Network Working Group
Internet-Draft
Intended status: Experimental
Expires: April 21, 2014

F. Maino
V. Ermagan
Y. Hertoghs
Cisco Systems
D. Farinacci
lispers.net
M. Smith
Insieme Networks
October 18, 2013

LISP Control Plane for Network Virtualization Overlays
draft-maino-nvo3-lisp-cp-03

Abstract

The purpose of this draft is to analyze the mapping between the Network Virtualization over L3 (NVO3) requirements and the capabilities of the Locator/ID Separation Protocol (LISP) control plane. This information is provided as input to the NVO3 analysis of the suitability of existing IETF protocols to the NVO3 requirements.

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].

Table of Contents

1. Introduction

The purpose of this draft is to analyze the mapping between the Network Virtualization over L3 (NVO3) [I-D.ietf-nvo3-overlay-problem-statement] requirements and the capabilities of the Locator/ID Separation Protocol (LISP) [RFC6830] control plane. This information is provided as input to the NVO3 analysis of the suitability of existing IETF protocols to the NVO3 requirements.

LISP is a flexible map and encap framework that can be used for overlay network applications, including Data Center Network Virtualization.

The LISP framework provides two main tools for NVO3: (1) a Data Plane that specifies how Endpoint Identifiers (EIDs) are encapsulated in Routing Locators (RLOCs), and (2) a Control Plane that specifies the interfaces to the LISP Mapping System that provides the mapping between EIDs and RLOCs.

This document focuses on the control plane for L2 over L3 LISP encapsulation, where EIDs are associated with MAC addresses. As such the LISP control plane can be used with the data path encapsulations defined in VXLAN [I-D.mahalingam-dutt-dcops-vxlan] and in NVGRE [I-D.sridharan-virtualization-nvgre]. The LISP control plane can, of course, be used with the L2 LISP data path encapsulation defined in [I-D.smith-lisp-layer2].

The LISP control plane provides the Mapping Service for the Network Virtualization Edge (NVE), mapping per-tenant end system identity information on the corresponding location at the NVE. As required by NVO3, LISP supports network virtualization and tenant separation to hide tenant addressing information, tenant-related control plane activity and service contexts from the underlay network.

The LISP control plane is extensible, and can support non-LISP data path encapsulations such as [I-D.sridharan-virtualization-nvgre], or other encapsulations that provide support for network virtualization. [RFC6832] specifies an open interworking framework to allow LISP to non-LISP sites communication.

Broadcast, unknown unicast, and multicast in the overlay network are supported by either replicated unicast, or core-based multicast as specified in [RFC6831], [I-D.farinacci-lisp-mr-signaling], and [I-D.farinacci-lisp-te].

Finally, the LISP architecture has a modular design that allows the use of different Mapping Databases, provided that the interface to the Mapping System remains the same [RFC6833]. This allows for different Mapping Databases that may fit different NVO3 deployments. As an example of the modularity of the LISP Mapping System, a worldwide LISP pilot network is currently using an hierarchical Delegated Database Tree [I-D.ietf-lisp-ddt], after having been operated for years with an overlay BGP mapping infrastructure [RFC6836].

The LISP mapping system supports network virtualization, and a single mapping infrastructure can run multiple instances, either public or private, of the mapping database.

The rest of this document, after giving a quick a LISP overview in Section 3, follows the functional model defined in [I-D.ietf-nvo3-framework] that provides in Section 4 an overview of the LISP NVO3 reference model, and in Section 5 a description of its functional components. Section 6 contains various considerations on key aspects of LISP NVO3, followed by security considerations in Section 7.

2. Definition of Terms

Flood-and-Learn: the use of dynamic (data plane) learning in VXLAN to discover the location of a given Ethernet/IEEE 802 MAC address in the underlay network.

ARP-Agent Reply: the ARP proxy-reply of an agent (e.g. an ITR) with a MAC address of some other system in response to an ARP request to a target which is not the agent's IP address

For definition of NVO3 related terms, notably Virtual Network (VN), Virtual Network Identifier (VNI), Network Virtualization Edge (NVE), Data Center (DC), please consult [I-D.ietf-nvo3-framework].

For definitions of LISP related terms, notably Map-Request, Map-Reply, Ingress Tunnel Router (ITR), Egress Tunnel Router (ETR), Map-Server (MS) and Map-Resolver (MR) please consult the LISP specification [RFC6830].

3. LISP Overview

This section provides a quick overview of L2 LISP, with focus on control plane operations.

