspective of the control, data and administrative planes. Structuring the ATN/IPS in this manner allows the combination of different implementations and even different mobility methods in the ATN/IPS. The structure proposed also improves resiliency by isolating events and failures across the different administrative domains and improves the scale of the ATN/IPS by distributing the responsibility of maintaining granular aircraft state across the different administrative domains.

The uberlay may be a BGP network as defined in [I-D.templin-atn-bgp]. Following the definitions put forth in [I-D.templin-atn-bgp], the uberlay transit is the core autonomous system and the different administrative domains that conform the ATN/IPS are what [I-D.templin-atn-bgp] defines as stub autonomous systems.

8. Security Considerations

For LISP control-plane message security, please refer to [I-D.ietf-lisp-sec]. This addresses the control-plane threats that target EID-to-RLOC mappings, including manipulations of Map-Request and Map-Reply messages, and malicious ETR EID prefix overclaiming.

8.1. LISP Basic Security Mechanisms

The LISP specification, documented in [RFC6830bis] and [RFC6833bis], includes basic security mechanisms for the control plane. The base mechanisms are designed to prevent rogue unauthorized ETRs from registering mappings into the Mapping System and to protect ITRs from receiving unsolicited mapping information. To authenticate EID-to-RLOC mapping registrations and ensure that they are from an authorized ETR, LISP uses shared secret keys between ETRs and the Mapping System. Only ETRs that have the shared secret key are able to register EID-to-RLOC mappings to the Mapping System. Without the correct key, the authenticity of the map-register message cannot be verified, and the Mapping System must reject the map-register. The shared keys used to authenticate map-registers are distributed across ETRs and MS/MRs by the orchestration/configuration infrastructure.
The shared keys need to be distributed between the xTR and the Mapping System. Since these components will be in the same administrative domain in GB-LISP, it would be feasible to implement a method for this key exchange (see Clause 6.5 in [LISP-SEC]). In addition to authenticating EID registrations, it is recommended that the Mapping System restricts EID registrations to configured EID prefix ranges. Thus, an authorized ETR is allowed to register EID prefixes only within the EID prefix range configured in the Map-Server. The confidentiality of the LISP control plane messages can be ensured by protecting the transport of control messages with DTLS (over UDP) [RFC6347] or LISP-crypto [RFC8061]. DTLS is also proposed in Clause 6.7 of [LISP-SEC] for providing message privacy.

8.2. Control Plane overload protection

Data-plane gleaning [Clause 9 in RFC6830bis] might need to be turned off for avoiding potential attacks by forged data plane packets that could overload the control plane. Another approach is data fusion between multiple reachability verification mechanisms. Generic control plane protection mechanism, such as packet filtering and rate control, should be also deployed for GB-LISP nodes based on a risk assessment. This could mitigate such attacks that try to misuse the Map-Versioning mechanism in the data-plane for overloading the control-plane.

8.3. Protecting the LISP control plane from overclaim attacks

The Internet Draft [LISP-SEC] defines a set of security mechanisms (usually referred as LISP-SEC) to provide origin authentication, integrity, and antireplay protection to the EID-to-RLOC mapping data conveyed in the map-resolution process. It includes the usage of multiple one-time-keys (OTK) and hash based message authentication. LISP-SEC also enables authorization verification on EID-prefixes claims made by ETRs, preventing so-called "overclaiming attacks" in which an ETR attempts to claim EID-prefixes for which it is not authoritative. A LISP-SEC protected map-reply, in fact, includes metadata authenticated by the map-server that specify which

8.4. LISP Reliable Transport

The communication with the Mappin Systems is originally proposed based on UDP that is not a reliable transport. For a proper synchronization between the ETR and the Map-Servers periodic message transmission would be needed. Usually, Map-Register messages are retransmitted every minute by the ETR. The Map-Server removes the EID entries if they are not refreshed for three successive periods.
In mobility solutions, typically a large number of EID entries needs to be registered. Because of packet size limitations these entries can be transported only by a significant number of Map-Register messages in each period. A new reliable transport option has been defined in [LISP-RELT] to solve these issues. Although this Internet Draft has been expired, the new method is used in the latest widely deployed LISP solution for Software Defined Access (SDA) by Cisco Systems. The reliable transport is composed by new message formats and the support for other then UDP as a transport in the control plane. Both TCP and SCTP is addressed by the specification. The TCP implementation could be traced in the labs. The messages are based on a TLV format where a type filed support the future extensions of the protocol. A message end marker provides extra integrity check possibility for complex aggregated messages. Error notification messages are also specified for notifying situations when the receiver does not recognize or cannot parse message contents. The following message types are specified for the reliable transport mechanism: o Map-Register, o Registration acknowledgement, o Registration rejection, o Registration refresh, o Mapping notification, The session establishment has to be backward compatible. The Map Server authenticates the ETR first using UDP based messages. Once the ETR is authenticated, the Map Server performs a passive open by listening on TCP port 4342. TCP connections are accepted only from the already authenticated ETRs. The ETR has to open the TCP connection actively towards the Map Server one it has received the Map-Notify message on the UDP transport. If the TCP session goes down, the same UDP based procedure has to be repeated. The Map-Server will also revert to the expiration mechanism used for UDP transport until the TCP based session would be fully restored. A single TCP session is built up for all subsequent control plane messages. This applies even when multiple address families are used in the EID space. Once the reliable transport can be used, the periodic refresh is not needed anymore. Mapping information is sent only when there is new information to share. Time-out based removal of registrations are not used in this case. An explicit de-registration is needed by carrying a zero TTL. The reliable transport session should be authenticated. In the simpler case, it could be an RLOC spoofing mitigation. If this is not reliable, then the TCP Authentication Option [RFC5925], or the SCTP Authenticated Chunks [RFC4895] are recommended.
8.5. Reachability Control

