Enhancing the Internet Architecture:

LISP - Locator/Id Separation protocol

Luigi Iannone
Tutorial Road Map

• Why LISP??
• LISP Principles
• LISP Data Plane
  - RFC 6830
• LISP Control Plane
  - LISP-MS (RFC 6833)
  - LISP-NERD (RFC 6837)
  - LISP+ALT (RFC 6836)
  - LISP-DDT (draft-ietf-lisp-ddt-03.txt)
• LISP-Versioning
  - RFC 6834
• LISP-Interworking
  - RFC 6832
Why LISP ???
The origins: back in 70s
Do you ever heard of HOSTS.TXT ????

• In the 70s HOSTS.TXT contained all translation from hosts names to IP addresses
• Same approach as routing today:
  - everybody knows everybody
• In the early 80s it was a nightmare to maintain
• ... and then came DNS (Domain Name System)
  - On demand
  - Distributed
  - Hierarchical
  - .... in other words scalable
BGP = HOSTS.TXT ????

 IPv4

 IPv6

 2008's Economic Backdrop

 Commercial Internet

 DotCom Bubble

 CIDR

 Growth Fear!!!

http://bgp.potaroo.net/as2.0/bgp-active.html
Root of the BGP’s FIB inflation problem
No… is not just natural Internet growth

BGP Forwarding Information Base (FIB) and Churn Explosion:
PI (Provider Independent) prefix assignment
Multi-homing
Traffic-Engineering
Security
Remember the YouTube incident?

Churn can have peaks of thousands per seconds
Churn increases the need processing power
So where the problem comes from???

“Addressing can follow topology
or
topology can follow addressing. Choose one.”

Rekhter’s Law
The overloaded IP address Semantic

• An IP Address tells:
  – Who you are
    – Hi! I am Luigi Iannone RIPE 70….
  – Where you are
    – …and you can find me at Luigi Iannone RIPE 70

• This design was OK in the 70s-80s
  – Because was easier to implement
  – Because the Internet was a small academic network of networks
Internet’s Scaling Issues


“It is commonly recognized that today’s Internet routing and addressing system is facing serious scaling problems.”
“The Research Group has rough consensus that separating identity from location is desirable and technically feasible. However, the Research Group does NOT have consensus on the best engineering approach to such an identity/location split.”

Along with a plethora of proposals:

Locator/ID Separation Protocol (LISP)
Routing Architecture for the Next Generation Internet (RANGI)
Internet Vastly Improved Plumbing (IPv6)
Hierarchical IPv4 Framework (hIPv4)
Name Overlay (NOL) Service for Scalable Internet Routing
Compact Routing in a Locator Identifier Mapping System (CRM)
Layered Mapping System (LMS)
Two-Phased Mapping
Global Locator, Local Locator, and Identifier Split (GLI-Split)
Tunneled Inter-Domain Routing (TIDR)
Identifier-Locator Network Protocol (ILNP)
Enhanced Efficiency of Mapping Distribution protocols in Map-and-Encap Schemes (EEMDP)
Evolution
Name-Based Sockets
Routing and Addressing in Networks with Global Enterprise Recursion (IRON-RANGER)
Hierarchical Architecture for Internet Routing (HAIR)
LISP Principles
Who are all those prefixes?

- Number of Active ASes: 48349
- Number of Origin Only ASes: 41274 (85%)
- Average entries per Origin AS: ~11
- Roughly ~454 000 Prefixes are Stub Networks

[http://bgp.potaroo.net/as2.0/bgp-active.html](http://bgp.potaroo.net/as2.0/bgp-active.html)
The Separated Core

Internet Routing Location
The Separated Edge
Packets in Core/Edge Separation

Oracle

Internet (DFZ)

AS_s

AS_z

AS_k

AS_d

Core (Push Routing Model)
Edge (Pull Routing Model)
Payload
From Core/Edge Separation to LISP

**Internet** (DFZ)

- **Core** (Push Routing Model)
  - **RLOC**
- **Edge** (Pull Routing Model)
  - **EID**

**Mapping System**

- Oracle System

**Routing**

- **Routing LOC** (Push Routing Model)
- **Endpoint ID** (Pull Routing Model)