The modular and extensible architecture of the LISP control plane allows its use with both L2 or L3 LISP data path encapsulation. In

fact, the LISP control plane can be used even with other L2 overlay data path encapsulations such as VXLAN and NVGRE. When used with VXLAN, the LISP control plane replaces the use of dynamic data plane learning (Flood-and-Learn), as specified in [I-D.mahalingam-dutt-dcops-vxlan] improving scalability and mitigating multicast requirements in the underlay network.

For a detailed LISP overview please refer to [RFC6830] and related drafts.

To exemplify LISP operations let's consider two data centers (LISP sites) A and B that provide L2 network virtualization services to a number of tenant end systems, as depicted in Figure 1. The Endpoint Identifiers (EIDs) are encoded according to [I-D.ietf-lisp-lcaf] as an <IID,MAC> tuple that contains the Instance ID, or Virtual Network Identifier (VNI), and the endpoint Ethernet/IEEE 802 MAC address.

The data centers are connected via a L3 underlay network, hence the Routing Locators (RLOCs) are IP addresses (either IPv4 or IPv6) encoded according to [I-D.ietf-lisp-lcaf].

In LISP the network virtualization edge function is performed by Ingress Tunnel Routers (ITRs) that are responsible for encapsulating the LISP ingress traffic, and Egress Tunnel Routers (ETRs) that are responsible for decapsulating the LISP egress traffic. ETRs are also responsible to register the EID-to-RLOC mapping for a given LISP site in the LISP mapping database system. ITRs and ETRs are collectively referred as xTRs.

The EID-to-RLOC mapping is stored in the LISP mapping database, a distributed mapping infrastructure accessible via Map Servers (MS) and Map Resolvers (MR). [I-D.ietf-lisp-ddt] is an example of a mapping database used in many LISP deployments. Another example of of mapping database is [RFC6836].

For small deployments the mapping infrastructure can be very minimal, in some cases even a single system running as MS/MR.

```
                             ,---------.
                           ,'           `.
                          (Mapping System )
                           `.           ,'
                             `-+------+'
                          +--+--+   +-+---+
                          |MS/MR|   |MS/MR|
                          +-+---+   +-----+
                            |         |
                          .--..--. .--. ..
```

```
                                  (    '           '.--.
                               .-.'        L3          '
                              (         Underlay       )
                               (                     '-'
                                ._.'--'._.'.-._.'.-._)
                       RLOC=IP_A //                  \\ RLOC=IP_B
                              +---+--+              +-+--+--+
                        .--.-.|xTR A |'.-.         .| xTR B |.-.
                       (      +---+--+    )       ( +-+--+--+   )
                      (                __.       (              '.
                    ..'  LISP Site A  )         .'   LISP Site B  )
                   (             .'-'          (             .'-'
                     '--'._.'.    )\            '--'._.'.    )\
                      /       '--'  \            /       '--'  \
                  '--------'    '--------'   '--------'   '--------'
                  :  End   :    :  End   :   :  End   :   :  End   :
                  : Device :    : Device :   : Device :   : Device :
                  '--------'    '--------'   '--------'   '--------'
                     EID=          EID=         EID=         EID=
                 <IID1,MAC_W>  <IID2,MAC_X> <IID1,MAC_Y> <IID1,MAC_Z>
```

Figure 1: Example of L2 NVO3 Services

3.1. LISP Site Configuration

In each LISP site the xTRs are configured with an IP address (the site RLOCs) per each interface facing the underlay network.

Similarly the MS/MR are assigned an IP address in the RLOC space.

The configuration of the xTRs includes the RLOCs of the MS/MR and a shared secret that is optionally used to secure the communication between xTRs and MS/MR.

To provide support for multi-tenancy multiple instances of the mapping database are identified by a LISP Instance ID (IID), that is equivalent to the 24-bit VXLAN Network Identifier (VNI) or Tenant Network Identifier (TNI) that identifies tenants in [I-D.mahalingam-dutt-dcops-vxlan].

3.2. End System Provisioning

We assume that a provisioning framework will be responsible for provisioning end systems (e.g. VMs) in each data center. The provisioning system configures each end system with an Ethernet/IEEE 802 MAC address and/or IP address and provisions the NVE with other end system specific attributes such as VLAN information, and TS/VLAN to VNI mapping information. LISP does not introduce new addressing requirements for end systems.

The provisioning infrastructure is also responsible to provide a network attach function, that notifies the network virtualization edge (the LISP site ETR) that the end system is attached to a given virtual network (identified by its VNI/IID) and that the end system is identified, within that virtual network, by a given Ethernet/IEEE 802 MAC address.

3.3. End System Registration

Upon notification of end system network attach, that includes the EID=<IID,MAC> tuple that identifies that end system, the ETR sends a LISP Map-Register to the Mapping System. The Map-Register includes the EID and RLOCs of the LISP site. The EID-to-RLOC mapping is now available, via the Mapping System Infrastructure, to other LISP sites that are hosting end systems that belong to the same tenant.