The communication with the Mappin Systems is originally proposed based on UDP that is not a reliable transport. For a proper synchronization between the ETR and the Map-Servers periodic message transmission would be needed. Usually, Map-Register messages are retransmitted every minute by the ETR. The Map-Server removes the EID entries if they are not refreshed for three successive periods. In mobility solutions, typically a large number of EID entries needs to be registered. Because of packet size limitations these entries can be transported only by a significant number of Map-Register messages in each period. A new reliable transport option has been defined in [LISP-RELT] to solve these issues. Although this Internet Draft has been expired, the new method is used in the latest widely deployed LISP solution for Software Defined Access (SDA) by Cisco Systems. The reliable transport is composed by new message formats and the support for other then UDP as a transport in the control plane. Both TCP and SCTP is addressed by the specification. The TCP implementation could be traced in the labs. The messages are based on a TLV format where a type filed support the future extensions of the protocol. A message end marker provides extra integrity check possibility for complex aggregated messages. Error notification messages are also specified for notifying situations when the receiver does not recognize or cannot parse message contents. The following message types are specified for the reliable transport mechanism: o Map-Register, o Registration acknowledgement, o Registration rejection, o Registration refresh, o Mapping notification, The session establishment has to be backward compatible. The Map Server authenticates the ETR first using UDP based messages. Once the ETR is authenticated, the Map Server performs a passive open by listening on TCP port 4342. TCP connections are accepted only from the already authenticated ETRs. The ETR has to open the TCP connection actively towards the Map Server one it has received the Map-Notify message on the UDP transport. If the TCP session goes down, the same UDP based procedure has to be repeated. The Map-Server will also revert to the expiration mechanism used for UDP transport until the TCP based session would be fully restored. A single TCP session is built up for all subsequent control plane messages. This applies even when multiple address families are used in the EID space. Once the reliable transport can be used, the periodic refresh is not needed anymore. Mapping information is sent only when there is new information to share. Time-out based removal of registrations are not used in this case. An explicit de-registration is needed by carrying a zero TTL. The reliable transport session should be authenticated. In the simpler case, it could be an RLOC spoofing mitigation. If this is not reliable, then the TCP Authentication Option [RFC5925], or the SCTP Authenticated Chunks [RFC4895] are
recommended.

8.6. Data Plane Security

8.6.1. Segmentation

LISP inherently delivers segmentation by using extended endpoint identifiers (EIDs) and Instance-IDs to partition the EID space, segment the map-caches, and color the control and data plane messages to create virtual networks. These virtual networks are a seamless extension of the way EIDs are normally handled in LISP and therefore enjoy all the benefits of scale, mobility, and address family independence that LISP provides.

8.6.2. Automated RLOC Filtering

The communication on between the xTRs and Map-Servers use the RLOC space data plane. Only those communications attempts shall be accepted that are coming from valid RLOC addresses. Manual configuration of such access lists would be too difficult to manage. An automated RLOC membership mechanism is proposed in [LISP-RFIL]. Although this Internet Draft has been expired, it is still included in some LISP implementations. The Map-Server can authenticate each xTR that wants to communicate. It will build up a list of xTRs that are valid members of this LISP administrative domain. An xTR can specifically subscribe to this membership information. Membership can be maintained by address family and instance ID (VPN). This allows an easy management of both RLOC and EID space segmentation by VPNs. It also supports gateway functions between separated RLOC spaces. Only valid xTR members can apply for notifications of membership information. The xTR receiving the membership information might use it for building internal access control lists automatically. Proxy xTR information is not included in the membership list, so communication with such nodes need to be configured manually. A membership message format is defined in [LISP-RFIL]. The following message type are specified: o Membership subscribe, o Membership subscribe acknowledgement, o Membership subscribe negative acknowledgement, o Membership unsubscribe, o Membership element add, o Membership element delete, o Membership refresh request, o Membership refresh begin, o Membership refresh end. The membership information could be used by the xTR for other future functions, too. Automated RLOC filtering is just one example.
8.6.3. Confidentiality, Integrity and Anti-replay protection

In those sections of the ATN/IPS network where data plane confidentiality, integrity and anti-replay protection may be required, the LISP data plane can be secured as any other IP traffic by leveraging IPsec. The provisioning of an IPsec VPN to secure IP encapsulated LISP frames is orthogonal to deployment of LISP and can be done using well known IPsec key negotiation mechanisms such as IKEv2 [RFC7296].

IKEv2 uses X.509 certificates for authentication. A PKI is needed for managing the certificates. The certificates are used for generating the exchanged symmetric encryption keys.

9. IANA Considerations

No IANA considerations.

10. Acknowledgements

The original authors would like to thank Dino Farinacci and Bela Varkonyi for their review of the document and deep insights.

The following people have contributed, over time, to the authorship of this document: Bernhard Haindl, Manfred Lindner, Reshad Rahman, Marc Portoles-Comeras, Victor Moreno, Fabio Maino, Balaji Venkatachalapathy.

11. References

11.1. Normative References


11.2. Informative References


Authors’ Addresses

Bernhard Haindl
Frequentis
Email: bernhard.haindl@frequentis.com

Manfred Lindner
Frequentis
Email: manfred.lindner@frequentis.com

Victor Moreno
Google
Email: vimoreno@google.com

Marc Portoles Comeras
Cisco Systems
Email: mportole@cisco.com

Fabio Maino
Cisco Systems
Email: fmaino@cisco.com

Balaji Venkatachalapathy
Cisco Systems
Email: bvenkata@cisco.com
LISP L2/L3 EID Mobility Using a Unified Control Plane

draft-ietf-lisp-eid-mobility-10

Abstract

The LISP control plane offers the flexibility to support multiple overlay flavors simultaneously. This document specifies how LISP can be used to provide control-plane support to deploy a unified L2 and L3 overlay solution for End-point Identifier (EID) mobility, as well as analyzing possible deployment options and models.