**Payload**
LISP Data Plane

RFC 6830
At the border of ID and Location

**Identifier Space**

**Locator Space**

Internet (DFZ)

**Bindings between Identifier and Locator Spaces: Mappings**

**ITR/ETR**
Ingress/Egress Tunnel Routers (xTRs)
Life of a Packet in a LISP World..

Bindings between ID and Locators: *Mappings*

![Diagram showing the life of a packet in a LISP world.](image-url)
Map & Encap Operations: source side

LISP-Database:
- Contains mappings “owned” locally
- Used to select source RLOC

\[ EID_s - Prefix \Rightarrow (RLOC_{EID_s}^1, RLOC_{EID_s}^2) \]
Where does LISP find the Mappings?

- Mapping Distribution System:
  - Queried to retrieve mappings
  - Used to select Destination RLOC

EID_d – Prefix \( \Rightarrow (RLOC^{1}_{EID_{d}}, RLOC^{2}_{EID_{d}}) \)

Where is located EID_d?
Where does LISP store the Mappings?

LISP-Cache:
- Queried before the Mapping system
- Mapping system queried only in case of miss
- Used to select Destination RLOC
Map & Encap Operations: destination side

Consistency Checks:
- Check DB: Am I the correct RLOC for the destination EID?
LISP Header

RFC 6830
LISP Header
LISP Header: Nonce

| L | N | E | 0 | | | Nonce |
| I | +---------------------------------+ |
| S \ | | |
| P | +---------------------------------+ |

- **N-bit**: Nonce Present bit
  - V-bit must be 0
- **E-Bit**: Echo-Nonce Request Bit
- **Echo-Nonce Algorithm**:
  - E-Bit & N-Bit Set
  - Echo back the Nonce
  - Purpose: Reachability
LISP Header: Nonce in practice
Echo-Nonce may be tricky….

• How to understand the difference????
LISP Header: Loc-Status

- L-bit: Locator-Status Present bit
  - L-bit must be 0

- Purpose: provide in-band information about RLOC status/reachability
  - How to define the status/reachability?
LISP Header: Loc-Status in practice
LISP Header: Instance ID

- I-bit: Instance ID Present bit
  - L-bit must be 0

- Second long-word split in two fields:
  - 24 high-order bits: Instance ID
  - 8 low-order bits: Loc-Status bits

- Purpose: Virtualization (VPN) and Segmentation
LISP Header: Instance ID in practice

- ETR uses IID as VRF selector
**LISP Header: Versioning**

- **V-bit**: Map-Version Present bit
  - E-bit must be 0
- **RFC 6834 (later)**
LISP Map-Server
RFC 6833
MS: A Mapping System API

\[ EID_d - Prefix \Rightarrow \]
Map-Reply(EID: \(<RLOC_1, Priority, Weight; \ldots \>)\)

Map-Register(EID: \(<RLOC_1, Priority, Weight; \ldots \>)\)

Map-Notify(EID: \(<\ldots\>\)\)

Map-Resolver

Map-Server

Distribution System
(ALT, CONS, EMACS, DHT, NERD, TREE => DDT)

Where is located EID\(_d\)?

RLOC\(_{EID_s}^2\)

RLOC\(_{EID_d}^2\)

DB

RLOC\(_{EID_d}\)

EID\(_d\)

AS\(_x\)

AS\(_y\)

AS\(_s\)

EID\(_s\)
LISP Control Message Outer Header

- RLOC space
- UDP Destination port 4342

- Reserved: 0 b'0000'
- LISP Map-Request: 1 b'0001'
- LISP Map-Reply: 2 b'0010'
- LISP Map-Register: 3 b'0011'
- LISP Map-Notify: 4 b'0100'
- LISP Encapsulated Control Message: 8 b'1000'
Map-Request
### Map-Request: Flags

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Type=1</td>
<td>A</td>
<td>M</td>
</tr>
<tr>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Nonce . . .</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Source-EID-AFI</td>
<td>Source EID Address ...</td>
</tr>
<tr>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ITR-RLOC-AFI 1</td>
<td>ITR-RLOC Address 1 ...</td>
</tr>
<tr>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>ITR-RLOC-AFI n</td>
<td>ITR-RLOC Address n ...</td>
</tr>
<tr>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Reserved</td>
<td>EID mask-len</td>
</tr>
<tr>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>EID-Prefix ...</td>
<td></td>
</tr>
<tr>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>Map-Reply Record ...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>+---------------------------------------------------------------+</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