For more details on end system registration see [RFC6833].

3.4. Packet Flow and Control Plane Operations

This section provides an example of the unicast packet flow and the control plane operations when in the topology shown in Figure 1 end system W, in LISP site A, wants to communicate to end system Y in LISP site B. We'll assume that W knows Y's EID MAC address (e.g. learned via ARP).

- W sends an Ethernet/IEEE 802 MAC frame with destination EID=<IID1,MAC_Y> and source EID=<IID1,MAC_W>.

- ITR A does a lookup in its local map-cache for the destination EID=<IID1, MAC_Y>. Since this is the first packet sent to MAC_Y, the map-cache is a miss, and the ITR sends a Map-request to the mapping database system looking up the EID=<IID1,MAC_Y>.

- The mapping systems forwards the Map-Request to ETR B, that is aware of the EID-to-RLOC mapping for <IID1,MAC_Y>. 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 A.

o ETR B sends a Map-Reply to ITR A that includes the EID-to-RLOC mapping: EID=<IID1,MAC_Y> -> RLOC=IP_B, where IP_B is the locator of ETR B, hence the locator of LISP site B. In order to facilitate interoperability, the Map-Reply may also include attributes such as the data plane encapsulations supported by the ETR.

o ITR A populates the local map-cache with the EID to RLOC mapping, and either L2 LISP, VXLAN, or NVGRE encapsulates all subsequent packets with a destination EID=<IID1,MAC_Y> with a destination RLOC=IP_B.

It should be noted how the LISP mapping system replaces the use of Flood-and-Learn based on multicast distribution trees instantiated in the underlay network (required by VXLAN's dynamic data plane learning), with a unicast control plane and a cache mechanism that "pulls" on-demand the EID-to-RLOC mapping from the LISP mapping database. This improves scalability, and simplifies the configuration of the underlay network.

3.4.1. Supporting ARP Resolution with LISP Mapping System

A large majority of data center applications are IP based, and in those use cases end systems are provisioned with IP addresses as well as MAC addresses.

In this case, to eliminate the flooding of ARP traffic and further reduce the need for multicast in the underlay network, the LISP mapping system is used to support ARP resolution at the ITR. We assume that as shown in Figure 2: (1) end system W has an IP address IP_W, and end system Y has an IP address IP_Y, (2) end system W knows Y's IP address (e.g. via DNS lookup). We also assume that during registration Y has registered both its MAC address and its IP address as EID. End system Y is then identified by the EID = <IID1,IP_Y,MAC_Y>.

```
                                      ,---------.
                                    ,'           `.
                                   (Mapping System )
                                    `.           ,'
                                      `-+-------+'
                                   +--+--+   +-+---+
                                   |MS/MR|   |MS/MR|
                                   +-+---+   +-----+
                                       |        |
                                   .--..--. .--. ..
                                  (    '           '.--.
                              .-.'        L3           '
                             (         Underlay         )
```

```
                               (                     '-'
                                ._.'--'._.'.-._.'.-._)
                       RLOC=IP_A //                  \\ RLOC=IP_B
                              +---+--+             +-+--+--+
                        .--.-.|xTR A |'.-.        .| xTR B |.-.
                       (      +---+--+    )      ( +-+--+--+   )
                      (                __.      (              '.
                    ..'  LISP Site A  )        .'   LISP Site B  )
                   (             .'-'         (              .'-'
                     '--'._.'.    )\           '--'._.'.    )\
                      /       '--'  \           /       '--'  \
                  '--------'    '--------'   '--------'   '--------'
                  :  End   :    :  End   :   :  End   :   :  End   :
                  : Device :    : Device :   : Device :   : Device :
                  '--------'    '--------'   '--------'   '--------'
                     EID=          EID=          EID=          EID=
                  <IID1,IP_W,  <IID2,IP_X,   <IID1,IP_Y,  <IID1,IP_Z,
                    MAC_W>        MAC_X>        MAC_Y>        MAC_Z>
```

Figure 2: Example of L3 NVO3 Services

The packet flow and control plane operation are as follows:

- End system W sends a broadcast ARP message to discover the MAC address of end system Y. The message contains IP_Y in the ARP message payload.

- ITR A, acting as a L2 switch, will receive the ARP message, but rather than flooding it on the overlay network sends a Map-Request to the mapping database system for EID = <IID1,IP_Y,*>.

- The Map-Request is routed by the mapping system infrastructure to ETR B, that will send a Map-Reply back to ITR A containing the mapping EID=<IID1,IP_Y,MAC_Y> -> RLOC=IP_B, (the locator of ETR B). Alternatively, depending on the mapping system configuration, a Map-Server in the mapping system may send directly a Map-Reply to ITR A.