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 11 January 2023.
Table of Contents

1. Introduction .................................................. 3
2. Definition of Terms ........................................... 4
3. Reference System ............................................. 4
4. L3 Overlays and Mobility Support .............................. 5
   4.1. Reference Architecture and packet flows ................. 5
       4.1.1. Routed Traffic Flow: L3 Overlay use .............. 6
       4.1.2. L3 Mobility Flow .................................. 6
           4.1.2.1. L3 Mobility Flow using Data Driven SMRs .. 7
           4.1.2.2. L3 Mobility Flow using Publish/Subscribe   
               Mechanisms ..................................... 8
   4.2. Implementation Considerations .......................... 9
       4.2.1. L3 Segmentation .................................. 9
       4.2.2. L3 Database-Mappings ............................. 9
       4.2.3. LISP Mapping System support ........................ 10
       4.2.4. Using SMRs to Track Moved-Away Hosts ........... 10
       4.2.5. L3 multicast support ............................. 11
       4.2.6. Time-to-Live Handling in Data-Plane .............. 11
5. L2 Overlays and Mobility Support ............................ 11
   5.1. Reference Architecture and packet flows ................. 11
       5.1.1. Bridged Traffic Flow: L2 Overlay use ............ 12
       5.1.2. L2 Mobility Flow ................................ 13
           5.1.2.1. L2 Mobility Flow using Data Driven SMRs .. 13
           5.1.2.2. L2 Mobility Flow using Publish/Subscribe   
               mechanisms ...................................... 14
   5.2. Implementation Considerations .......................... 14
       5.2.1. L2 Segmentation .................................. 15
       5.2.2. L2 Database-Mappings ............................. 15
       5.2.3. Interface to the LISP Mapping System .............. 16
       5.2.4. SMR support to track L2 hosts that moved away .... 16
       5.2.5. L2 Broadcast and Multicast traffic ................ 17
       5.2.6. L2 Unknown Unicast Support ........................ 17
       5.2.7. Time-to-Live Handling in Data-Plane .............. 18
1. Introduction

This document describes the architecture and design options required to offer a unified L2 and L3 overlay solution for End-point Identifier (EID) mobility using the LISP control-plane.

The architecture takes advantage of the flexibility that LISP provides to simultaneously support different overlay types. While the LISP specification defines both the data-plane and the control-plane, this document focuses on the use of the control-plane to provide L2 and L3 overlay services with EID mobility. The control plane may be combined with a data-plane of choice e.g., [LISP], [VXLAN-GPE], or [VXLAN].

The recommendation on whether a flow is sent over the L2 or the L3 overlay is based on whether the traffic is bridged (intra-subnet or non-IP) or routed (inter-subnet), respectively. This allows treating both overlays as separate segments, and enables L2-only and L3-only deployments (and combinations of them) without modifying the architecture.

The unified solution for L2 and L3 overlays offers the possibility to extend subnets and routing domains (as required in state-of-art Datacenter and Enterprise architectures) with mobility support and traffic optimization.

An important use-case of the unified architecture is that, while most data centers are complete layer-3 routing domains, legacy applications either have not converted to IP or still use auto-discovery at layer-2 and assume all nodes in an application cluster belong to the same subnet. For these applications, the L2-overlay limits the functionality to where the legacy app lives versus having to extend layer-2 into the underlay network.
Broadcast, Unknown and Multicast traffic on the overlay are supported by either replicated unicast, or underlay (RLOC) multicast as specified in [RFC6831] and [RFC8378].

2. Definition of Terms

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

3. Reference System

The following figure illustrates the reference system used to support the packet flow description throughout this document. The system presents 4 sites. Site A and Site D provide access to different subnets (non-extended), while Site B and Site C extend a common subnet. The xTR in each one of the sites registers EIDs from the sites with the LISP Mapping System and provides support to encapsulate overlay (EID) traffic through the underlay (RLOC space).
The recommended selection between the use of L2 and L3 overlays is to map them to bridged (intra-subnet or non-IP) and routed (inter-subnet) traffic. The rest of the document follows this recommendation to describe the packet flows.

However, note that in a different selection approach, intra-subnet traffic MAY also be sent over the L3 overlay. Section 6.1 specifies the changes needed to send all IP traffic using the L3 overlay and restricting the use of the L2 overlay to non-IP traffic.

When required, the control plane makes use of two basic types of EID-to-RLOC mappings associated to end-hosts and in order to support the unified architecture:

* EID = <IID, MAC> to RLOC=<IP>. This is used to support the L2 overlay.

* EID = <IID, IP> to RLOC=<IP>. This is the traditional mapping as defined in the original LISP specification and supports the L3 overlay.

4. L3 Overlays and Mobility Support

4.1. Reference Architecture and packet flows

In order to support the packet flow descriptions in this section we use Figure 1 as reference. This section uses Sites A and D to describe the flows.

```
|     |     |     |     |
| RLOC=IP_A | RLOC=IP_B | RLOC=IP_C | RLOC=IP_D |
| ++++++++ | ++++++++ | ++++++++ | ++++++++ |
| . | xTR A | . | xTR B | . | xTR C | . | xTR D |
| ( ++++++ ) | ( ++++++ ) | ( ++++++ ) | ( ++++++ ) |
| Site A | Site B | Site C | Site D |
| ( 1.0.0.0/24 | 3.0.0.0/24 | 3.0.0.0/24 | 2.0.0.0/24 |
| '---'.'.' | '---'.'.' | '---'.'.' | '---'.'.' |
|   |   |   |   |
| '--------' | '--------' | '--------' | '--------' |
| '--------' | '--------' | '--------' | '--------' |
| (IID1,1.0.0.1) | (IID1,3.0.0.2) | (IID1,3.0.0.3) | (IID1,2.0.0.4) |
| (IID2,0:0:3:0:0:2) | (IID2,0:0:3:0:0:3) |
```

4.1.1. Routed Traffic Flow: L3 Overlay use

Inter-subnet traffic is encapsulated using the L3 overlay. The process to encapsulate this traffic is the same as described in [I-D.ietf-lisp-rfc6830bis] and [I-D.ietf-lisp-rfc6833bis]. We describe the packet flow here for completeness.

The following is a sequence example of the unicast packet flow and the control plane operations when in the topology shown in Figure 1 End-Device 1, in LISP site A, wants to communicate with End-Device 4 in LISP site D. Note that both end systems reside in different subnets. We'll assume that End-Device 1 knows the EID IP address of End-Device 4 (e.g. it is learned using a DNS service).

* End-Device 1 sends an IP packet frame with destination 2.0.0.4 and source 1.0.0.1. As the destination address lies on a different subnet End-Device 1 sends the packet following its routing table to ITR A (e.g., it is its default gateway).