- **A** - Authoritative bit: Request authoritative answer.
- **M** - Map Data bit: Indicates Map-Reply record present
- **P** - Probe bit: Message to be used as locator reachability probe (sent directly to xTRs)
- **S** - Solicit Map Request: Solicit Map Request
- **p** - PITR bit: set when PITR sends the request (RFC 6832)
- **s** - SMR-invoked Map-Request: set when Map-Request has been solicited by SMR.
Map-Request: The not flags part.....

- IRC - ITR-RLOC Count: number of <ITR-RLOC-AFI, ITR-RLOC-Addr> present
- Record Count: Number of actual Requests in the message
- Nonce: Security Random value
- Source EID (AFI & Addr)
- Rec - AFI, Prefix Mask Length, EID prefix request
Map-Reply: Flags + Nonce + RecCount

- **P** - Probe bit: Message is a reply to a locator reachability probe (sent directly to xTRs)
- **E** - Echo None Bit: ETR is Echo-Nonce Capable
- **S** - Security bit: If present security material is appended to the packet
- **Record Count**: number of mappings
- **Nonce**: Security Random value
Map-Reply: Record

- TTL - Time To Live: How many minutes the mapping can be stored
  - 0: Remove
  - 0xFFFFFFFF: local decision
- Loc Count: Number of locators
- EID Mask Length
- ACT: Action bits (No-Action; Forward Native; Send-Map-Request; Drop)
- A - Authoritative Bit: set if the sender is an ETR
Map-Reply: Record

- Map Version Number
- EID (AFI & Prefix)
- Priority & Weight (also for multicast)
- L: Local Locator
- p: Probed Locator
- R: Route for Locator
- Locator (AFI & Addr)
Map-Register