- ITR A populates the map-cache with the received entry, and sends an ARP-Agent Reply to W that includes MAC_Y and IP_Y.

- End system W learns MAC_Y from the ARP message and can now send a packet to end system Y by including MAC_Y, and IP_Y, as destination addresses.

- ITR A will then process the packet as specified in Section 3.4.

This example shows how LISP, by replacing dynamic data plane learning (Flood-and-Learn) largely reduces the need for multicast in the underlay network, that is needed only when broadcast, unknown unicast or multicast are required by the applications in the overlay. In practice, the LISP mapping system, constrains ARP within the boundaries of a link-local protocol. This simplifies the configuration of the underlay network and removes the significant scalability limitation imposed by VXLAN Flood-and-Learn.

It's important to note that the use of the LISP mapping system, by pulling the EID-to-RLOC mapping on demand, also improves end system mobility across data centers.

3.5. End System Mobility

This section shows how the LISP control plane deals with mobility when end systems are migrated from one Data Center to another. We'll assume that a signaling protocol, as described in [I-D.kompella-nvo3-server2nve], signals to the NVE operations such as creating/terminating/migrating an end system. The signaling protocol consists of three basic messages: "associate", "dissociate", and "pre-associate".

Let's consider the scenario shown in Figure 3 where end system W moves from data center A to data center B.

```
                                     ,---------.
                                   ,'           `.
                                  (Mapping System )
                                   `.           ,'
                                     `-+------+'
                                  +--+--+   +-+---+
                                  |MS/MR|   |MS/MR|
                                  +-+---+   +-----+
                                      |        |
                                  .--..--. .--. ..
                                 (    '           '.--.
                              .-.'        L3          '
                             (         Underlay       )
                              (                     '-'
                               ._.'--'._.'.-._.'.-._)
                      RLOC=IP_A //                  \\ RLOC=IP_B
                             +---+--+              +-+--+--+
                       .--.-.|xTR A |'.-.         .| xTR B |.-.
                      (      +---+--+    )       ( +-+--+--+   )
                     (                __.       (              '.
                   ..'  LISP Site A  )         .'   LISP Site B  )
                  (              .'-'         (              .'-'
```

```
       '--'._.'.    )\           '--'._.'.    )\
         /       '--'  \           /       '--'  \
     '--------'   '--------'     '--------'   '--------'
     :  End   :   :  End   : ==> :  End   :   :  End   :
     : Device :   : Device : ==> : Device :   : Device :
     '--------'   '--------'     '--------'   '--------'
        EID=          EID=<IID1,MAC_W>          EID=
   <IID2,MAC_X>                             <IID1,MAC_Z>
```

Figure 3: End System Mobility

As a result of the end system registration, described in Section 3.3, the Mapping System contains the EID-to-RLOC mapping for end system W that associates EID=<IID1, MAC_W> with the RLOC(s) associated with LISP site A (IP_A).

The process of migrating end system W from data center A to data center B is initiated.

ETR B receives a pre-associate message that includes EID=<IID1, MAC_W>. ETR B sends a Map-Register to the mapping system registering RLOC=IP_B as an additional locator for end system W with priority set to 255. This means that the RLOC MUST NOT be used for unicast forwarding, but the mapping system is now aware of the new location.

During the migration process of end system W, ETR A receives a dissociate message, and sends a Map-Register with Record TTL=0 to signal the mapping system that end system W is no longer reachable at RLOC=IP_A. xTR A will also add an entry in its forwarding table that marks EID=<IID1, MAC_W> as non-local. When end system W has completed its migration, ETR B receives an associate message for end system W, and sends a Map-Register to the mapping system setting a non-255 priority for RLOC=IP_B. Now the mapping system is updated with the new EID-to-RLOC mapping for end system W with the desired priority.

The remote ITRs that were corresponding with end system W during the migration will keep sending packets to ETR A. ETR A will keep forwarding locally those packets until it receives a dissociate message, and the entry in the forwarding table associated with EID=<IID1, MAC_W> is marked as non-local. Subsequent packets arriving at ETR A from a remore ITR, and destined to end system W will hit the entry in the forwarding table that will generate an exception, and will generate a Solicit-Map-Request (SMR) message that is returned to the remote ITR. Upon receiving the SMR the remote ITR will invalidate its local map-cache entry for EID=<IID1, MAC_W> and send a new Map-Request for that EID. The Map-Request will generate a Map-Reply that includes the new EID-to-RLOC mapping for end system W

with RLOC=IP_B. Similarly, unencapsulated packets arriving at ITR A from local end systems and destined to end system W will hit the entry in the forwarding table marked as non-local, and will generate an exception that by sending a Map-Request for EID=<IID1, MAC_W> will populate the map-cache of ITR A with an EID-to-RLOC mapping for end system W with RLOC=IP_B.