* ITR A does a L3 lookup in its local map-cache for the destination IP 2.0.0.4. When the lookup of 2.0.0.4 is a miss, the ITR sends a Map-request to the mapping database system looking up for EID=<IID1,2.0.0.4>.

* The mapping systems forwards the Map-Request to ETR D, that has registered the EID-to-RLOC mapping of EID=<IID1,2.0.0.4>.

* ITR A populates the local map-cache with the EID to RLOC mapping, and encapsulates all subsequent packets with a destination IP 2.0.0.4 using destination RLOC=IP_D.

4.1.2. L3 Mobility Flow

The support to L3 mobility covers the requirements to allow an end-host to move from a given site to another and maintain correspondence with the rest of end-hosts that are connected to the same L3 routing domain. This support MUST ensure convergence of L3 forwarding (IPv4/IPv6 based) from the old location to the new one when the host completes its move.
The update of the ITR map-caches when EIDs move MAY be achieved using Data Driven SMRs or the Publish/Subscribe mechanisms defined in [I-D.ietf-lisp-pubsub]. The following two sections are sequence descriptions of the packet flow when End-Device 1 in the reference figure roams to site D.

4.1.2.1. L3 Mobility Flow using Data Driven SMRs

The following is a sequence description of the packet flow when End-Device 1 in the reference figure roams to site D. This sequence uses Data Driven SMRs to trigger the updates of the ITR map-caches.

* When End-Device 1 is attached or detected in site D, ETR D sets up the database mapping corresponding to EID=<IID1, 1.0.0.1>. ETR D sends a Map-Register to the mapping system registering RLOC=IP_D as locator for EID=<IID1, 1.0.0.1>. Now the mapping system is updated with the new EID-to-RLOC mapping (location) for End-Device 1.

* The Mapping System, after receiving the new registration for EID=<IID1, 1.0.0.1> sends a Map-Notify to the departure ETR(s) (ETR A) to inform it of the move. Then, ETR A removes its local database mapping information and stops registering EID=<IID1, 1.0.0.1>.

* Any ITR or PiTR participating in the L3 overlay (corresponding to IID1) that were sending traffic to 1.0.0.1 before the migration keep sending traffic to ETR A.

* Once ETR A is notified by the Mapping system, when it receives traffic from an ITR with destination 1.0.0.1, it generates a Solicit-Map-Request (SMR) back to the ITR (or PiTR) for EID=<IID1, 1.0.0.1>.

* Upon receiving the SMR the ITR invalidates its local map-cache entry for EID=<IID1, 1.0.0.1> and sends a new Map-Request for that EID. The Map-Reply includes the new EID-to-RLOC mapping for End-Device 1 with RLOC=IP_D.

* Similarly, once the local database mapping is removed from ITR A, non-encapsulated packets arriving at ITR A from a local End-Device and destined to End-Device 1 result in a cache miss, which triggers sending a Map-Request for EID=<IID1, 1.0.0.1> to populate the map-cache of ITR A.
4.1.2.2. L3 Mobility Flow using Publish/Subscribe Mechanisms

When Publish/Subscribe ([I-D.ietf-lisp-pubsub]) mechanisms are used, the flow of signaling to achieve EID mobility is modified. In this case, when a local end-device connected via an ITR establishes communication with a remote mobile end-device (connected to a remote ETR), the ITR will issue a Map-Request for the mobile end-device. Following the Mapping Request Subscribe Procedures defined in [I-D.ietf-lisp-pubsub], the Map-request will be issued with the N-bit set on the EID-Record so that the ITR is notified of any RLOC-set changes for the mobile EID-prefix.

The following is a sequence description of the packet flow when End-Device 1 in the reference figure roams to site D. This sequence leverages Publish/Subscribe mechanisms to update the ITR map-caches.

* When an end-Device connected via an ITR establishes communication with a mobile end-device (e.g. end-device 1), the ITR will issue a Map-Request for the mobile end-device. Following the Mapping Request Subscribe Procedures defined in [I-D.ietf-lisp-pubsub], the Map-request will be issued with the N-bit set on the EID-Record so that the ITR is notified of any RLOC-set changes for the mobile EID-prefix.

* When the mobile end-device (End-Device 1) is attached or detected in a new site (e.g. site D), The ETR at the new site (e.g. ETR D) sets up the database mapping corresponding to the EID of the mobile end-device (e.g. EID=<IID1, 1.0.0.1>). The ETR at the new site (e.g. ETR D) sends a Map-Register to the mapping system registering its RLOCs (e.g. RLOC=IP_D) as locator for the EID of the mobile end-device (e.g. EID=<IID1, 1.0.0.1>). Now the mapping system is updated with the new EID-to-RLOC mapping (location) for the mobile end-device (e.g. End-Device 1).

* The Mapping System, after receiving the new registration for the EID of the mobile end-device (EID=<IID1, 1.0.0.1>) sends a Map-Notify to the departure site (ETR A) to inform it of the move. Then, the ETR at the departure site (ETR A) removes its local database mapping information and stops registering the EID for the mobile end-device (EID=<IID1, 1.0.0.1>).

* Any ITR or PiTR participating in the L3 overlay (corresponding to IID1) that were sending traffic to the mobile end-device (1.0.0.1) would have Subscribed to receive notifications of any changes in the RLOC-set for the EID of the mobile end-device (1.0.0.1). The Mapping System publishes the updated RLOC-set to the Subscribed ITRs by sending a Map-Notify to the ITRs or PiTRs per the Mapping Notification Publish Procedures defined in [I-D.ietf-lisp-pubsub].
Upon receiving the Map-Notify message, the ITR updates the RLOC-set in its local map-cache entry for the EID of the mobile end-device (EID=<IID1, 1.0.0.1>). Once the map-cache is updated, traffic is tunneled by the ITR to the new location.

4.2. Implementation Considerations

4.2.1. L3 Segmentation

LISP support of segmentation and multi-tenancy is structured around the propagation and use of Instance-IDs, and handled as part of the EID in control plane operations. The encoding is described in [RFC8060] and its use in [RFC8111].

Instance-IDs can be used to support L3 overlay segmentation, such as in extended VRFs or multi-VPN overlays ([I-D.ietf-lisp-vpn]).