- P - Proxy Bit
- M - Map-Notify requested:
### Map-Notify

<table>
<thead>
<tr>
<th>0</th>
<th>1</th>
<th>2</th>
<th>3</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>4</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>8</td>
<td>9</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>2</td>
<td>3</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td>6</td>
<td>7</td>
<td>8</td>
<td>9</td>
</tr>
<tr>
<td>0</td>
<td>1</td>
<td>2</td>
<td>3</td>
</tr>
</tbody>
</table>

```
+---------------------------------+
| Type=4 | Reserved | Record Count |  
+---------------------------------+
| Nonce . . .                       |
+---------------------------------+
| Key ID | Key ID | Authentication Data Length |
+---------------------------------+
| Authentication Data              |
```

```
+---------------------------------+
| Authentication Data              |
```

```
+---------------------------------+
| Record TTL                      |
+---------------------------------+
| Locator Count | EID mask-len | ACT |A|      Reserved         |
+---------------------------------+
| EID-Prefix-AFI | EID-Prefix-AFI |
+---------------------------------+
| EID-Prefix | EID-Prefix |
+---------------------------------+
| Priority | Weight | M Priority | M Weight |
+---------------------------------+
| Loc-AFI | Loc-AFI |
+---------------------------------+
LISP Control Plane

• LISP+ALT (RFC 6836)
• LISP-NERD (RFC 6837)
• LISP-DDT (draft-ietf-lisp-ddt-01.txt)
LISP NERD

RFC 6837
LISP-NERD

Not-so-novel Eid to Rloc Database

- Centralized approach based on one (or more) authority(ies)
  - xTRs periodic update request
  - Updates contain all mappings

- Pros
  - Simple
  - Bootstrap
  - Data Center
  - Private Deployments

- Cons
  - Not scalable
  - Static
LISP+ALT

- BGP over GRE overlay
  - EID-Prefix (aggregated) advertisements
  - Map-Request routed on the overlay
- Pros
  - Well-known technology
- Cons
  - PITA
LISP+ALT Failure

- Used on an international testbed (www.lisp4.net)
  - Rapidly resulted in very cumbersome maintenance

www.lisp4.net
LISP-DDT: Delegated Database Tree

- DNS-Like logical hierarchy
  - EID-Prefix more specific delegated
  - Delegation is statically configured

- Pros
  - Easy maintenance
  - Economics
  - Security

- Cons
  - Manual configuration
  - Static tree
LISP-DDT

- **MR**: Static pointer to root
- **DDT Nodes**: Static pointers to children authoritative of more-specific (leaves are MS)
- **ETR**: Registers to statically configured MS
1. ITR sends Map-Request to its configured MR
2. MR sends Map-Request to configured Root Server
3. Root Server sends back Map-Referral to configured DDT Node authoritative on more specific
4. Repeat 2 and 3 until MS reached
5. MR sends Map-Request to authoritative MS
6. MS forward Map-Request to authoritative ETR
7. MS sends Map-ACK to MR
8. ETR sends Map-Reply with requested mapping to ITR
1. MR sends Map-Request to *cached* authoritative MS

2. MS sends Map-ACK to MR

3. MS forward Map-Request to authoritative ETR

4. ETR sends Map-Reply with requested mapping to ITR
LISP-DDT: Mapping Retrieval (Proxy Case)

1. MR sends Map-Request to cached authoritative MS
2. MS act as proxy and sends Map-Reply to MR
3. MR sends Map-Reply with requested mapping to ITR
**Map-Referral**

| 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 0 | 1 |
| +---------------------------------------------------------------------+ |
| | Type=6 | Reserved | Record Count | |
| +---------------------------------------------------------------------+ |
| | Nonce . . . | |
| +---------------------------------------------------------------------+ |
| | . . . Nonce | |
| +---------------------------------------------------------------------+ |
| | Record TTL | |
| +---------------------------------------------------------------------+ |
| | Referral Count| EID mask-len | ACT |A|I| Reserved | |
| +---------------------------------------------------------------------+ |
| | SigCnt | Map Version Number | EID-AFI | |
| +---------------------------------------------------------------------+ |
| | EID-prefix ... | |
| +---------------------------------------------------------------------+ |
| | Priority | Weight | M Priority | M Weight | |
| +---------------------------------------------------------------------+ |
| | Unused Flags | R | Loc/LCAF-AFI | |
| +---------------------------------------------------------------------+ |
| | Locator ... | |
| +---------------------------------------------------------------------+ |

- **Action Bits:**
  - Node Referral (0): sent by authoritative node
  - MS-Referral (1): DDT-Node sending the referral is not a Map Server.
  - MS-Ack (2): Sent by DDT Map Server
  - MS-Not-Registered (3): Sent by a DDT Map Server with no ETRs registered
  - Delegation-Hole (4): No child delegation for the EID. Also sent by a DDT Map Server
  - Not-Authoritative (5): Sent by a DDT node not authoritative for the requested EID
    - May send incomplete referrals (bit I set)
LISP Header: Versioning

- V-bit: Map-Version Present bit
  - E-bit must be 0
- RFC 6834 (Now)
- LISP Map Versioning:
  - Bind a Version number to mappings
  - Embed Map Version in LISP header
The problem of Mapping “Freshness”

- Mappings can change due to:
  - Mobility (for some definition of it….)
  - TE
  - Outage....
LISP Versioning
The problem of Mapping “Freshness”

To update a mapping in a cache there is always a Map-Request/Map-Reply exchange.
LISP Interworking

RFC 6832
Interworking LISP and the Legacy Internet

• How to forward packets from an ITR to a non-LISP site?
• How to forward packets from a non-LISP site to an EID?
Proxy-xTRs (PxTRs)

- **Proxy-ITR (PITR)** - Gateway from Legacy to LISP Internet
  - Announce large aggregates of EID-Prefixes into the legacy BGP
- **Proxy-ETR (PETR)** - Gateway from LISP to Legacy Internet
  - Could be a static configuration (similar to default route)
  - Could be obtained from Mapping System
PxTRs and asymmetric traffic flow
Few Pointers...

- www.lisp-lab.org
- datatracker.ietf.org/wg/lisp/
- www.openlisp.org
- www.lisp4.net
Questions?