3.6. L3 LISP

The two examples above shows how the LISP control plane can be used in combination with either L2 LISP, VXLAN, or NVGRE encapsulation to provide L2 network virtualization services across data centers.

There is a trend, led by Massive Scalable Data Centers, that is accelerating the adoption of L3 network services in the data center, to preserve the many benefits introduced by L3 (scalability, multi-homing, ...).

LISP, as defined in [RFC6830], provides L3 network virtualization services over an L3 underlay network that, as an alternative to L2 overlay solutions, matches the requirements for DC Network Virtualization. L2 overlay solutions are necessary for Data Centers that rely on non IPv4/IPv6 protocols, but when IP is pervasive L3 LISP provides a better and more scalable overlay.

4. Reference Model

Figure 4, taken from [I-D.ietf-nvo3-framework], introduces the NVO3 reference model.

In a LISP NVO3 network the Tenant Systems (TS) that are homed to a common NVE, having specific Endpoint Identifiers (EIDs), are part of a 'LISP site'.

The network virtualization edge (NVE) function is performed by Ingress Tunnel Routers (ITRs) that are responsible for encapsulating the LISP ingress traffic, and Egress Tunnel Routers (ETRs) that are responsible for de-encapsulating the LISP egress traffic.

The outer tunnel IP addresses (either IPv4 or IPv6) on the ITR and ETR NVE function are known as Routing Locators (RLOCs).

ETRs are also responsible to register the EID-to-RLOC mapping for a given LISP site in the LISP mapping database system [RFC6833] .

ITRs and ETRs, collectively referred as xTRs, provide for tenant separation, perform the encap/decap function, and interface with the LISP Mapping System that maps tenant addressing information (in the

EID name space) on the underlay L3 infrastructure (in the RLOC name space), with the encoding defined in [I-D.ietf-lisp-lcaf].

The LISP Mapping system is a distributed mapping infrastructure, accessible via Map Servers (MS) and Map Resolvers (MR), that performs the NVA function.

The LISP Mapping system can be scaled across various physical components e.g. across an EID based hierarchy as described in [I-D.ietf-lisp-ddt]. EID prefixes and/or address families can thus be scaled across the mapping infrastructure if needed.

The NVA function can offer a northbound SDN interface in order to program the EID to RLOC mapping from e.g. an orchestration system. An example is given in [I-D.barkai-lisp-nfv] .

As traffic reaches An ingress NVE, the corresponding ITR uses the LISP Map-Request/Reply service to determine the location of the destination End System.

The LISP mapping system combines the distribution of address advertisement and (stateless) tunneling provisioning.

LISP defines several mechanisms for determining RLOC reachability, including Locator Status Bits, "nonce echoing", and RLOC probing. Please see Sections 5.3 and 6.3 of [RFC6830]. However, given the fact that DC's are typically deployed with a single stage IGP hierarchy, the IGP responsible for the RLOC space offers enough reachability information.

```
                +--------+                                    +--------+
                | Tenant +--+                            +----| Tenant |
                | System |  |                           (')   | System |
                +--------+  |    .................     (   )  +--------+
                            |  +---+           +---+    (_)
                            +--|NVE|---+   +---|NVE|-----+
                               +---+   |   |   +---+
                               / .    +-----+      .
                              /  . +--| NVA |      .
                             /   . |  +-----+      .
                            |    . |               .
                            |    . |  L3 Overlay +--+--++--------+
                +--------+  |    . |   Network   | NVE || Tenant |
                | Tenant +--+    . |             |     || System |
                | System |       .  \ +---+      +--+--++--------+
                +--------+       .....|NVE|.........
                                      +---+
                                        |
```

```
                         |
            =====================
              |               |
          +--------+      +--------+
          | Tenant |      | Tenant |
          | System |      | System |
          +--------+      +--------+
```

Figure 4: NVO3 Generic Reference Model

4.1. LISP NVE Service Types

L2 NVE and L3 NVE service types thanks to the flexibility provided by the LISP Canonical Address Format [I-D.ietf-lisp-lcaf], that allows for EIDs to be encoded either as MAC addresses or IP addresses.

4.1.1. LISP L2 NVE Services

The frame format defined in [I-D.mahalingam-dutt-dcops-vxlan], has a header compatible with the LISP data path encapsulation header, when MAC addresses are used as EIDs, as described in section 4.12.2 of [I-D.ietf-lisp-lcaf].