4.2.2. L3 Database-Mappings

When an end-host is attached or detected in an ETR that provides L3-overlay services and mobility, a database Mapping is registered to the mapping system with the following structure:

* The EID 2-tuple (IID, IP) with its binding to a corresponding ETR locator set (IP RLOC)

The registration of these EIDs MUST follow the LCAF format as defined in [RFC8060] and the specific EID record to be used is illustrated in the following figure:
The L3 EID record follows the structure as described in [I-D.ietf-lisp-rfc6833bis].

4.2.3. LISP Mapping System support

The interface between the xTRs and the Mapping System is described in [I-D.ietf-lisp-rfc6833bis] and this document follows the specification as described there. When available, the registrations MAY be implemented over a reliable transport as described in [I-D.kouvelas-lisp-map-server-reliable-transport].

In order to support system convergence after mobility, when the Map-Server receives a new registration for a specific EID, it MUST send a Map-Notify to the entire RLOC set in the site that last registered this same EID. This Map-Notify is used to track moved-away state of L3 EIDs as described in Section 4.2.4.

4.2.4. Using SMRs to Track Moved-Away Hosts

One of the key elements to support end-host mobility using the LISP architecture is the Solicit-Map-Request (SMR). This is a special message by means of which an ETR can request an ITR to send a new Map-Request for a particular EID record. In essence the SMR message is used as a signal to indicate a change in mapping information and it is described in [I-D.ietf-lisp-rfc6833bis].
In order to support mobility, an ETR SHALL maintain a list of EID records for which it has to generate a SMR message whenever it receives traffic with that EID as destination.

The particular strategy to maintain an Away Table is implementation specific and it will be typically based on the strategy to detect the presence of hosts and the use of Map-Notify messages received from the Map-Server. In essence the table SHOULD provide support to the following:

* Keep track of end-hosts that were once connected to an ETR and have moved away.

* Support for L3 EID records, the 2-tuple (IID, IP), for which a SMR message SHOULD be generated.

4.2.5. L3 multicast support

L3 Multicast traffic on the overlay MAY be supported by either replicated unicast, or underlay (RLOC) multicast. Specific solutions to support L3 multicast over LISP controlled overlays are specified in in [RFC6831], and [RFC8378].

4.2.6. Time-to-Live Handling in Data-Plane

The LISP specification ([I-D.ietf-lisp-rfc6830bis]) describes how to handle Time-to-Live values of the inner and outer headers during encapsulation and decapsulation of packets when using the L3 overlay.

5. L2 Overlays and Mobility Support

5.1. Reference Architecture and packet flows

In order to support L2 packet flow descriptions in this section we use Figure 1 as reference. This section uses Sites B and C to describe the flows.
5.1.1. Bridged Traffic Flow: L2 Overlay use

Bridged traffic is encapsulated using the L2 overlay. This section provides an example of the unicast packet flow and the control plane operations when in the topology shown in Figure 1, the End-Device 2 in site B communicates with the End-Device 3 in site C. In this case we assume that End Device 2, knows the MAC address of End-Device 3 (e.g., learned through ARP).

* End-Device 2 sends an Ethernet/IEEE 802 MAC frame with destination 0:0:3:0:0:3 and source 0:0:3:0:0:2.

* ITR B does a L2 lookup in its local map-cache for the destination MAC 0:0:3:0:0:3. When the lookup of 0:0:3:0:0:3 is a miss, the ITR sends a Map-Request to the mapping database system looking up for EID=<IID2,0:0:3:0:0:3>.

* The mapping systems forwards the Map-Request to ETR C, that has registered the EID-to-RLOC mapping for EID=<IID2,0:0:3:0:0:3>. Alternatively, depending on the mapping system configuration, a Map-Server which is part of the mapping database system MAY send a Map-Reply directly to ITR B.

* ETR C sends a Map-Reply to ITR B that includes the EID-to-RLOC mapping: EID=<IID2, 0:0:3:0:0:3> -> RLOC=IP_C, where IP_C is the locator of ETR C.
* ITR B populates the local map-cache with the EID to RLOC mapping, and encapsulates all subsequent packets with a destination MAC 0:0:3:0:0:3 using destination RLOC=IP_C.

5.1.2. L2 Mobility Flow

The support to L2 mobility covers the requirements to allow an end-host to move from a given site to another and maintain correspondence with the rest of end-hosts that are connected to the same L2 domain (e.g. extended VLAN). This support MUST ensure convergence of L2 forwarding (MAC based) from the old location to the new one, when the host completes its move.

The update of the ITR map-caches when EIDs move MAY be achieved using Data Driven SMRs or the Publish/Subscribe mechanisms defined in [I-D.ietf-lisp-pubsub]. The following two sections are sequence descriptions of the packet flow when End-Device 2 in the reference figure roams to site C, which is extending its own subnet network.

5.1.2.1. L2 Mobility Flow using Data Driven SMRs

The following is a sequence description of the packet flow when End-Device 2 in the reference figure roams to site C. This sequence uses Data Driven SMRs to trigger the updates of the ITR map-caches.

* When End-Device 2 is attached or detected in site C, ETR C sets up the database mapping corresponding to EID=<IID2, 0:0:3:0:0:2>. ETR C sends a Map-Register to the mapping system registering RLOC=IP_B as locator for EID=<IID2, 0:0:3:0:0:2>.

* The Mapping System, after receiving the new registration for EID=<IID1, 0:0:3:0:0:2> sends a Map-Notify to ETR B with the new locator set (IP_B). ETR B removes then its local database mapping and stops registering <IID2, 0:0:3:0:0:2>.

* Any PiTR or ITR participating in the same L2-overlay (corresponding to IID2) that was encapsulating traffic to 0:0:3:0:0:2 before the migration continues encapsulating this traffic to ETR B.

* Once ETR B is notified by the Mapping system, when it receives traffic from an ITR which is destined to 0:0:3:0:0:2, it will generate a Solicit-Map-Request (SMR) message that is sent to the ITR for (IID2,0:0:3:0:0:2).