The LISP control plane is extensible, and can support non-LISP data path encapsulations such as NVGRE [I-D.sridharan-virtualization-nvgre], or other encapsulations that provide support for network virtualization.

4.1.2. LISP L3 NVE Services

LISP is defined as a virtualized IP routing and forwarding service in [RFC6830], and as such can be used to provide L3 NVE services.

5. Functional Components

This section describes the functional components of a LISP NVE as defined in Section 3 of [I-D.ietf-nvo3-framework].

5.1. Generic Service Virtualization Components

The generic reference model for NVE is depicted in [I-D.ietf-nvo3-framework].

```
                    +-------- L3 Network -------+
                    |                           |
                    |        Tunnel Overlay     |
        +------------+---------+       +---------+------------+
        | +----------+-------+ |       | +---------+--------+ |
```

```
                | |  Overlay Module  | |      | |  Overlay Module  | |
                | +---------+--------+ |      | +---------+--------+ |
                |           |VN context|      | VN context|          |
                |           |          |      |           |          |
                |  +--------+-------+  |      |  +--------+-------+  |
                |  | |VNI|   .  |VNI|  |      |  | |VNI|   .  |VNI|  |
           NVE1 |  +-+------------+-+  |      |  +-+-----------+--+  | NVE2
                |    |   VAPs     |    |      |    |    VAPs   |     |
                +----+------------+----+      +----+-----------+-----+
                     |            |                |           |
                     |            |                |           |
                    Tenant Systems                Tenant Systems
```

Figure 5: Generic reference model for NV Edge

5.1.1. Virtual Attachment Points (VAPs)

In a LISP NVE, Tunnel Routers (xTRs) implement the NVE functionality on ToRs or Virtual Switches. Tenant Systems attach to the Virtual Access Points (VAPs) provided by the xTRs (either a physical port or a virtual interface).

The VAPs are identified by either a physical port or a virtual interface, e.g. Indexed by VLAN tags, a set, range, or set of ranges of VLAN tags in the case of a L2 service, or a virtual routed interface Indexed by a VLAN in case of a L3 service, or a combination of them in case of An L2/L3 service.

5.1.2. Overlay Modules and Tenant ID

The xTR also implements the function of NVE Overlay Module, by mapping the addressing information (EIDs) of the tenant packet on the appropriate locations (RLOCs) in the underlay network. The Virtual Network Identifier (VNI) is encoded in the encapsulated packet (either in the 24-bit IID field of the LISP header for L2/L3 LISP encapsulation, or in the 24-bit VXLAN Network Identifier field for VXLAN encapsulation, or in the 24-bit NVGRE Tenant Network Identifier field of NVGRE). In a LISP NVE globally unique (per administrative domain) VNIs are used to identify the Tenant instances.

The mapping of the tenant packet address onto the underlay network location is "pulled" on-demand from the mapping system, and cached at the NVE in a per-VNI map-cache.

5.1.3. Tenant Instance

Tenants are mapped on LISP Instance IDs (IIDs), and the LISP Control Plane uses the IID to provide segmentation and virtualization. The

ETR is responsible to register the Tenant System to the LISP mapping system, via the Map-Register service provided by LISP Map-Servers (MS). The Map-Register includes the IID that is used to identify the tenant.

5.1.4. Tunnel Overlays and Encapsulation Options

The LISP control protocol, as defined today, provides support for L2 LISP and VXLAN L2 over L3 encapsulation, and LISP L3 over L3 encapsulation, as well as support for the Generic Protocol Extensions for LISP and VXLAN defined in [I-D.lewis-lisp-gpe] and [I-D.quinn-vxlan-gpe] respectively. The Generic Protocol Extensions can be used to offer a concurrent L2 and L3 overlay across the same dataplane.

We believe that the LISP control Protocol can be easily extended to support different IP tunneling options (such as NVGRE).

5.1.5. Control Plane Components

5.1.5.1. Auto-provisioning/Service Discovery

The LISP framework does not include mechanisms to provision the local NVE with the appropriate Tenant Instance for each Tenant System. Other protocols, such as VDP (in IEEE P802.1Qbg), should be used to implement a network attach/detach function.

The LISP control plane can take advantage of such a network attach/detach function to trigger the registration of a Tenant End System to the Mapping System. This is particularly helpful to handle mobility across DC of the Tenant End System.

It is possible to extend the LISP control protocol to advertise the tenant service instance (tenant and service type provided) to other NVEs, and facilitate interoperability between NVEs that are using different service types.

5.1.5.2. Address Advertisement and Tunnel mapping

As traffic reaches an ingress NVE, the corresponding ITR uses the LISP Map-Request/Reply service to determine the location of the destination End System.

The LISP mapping system combines the distribution of address advertisement and (stateless) tunneling provisioning.