* Upon receiving the SMR the ITR sends a new Map-Request for the EID=<IID2,0:0:3:0:0:2>. As a response ETR B sends a Map-Reply that includes the new EID-to-RLOC mapping for <IID2,0:0:3:0:0:2>
5.1.2.2. L2 Mobility Flow using Publish/Subscribe mechanisms

When Publish/Subscribe ([I-D.ietf-lisp-pubsub]) mechanisms are used, the flow of signaling to achieve EID mobility is modified. In this case, when an End-Device connected via an ITR establishes communication with a mobile EID-prefix, the ITR will issue a Map-Request for the mobile End-device. Following the Mapping Request Subscribe Procedures defined in [I-D.ietf-lisp-pubsub], the Map-request will be issued with the N-bit set on the EID-Record so that the ITR is notified of any RLOC-set changes for the mobile EID-prefix.

The following is a sequence description of the packet flow when End-Device 2 in the reference figure roams to site C. This sequence leverages Publish/Subscribe mechanisms to update the ITR map-caches.

* When End-Device 2 is attached or detected in site C, ETR C sets up the database mapping corresponding to EID=<IID2, 0:0:3:0:0:2>. ETR C sends a Map-Register to the mapping system registering RLOC=IP_B as locator for EID=<IID2, 0:0:3:0:0:2>.

* The Mapping System, after receiving the new registration for EID=<IID1, 0:0:3:0:0:2> sends a Map-Notify to the departure site (ETR B) with the new locator set (IP_B). ETR B removes then its local database mapping and stops registering <IID2, 0:0:3:0:0:2>.

* Any ITR or PiTR participating in the same L2-overlay (corresponding to IID2) that was encapsulating traffic to 0:0:3:0:0:2 before the migration would have subscribed to receive notifications of any changes in the RLOC-set for 0:0:3:0:0:2. The Mapping System publishes the updated RLOC-set to the Subscribed ITRs by sending a Map-Notify to the ITRs or PiTRs per the Mapping Notification Publish Procedures defined in [I-D.ietf-lisp-pubsub].

* Upon receiving the Map-Notify message, the ITR updates the RLOC-set in its local map-cache entry for EID=<IID2, 0:0:3:0:0:2>. Once the map-cache is updated, traffic is tunneled by the ITR to the new location.

5.2. Implementation Considerations
5.2.1. L2 Segmentation

As with L3 overlays, LISP support of L2 segmentation is structured around the propagation and use of Instance-IDs, and handled as part of the EID in control plane operations. The encoding is described in [RFC8060] and its use in [RFC8111]. Instance-IDs are unique to a Mapping System and MAY be used to identify the overlay type (e.g., L2 or L3 overlay).

An Instance-ID can be used for L2 overlay segmentation. An important aspect of L2 segmentation is the mapping of VLANs to IIDs. In this case a Bridge Domain (which is the L2 equivalent to a VRF as a forwarding context) maps to an IID, a VLAN-ID may map 1:1 to a bridge domain or different VLAN-IDs on different ports may map to a common Bridge Domain, which in turn maps to an IID in the L2 overlay. When ethernet traffic is double tagged, usually the outer 802.1Q tag will be mapped to a bridge domain on a per port basis, and the inner 802.1Q tag will remain part of the payload to be handled by the overlay. The IID should therefore be able to carry ethernet traffic with or without an 802.1Q header. A port may also be configured as a trunk and we may chose to take the encapsulated traffic and map it to a single IID in order to multiplex traffic from multiple VLANs on a single IID. These are all examples of local operations that could be effected on VLANs in order to map them to IIDs, they are provided as examples and are not exhaustive.

5.2.2. L2 Database-Mappings

When an end-host is attached or detected in an ETR that provides L2-overlay services, a database Mapping is registered to the mapping system with the following structure:

* The EID 2-tuple (IID, MAC) with its binding to a locator set (IP RLOC)

The registration of these EIDs MUST follow the LCAF format as defined in [RFC8060] and as illustrated in the following figure:
The L2 EID record follows the structure as described in [I-D.ietf-lisp-rfc6833bis].

5.2.3. Interface to the LISP Mapping System

The interface between the xTRs and the Mapping System is described in [I-D.ietf-lisp-rfc6833bis] and this document follows the specification as described there. When available, the registrations MAY be implemented over a reliable transport.

In order to support system convergence after mobility, when the Map-Server receives a new registration for a specific EID, it MUST send a Map-Notify to the entire RLOC set in the site that last registered this same EID. This Map-Notify is used to track moved-away state of L2 EIDs as described in Section 5.2.4.

5.2.4. SMR support to track L2 hosts that moved away

In order to support mobility, an ETR SHALL maintain a list of EID records for which it has to generate a SMR message whenever it receives traffic with that EID as destination.
The particular strategy to maintain a SMR table is implementation specific. In essence the table SHOULD provide support for the following:

* Keep track of end-hosts that were once connected to an ETR and have moved away.

* Support for L2 EID records, the 2-tuple (IID, MAC), for which a SMR message SHOULD be generated.

5.2.5. L2 Broadcast and Multicast traffic

Broadcast and Multicast traffic on the L2-overlay is supported by either replicated unicast, or underlay (RLOC) multicast.

xTRs that offer L2 overlay services and are part of the same Instance-ID join a common Multicast Group. When required, this group allows ITRs to send traffic that needs to be replicated (flooded) to all ETRs participating in the L2-overlay (e.g., broadcast traffic within a subnet). When the core network (RLOC space) supports native multicast ETRs participating in the L2-overlay join a (*,G) group associated to the Instance-ID.

When multicast is not available in the core network, each xTR that is part of the same instance-ID SHOULD register a (S,G) entry to the mapping system using the procedures described in [RFC8378], where S is 0000-0000-0000/0 and G is ffff-ffff-ffff/48. This strategy allows and ITR to know which ETRs are part of the L2 overlay and it can head-end replicate traffic to.

Following the same case, when multicast is not available in the core network, the procedures in [RFC8378] can be used to ensure proper distribution of link-local multicast traffic across xTRs participating in the L2 overlay. In such case, the xTRs SHOULD join a (S,G) entry with S being 0000-0000-0000/0 and where G is 0100-0000-0000/8.