When EIDs are mapped on both IP addresses and MACs, the need to flood ARP messages at the NVE is eliminated resolving the issues with explosive ARP handling.

5.1.5.3. Tunnel Management

LISP defines several mechanisms for determining RLOC reachability, including Locator Status Bits, "nonce echoing", and RLOC probing. Please see Sections 5.3 and 6.3 of [RFC6830].

6. Key Aspects of Overlay

6.1. Overlay Issues to Consider

6.1.1. Data Plane vs. Control Plane Driven

The use of LISP control plane minimizes the need for multicast in the underlay network overcoming the scalability limitations of VXLAN dynamic data plane learning (Flood-and-Learn).

Multicast or ingress replication in the underlay network are still required, as specified in [RFC6831], [I-D.farinacci-lisp-mr-signaling], and [I-D.farinacci-lisp-te], to support broadcast, unknown, and multicast traffic in the overlay, but multicast in the underlay is no longer required (at least for IP traffic) for unicast overlay services.

6.1.2. Data Plane and Control Plane Separation

LISP introduces a clear separation between data plane and control plane functions. LISP modular design allows for different mapping databases, to achieve different scalability goals and to meet requirements of different deployments.

6.1.3. Handling Broadcast, Unknown Unicast and Multicast (BUM) Traffic

Packet replication in the underlay network to support broadcast, unknown unicast and multicast overlay services can be done by:

- o Ingress replication

- o Use of underlay multicast trees

[RFC6831] specifies how to map a multicast flow in the EID space during distribution tree setup and packet delivery in the underlay network. LISP-multicast doesn't require packet format changes in multicast routing protocols, and doesn't impose changes in the internal operation of multicast in a LISP site. The only operational changes are required in PIM-ASM [RFC4601], MSDP [RFC3618], and PIM-SSM [RFC4607].

7. Security Considerations

[I-D.ietf-lisp-sec] defines a set of security mechanisms that provide origin authentication, integrity and anti-replay protection to LISP's EID-to-RLOC mapping data conveyed via mapping lookup process. LISP-SEC also enables verification of authorization on EID-prefix claims in Map-Reply messages.

Additional security mechanisms to protect the LISP Map-Register messages are defined in [RFC6833].

The security of the Mapping System Infrastructure depends on the particular mapping database used. The [I-D.ietf-lisp-ddt] specification, as an example, defines a public-key based mechanism that provides origin authentication and integrity protection to the LISP DDT protocol.

8. IANA Considerations

This document has no IANA implications

9. Acknowledgements

The authors want to thank Victor Moreno and Paul Quinn for the early review, insightful comments and suggestions.

10. References

10.1. Normative References

[RFC2119] Bradner, S., "Key words for use in RFCs to Indicate Requirement Levels", BCP 14, RFC 2119, March 1997.

[RFC3618] Fenner, B. and D. Meyer, "Multicast Source Discovery Protocol (MSDP)", RFC 3618, October 2003.

[RFC4601] Fenner, B., Handley, M., Holbrook, H., and I. Kouvelas, "Protocol Independent Multicast - Sparse Mode (PIM-SM): Protocol Specification (Revised)", RFC 4601, August 2006.

[RFC4607] Holbrook, H. and B. Cain, "Source-Specific Multicast for IP", RFC 4607, August 2006.

10.2. Informative References

[I-D.barkai-lisp-nfv]
sbarkai@gmail.com, s., Farinacci, D., Meyer, D., Maino, F., and V. Ermagan, "LISP Based FlowMapping for Scaling NFV", draft-barkai-lisp-nfv-02 (work in progress), July 2013.

[I-D.farinacci-lisp-mr-signaling]
Farinacci, D. and M. Napierala, "LISP Control-Plane Multicast Signaling", draft-farinacci-lisp-mr-signaling-03 (work in progress), September 2013.

[I-D.farinacci-lisp-te]
Farinacci, D., Lahiri, P., and M. Kowal, "LISP Traffic Engineering Use-Cases", draft-farinacci-lisp-te-03 (work in progress), July 2013.

[I-D.ietf-lisp-ddt]
Fuller, V., Lewis, D., Ermagan, V., and A. Jain, "LISP Delegated Database Tree", draft-ietf-lisp-ddt-01 (work in progress), March 2013.

[I-D.ietf-lisp-lcaf]
Farinacci, D., Meyer, D., and J. Snijders, "LISP Canonical Address Format (LCAF)", draft-ietf-lisp-lcaf-03 (work in progress), September 2013.

[I-D.ietf-lisp-sec]
Maino, F., Ermagan, V., Cabellos-Aparicio, A., Saucez, D., and O. Bonaventure, "LISP-Security (LISP-SEC)", draft-ietf-lisp-sec-04 (work in progress), October 2012.

[I-D.ietf-nvo3-dataplane-requirements]
Bitar, N., Lasserre, M., Balus, F., Morin, T., Jin, L., and B. Khasnabish, "NVO3 Data Plane Requirements", draft-ietf-nvo3-dataplane-requirements-01 (work in progress), July 2013.

[I-D.ietf-nvo3-framework]
Lasserre, M., Balus, F., Morin, T., Bitar, N., and Y. Rekhter, "Framework for DC Network Virtualization", draft-ietf-nvo3-framework-03 (work in progress), July 2013.

[I-D.ietf-nvo3-nve-nva-cp-req]
Kreeger, L., Dutt, D., Narten, T., and D. Black, "Network Virtualization NVE to NVA Control Protocol Requirements", draft-ietf-nvo3-nve-nva-cp-req-00 (work in progress), July 2013.

[I-D.ietf-nvo3-overlay-problem-statement]
Narten, T., Gray, E., Black, D., Fang, L., Kreeger, L., and M. Napierala, "Problem Statement: Overlays for Network Virtualization", draft-ietf-nvo3-overlay-problem-statement-04 (work in progress), July 2013.

[I-D.kompella-nvo3-server2nve]
Kompella, K., Rekhter, Y., Morin, T., and D. Black, "Signaling Virtual Machine Activity to the Network Virtualization Edge", draft-kompella-nvo3-server2nve-02 (work in progress), April 2013.

[I-D.lewis-lisp-gpe]
Lewis, D., Agarwal, P., Kreeger, L., Quinn, P., Smith, M., and N. Yadav, "LISP Generic Protocol Extension", draft-lewis-lisp-gpe-01 (work in progress), October 2013.

[I-D.mahalingam-dutt-dcops-vxlan]
Mahalingam, M., Dutt, D., Duda, K., Agarwal, P., Kreeger, L., Sridhar, T., Bursell, M., and C. Wright, "VXLAN: A Framework for Overlaying Virtualized Layer 2 Networks over Layer 3 Networks", draft-mahalingam-dutt-dcops-vxlan-04 (work in progress), May 2013.

[I-D.quinn-vxlan-gpe]
Quinn, P., Agarwal, P., Fernando, R., Lewis, D., Kreeger, L., Smith, M., and N. Yadav, "Generic Protocol Extension for VXLAN", draft-quinn-vxlan-gpe-01 (work in progress), October 2013.

[I-D.smith-lisp-layer2]
Smith, M., Dutt, D., Farinacci, D., and F. Maino, "Layer 2 (L2) LISP Encapsulation Format", draft-smith-lisp-layer2-03 (work in progress), September 2013.

[I-D.sridharan-virtualization-nvgre]
Sridharan, M., Greenberg, A., Wang, Y., Garg, P., Venkataramiah, N., Duda, K., Ganga, I., Lin, G., Pearson, M., Thaler, P., and C. Tumuluri, "NVGRE: Network Virtualization using Generic Routing Encapsulation", draft-sridharan-virtualization-nvgre-03 (work in progress), August 2013.

[RFC6830] Farinacci, D., Fuller, V., Meyer, D., and D. Lewis, "The Locator/ID Separation Protocol (LISP)", RFC 6830, January 2013.

[RFC6831] Farinacci, D., Meyer, D., Zwiebel, J., and S. Venaas, "The Locator/ID Separation Protocol (LISP) for Multicast Environments", RFC 6831, January 2013.

[RFC6832] Lewis, D., Meyer, D., Farinacci, D., and V. Fuller, "Interworking between Locator/ID Separation Protocol (LISP) and Non-LISP Sites", RFC 6832, January 2013.

[RFC6833] Fuller, V. and D. Farinacci, "Locator/ID Separation Protocol (LISP) Map-Server Interface", RFC 6833, January 2013.

[RFC6836] Fuller, V., Farinacci, D., Meyer, D., and D. Lewis, "Locator/ID Separation Protocol Alternative Logical Topology (LISP+ALT)", RFC 6836, January 2013.

Authors' Addresses

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

Email: fmaino@cisco.com

Vina Ermagan
Cisco Systems
170 Tasman Drive
San Jose, California 95134
USA

Email: vermagan@cisco.com

Yves Hertoghs
Cisco Systems
6a De Kleetlaan
Diegem 1831
Belgium

Phone: +32-2778-435
Fax: +32-2704-6000
Email: yves@cisco.com

Dino Farinacci
lispers.net

Email: farinacci@gmail.com

Michael Smith
Insieme Networks
California
USA

Email: michsmit@insiemenetworks.com