5.2.6. L2 Unknown Unicast Support

An ITR attempts to resolve MAC destination misses through the Mapping System. When the destination host remains undiscovered the destination is considered an Unknown Unicast.

A Map-Server SHOULD respond to a Map-Request for an undiscovered host with a Negative Map-Reply with action "Native Forward". Alternatively the action "Drop" may be used in order to suppress Unknown Unicast forwarding.
An ITR that receives a Negative Map-Reply with Action "Native Forward" will handle traffic for the undiscovered host as L2 Broadcast traffic and will be unicast replicated or flooded using underlay multicast to the rest of ETRs in the Layer-2 overlay.

Upon discovery of a previously unknown unicast MAC EID, a data triggered SMR for the discovered EID should be sent by the discovery ETR back to the ITRs that are flooding the unknown unicast traffic. This would allow ITRs to refresh their caches and stop flooding unknown unicast traffic as necessary.

5.2.7. Time-to-Live Handling in Data-Plane

When using a L2 overlay and the encapsulated traffic is IP traffic, the Time-to-Live value of the inner IP header MUST remain unmodified during encapsulation and decapsulation. Network hops traversed as part of the L2 overlay SHOULD be hidden to tools like traceroute and applications that require direct L2 connectivity.

5.3. Support to ARP resolution through Mapping System

5.3.1. Map-Server support to ARP resolution: Packet Flow

A large majority of applications are IP based and, as a consequence, end systems are typically provisioned with IP addresses as well as MAC addresses.

In this case, to limit the flooding of ARP traffic and reduce the use of multicast in the RLOC network, the LISP mapping system MAY be used to support ARP resolution at the ITR.

In order to provide this support, ETRs handle and register an additional EID-to-RLOC mapping as follows,

* EID-record contents = <IID, IP>, RLOC-record contents <MAC>.

There is a dedicated IID used for the registration of the ARP/ND related mappings. Thus, a system with L2 and L3 overlays as well as ARP/ND mappings would have three IIDs at play. In the spirit of providing clarity, we will refer to those IIDs as L2-IID, L3-IID and ARP-IID respectively. By using these definitions, we do not intend to coin new terminology, nor is there anything special about those IIDs that would make them differ from the generic definition of an IID. The types of mappings expected in such a system are summarized below:

* EID = <IID1, IP> to RLOC = <IP-RLOC> (L3-overlay)
The following packet flow sequence describes the use of the LISP Mapping System to support ARP resolution for hosts residing in a subnet that is extended to multiple sites. Using Figure 1, End-Device 2 tries to find the MAC address of End-Device 3. Note that both have IP addresses within the same subnet:

* End-Device 2 sends a broadcast ARP message to discover the MAC address of End-Device 3. The ARP request targets IP 3.0.0.3.

* ITR B receives the ARP message, but rather than flooding it on the overlay network sends a Map-Request to the mapping database system for EID = <IID2,3.0.0.3>.

* When receiving the Map-Request, the Map-Server sends a Proxy-Map-Reply back to ITR B with the mapping EID = <IID2,3.0.0.3> -> MAC 0:0:3:0:0:3.

* Using this Map-Reply, ITR B sends an ARP-Reply back to End-Device 2 with the tuple IP 3.0.0.3, MAC 0:0:3:0:0:3.

* End-Device 2 learns MAC 0:0:3:0:0:3 from the ARP message and can now send a L2 traffic to End-Device 3. When this traffic reaches ITR B is sent over the L2-overlay as described above in Section 5.1.1.

This example shows how LISP, by replacing dynamic data plane learning (such as Flood-and-Learn) can reduce the use of multicast in the underlay network.

Note that ARP resolution using the Mapping System is a stateful operation on the ITR. The source IP,MAC tuple coming from the ARP request have to be stored to generate the ARP-reply when the Map-Reply is received.

Note that the ITR SHOULD cache the ARP entry. In that case future ARP-requests can be handled without sending additional Map-Requests.

5.3.2. ARP registrations: MAC as a locator record

When an end-host is attached or detected in an ETR that provides L2-overlay services and also supports ARP resolution using the LISP control-plane, an additional mapping entry is registered to the mapping system:
* The EID 2-tuple (IID, IP) and its binding to a corresponding host MAC address.

In this case both the xTRs and the Mapping System MUST support an EID-to-RLOC mapping where the MAC address is set as a locator record.

In order to guarantee compatibility with current implementations of xTRs, the MAC locator record SHALL be encoded following the AFI-List LCAF Type defined in [RFC8060]. This option would also allow adding additional attributes to the locator record, while maintaining compatibility with legacy devices.

This mapping is registered with the Mapping System using the following EID record structure,

```
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Record TTL |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Locator Count | EID mask-len | ACT | A | Reserved |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Rsvd | Map-Version Number | AFI = 16387 |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Rsvd1 | Flags | Type = 2 | IID mask-len |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 4 + n | Instance-ID... |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| ...Instance-ID | EID-AFI = 1 or 2 |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| EID-Prefix (IPv4 or IPv6) |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Priority | Weight | M Priority | M Weight |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Unused Flags | AFI = 16387 |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Rsvd1 | Flags | Type = 1 | Rsvd2 |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| 2 + 6 | AFI = 6 |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| Layer-2 MAC Address ... |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
| ... Layer-2 MAC Address |
++---+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+-+
```

An EID record with a locator record that carries a MAC address follows the same structure as described in [I-D.ietf-lisp-rfc6833bis]. However, some fields of the EID record and the locator record require special consideration:
Locator Count: This value SHOULD be set to 1.

Instance-ID: This is the IID used to provide segmentation of the L2-overlays, L3 overlays and ARP tables.

Priority and Weight: IP to MAC bindings are one to one bindings. An ETR SHOULD not register more than one MAC address in the locator record together with an IP based EID. The Priority of the MAC address record is set to 255. The Weight value SHOULD be ignored and the recommendation is to set it to 0.

L bit: This bit of the locator record SHOULD only be set to 1 when an ETR is registering its own IP to MAC binding.

p bit: This bit of the locator record SHOULD be set to 0.

R bit: This bit of the locator record SHOULD be set to 0.

Note that an IP EID record that carries a MAC address in the locator record, SHALL be registered with the Proxy Map-Reply bit set.

5.3.3. Implementation Considerations

While ARP support through the LISP Mapping System fits the LISP Control-Plane there are a series of considerations to take into account when providing this feature:

* As indicated, when an end-host is attached the ETR maintains and registers a mapping with the binding EID = <IID, IP> -> RLOC = <MAC>.

* ARP support through the LISP Mapping System is OPTIONAL and the xTRs should allow the possibility to enable or disable the feature.

* When the ARP entry has not been registered, a Map Server SHOULD send a Negative Map-Reply with action "No Action" as a response to an ARP based Map Request.

* In case the ITR receives a Negative Map-Reply for an ARP request it should fallback to flooding the ARP packet as any other L2 Broadcast packet (as described in Section 5.2.5).

* When receiving a positive Map-Reply for an ARP based Map-Request, the ETR MUST recreate the actual ARP Reply, impersonating the real host. As a consequence, ARP support is a stateful operation where the ITR needs to store enough information about the host that generates an ARP request in order to recreate the ARP Reply.
* ARP replies learned from the Mapping System SHOULD be cached and
  the information used to reply to subsequent ARP requests to the
  same hosts.

6. Optional Deployment Models

The support of an integrated L2 and L3 overlay solution takes
multiple architectural design options, that depend on the specific
requirements of the deployment environment. While some of the
previous describe specific packet flows and solutions based on the
recommended solution, this section documents OPTIONAL design
considerations that differ from the recommended one but that MAY be
required on alternative deployment environments.

6.1. IP Forwarding of Intra-subnet Traffic

As pointed out at the beginning the recommended selection of the L2
and L3 overlays is not the only one possible. In fact, providing L2
extension to some cloud platforms is not always possible and subnets
need to be extended using the L3 overlay.

In order to send all IP traffic (intra- and inter-subnet) through the
L3 overlay the solution MUST change the ARP resolution process
described in Section 5.3.1 to the following one (we follow again
Figure 1 to drive the example. End-Device 2 queries about End-Device
3):

* End-Device 1 sends a broadcast ARP message to discover the MAC
  address of 3.0.0.3.

* ITR B receives the ARP message and sends a Map-Request to the
  Mapping System for EID = <IID1,3.0.0.3>.

* In this case, the Map-Request is routed by the Mapping system
  infrastructure to ETR C, that will send a Map-Reply back to ITR B
  containing the mapping EID = <IID1,3.0.0.3> -> RLOC=IP_C.

* ITR B populates its local cache with the received entry on the L3
  forwarding table. Then, using the cache information it sends a
  Proxy ARP-reply with its own MAC (MAC_xTR_B) address to end End-
  Device 1.

* End-Device 1 learns MAC_ITR_B from the proxy ARP-reply and sends
  traffic with destination address 3.0.0.3 and destination MAC,
  MAC_xTR_B.

* As the destination MAC address is the one from xTR B, when xTR B
  receives this traffic it is forwarded using the L3-overlay.
Note that when implementing this solution, when a host that is local to an ETR moves away, the ETR SHOULD locally send a Gratuitous ARP with its own MAC address and the IP of the moved host, to refresh the ARP tables of local hosts and guarantee the use of the L3 overlay when connecting to the remote host.

It is also important to note that using this strategy to extend subnets through the L3 overlay but still keeping the L2 overlay for the rest of traffic MAY lead to flow asymmetries. This MAY be the case in deployments that filter Gratuitous ARPs, or when moved hosts continue using actual L2 information collected before a migration.

6.2. Data-plane Encapsulation Options

The LISP control-plane offers independence from the data-plane encapsulation. Any encapsulation format that can carry a 24-bit instance-ID can be used to provide the unified overlay.

Common encapsulation formats that can be used are [VXLAN-GPE], [LISP] and [VXLAN]:

* VXLAN-GPE encap: This encapsulation format is defined in [I-D.ietf-lisp-gpe]. It allows encapsulation both L2 and L3 packets and the VNI field directly maps to the Instance-ID used in the control plane. Note that when using this encapsulation for a unified solution the P-bit is set and the Next-Protocol field is used (typically with values 0x1 (IPv4) or 0x2 (IPv6) in L3-overlays, and value 0x3 in L2-overlays).

* LISP encap: This is the encapsulation format defined in the LISP specification [I-D.ietf-lisp-rfc6830bis]. The encapsulation allows encapsulating both L2 and L3 packets. The Instance-ID used in the EIDs directly maps to the Instance-ID that the LISP header carries. At the ETR, after decapsulation, the IID MAY be used to decide between L2 processing or L3 processing.

* VXLAN encap: This is a L2 encapsulation format defined in [RFC7348]. While being a L2 encapsulation it can be used both for L2 and L3 overlays. The Instance-ID used in LISP signaling maps to the VNI field of the VXLAN header. Providing L3 overlays using VXLAN generally requires using the ETR MAC address as destination MAC address of the inner Ethernet header. The process to learn or derive this ETR MAC address is not included as part of this document.
7. IANA Considerations

This memo includes no request to IANA.

8. Acknowledgements

This draft builds on top of two expired drafts that introduced the concept of LISP L2/L3 overlays (draft-maino-nvo3-lisp-cp and draft-hertoghs-nvo3-lisp-controlplane-unified). Many thanks to the combined authors of those drafts, that SHOULD be considered main contributors of this draft as well: Vina Ermagan, Dino Farinacci, Yves Hertoghs, Luigi Iannone, Fabio Maino, Victor Moreno, and Michael Smith.

9. References

9.1. Normative References


9.2. Informative References


Authors’ Addresses

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

Vrushali Ashtaputre
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
United States of America
Email: vrushali@cisco.com

Fabio Maino
Cisco Systems
170 Tasman Drive
San Jose, CA 95134
United States of America
Email: fmaino@cisco.com

Victor Moreno
Google LLC
1600 Amphitheatre Pkwy
Mountain View, CA 94043
United States of America
Email: vimoreno@google.com

Dino Farinacci
lispers.net
San Jose, CA
United States of America
Email: farinacci@gmail.com