Vers un renforcement de l’architecture Internet : le protocole LISP ("Locator/ID Separation Protocol")

JCSA 2013

Luigi Iannone
Road Map

- Why LISP???

- LISP Data Plane
  - RFC 6830

- LISP Control Plane
  - RFC 6833; RFC 6836; draft-ietf-lisp-ddt-01.txt

- The Big Picture
Why LISP?????
Internet’s Scaling Issues

“It is commonly recognized that today’s Internet routing and addressing system is facing serious scaling problems.”

The BGP’s FIB inflation problem
No... it 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

01-Jan-94 to 14-March-11 Source: http://bgp.potaroo.net/index-bgp.html

IPv4
IPv6
Cause: The overloaded IP address semantic

- An IP Address tells:
  
  • Who you are
    — Hi! I am Luigi Iannone 46 Rue Barreault 75013 Paris France

  • Where you are
    — .... and I am in Luigi Iannone 46 Rue Barreault 75013 Paris France

- This design was OK in the 70s-80s
  
  • Because was easier to implement
  • Because the Internet was a small academic network of networks
“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 (Ipv)
- 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)

From RFC 6115: Recommendation for a Routing Architecture
LISP: Locator/ID Separation Protocol

- **RFC 6830**: The Locator/ID Separation Protocol (LISP)
- **RFC 6831**: The Locator/ID Separation Protocol (LISP) for Multicast Environments
- **RFC 6832**: Interworking between Locator/ID Separation Protocol (LISP) and Non-LISP Sites
- **RFC 6833**: Locator/ID Separation Protocol (LISP) Map-Server Interface
- **RFC 6834**: Locator/ID Separation Protocol (LISP) Map-Versioning
- **RFC 6835**: The Locator/ID Separation Protocol Internet Groper (LIG)
- **RFC 6836**: Locator/ID Separation Protocol (LISP) Alternative Logical Topology (LISP+ALT)
- **RFC 6837**: NERD: A not-so-novel Endpoint ID (EID) to Routing Locator (RLOC) Database
LISP Data Plane (RFC 6830)
Splitting ID and Location

Locator Space

Identifier Space

ITR/ETR
Ingress/Egress Tunnel Routers (xTRs)
Map & Encap Operations: source side

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

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

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

EID_d - Prefix \( \Rightarrow (RLOC_{EID_d}^1, RLOC_{EID_d}^2) \)

Where is located EID_d?

DB

AS_x

AS_y

AS_z

AS_d

EID_s

RLOC_{EID_d}^1

RLOC_{EID_d}^2

Internet

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

### Enapsulation

<table>
<thead>
<tr>
<th>0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1</th>
</tr>
</thead>
</table>
| +------------------------------------------+
<p>| | Version | IHL | Type of Service | Total Length |   |
| | +---------+-----+-----------------+--------------+---|
| | Identification | Flags | Fragment Offset |   |   |
| | +-----------------+--------------+-----------------+---+---|
| O | Time to Live | Protocol = 17 | Header Checksum |   |   |
| | | Source Routing Locator |   |   |
| | +-------------------------------+---+---|
| | | Destination Routing Locator |   |   |
| | +-----------------------------+---+---|
| L | | Source Port = xxxx | Dest Port = 4341 |   |   |
| | | UDP | UDP Length | UDP Checksum |   |   |
| | | | | |   |   |</p>
<table>
<thead>
<tr>
<th>0 1 2 3 4 5 6 7 8 9 0 1 2 3 4 5 6 7 8 9 0 1</th>
</tr>
</thead>
</table>
| +------------------------------------------+
| | Version | IHL | Type of Service | Total Length |   |
| | +---------+-----+-----------------+--------------+---|
| | Identification | Flags | Fragment Offset |   |   |
| | +-----------------+--------------+-----------------+---+---|
| O | Time to Live | Protocol | Header Checksum |   |   |
| | | Source Routing Locator |   |   |
| | +-------------------------------+---+---|
| | | Destination Routing Locator |   |   |
| | +-----------------------------+---+---|

### Consistency Checks:

- Check DB: Am I the correct RLOC for the destination EID?
LISP Mapping Distribution System

LISP-MS (RFC 6833)
LISP+ALT (RFC 6836)
LISP-DDT (draft-ietf-lisp-ddt-01.txt)
MS: A Mapping System API

Map-Request(EID)

Map-Reply(EID: <RLOC_{1,...}>)

Map-Register(EID: <RLOC_{1,...}>)

Map-Notify(EID: <RLOC_{1,...}>)

EID_s

RLOC_{EID_s}^2

EID_d

RLOC_{EID_d}^2

Map-Resolver

Map-Server

DB

AS_x

AS_y

AS_z

Internet

System

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

Where is located EID_d?

Map-Request(EID)

Map-Reply(EID: <RLOC_{1,...}>)

Map-Register(EID: <RLOC_{1,...}>)

Map-Notify(EID: <RLOC_{1,...}>)
Pros
- Based on existing well-known technology

Cons
- PITA

Mapping Distribution System

ALT-Router

BGP over GRE Session
LISP+ALT Failure

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

- 2nd Week March 2012
- LISP4.net switched to DDT

LISP+ALT Topology early 2012

Source: www.lispmont.net
- Pros
  - Easy maintenance
  - Economics
  - Security
- Cons
  - static manual configuration of the tree

DDT Node
...... Logical Delegation
LISP-DDT Configuration

**MR**
Static pointer to root

**DDT Nodes**
Static pointers to children authoritative of more-specific (leaves are MS)

**ETR**
Registers to statically configured MS

- **DDT Root 0/0**
- **DDT Node 10/8**
- **DDT Node 10.6/16**
- **MS 10.6.9/24**
- **ETR 10.6.9/24**

Luigi Iannone
Telecom ParisTech
LISP - JCSA 2013
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 MS
6. MS forwards Map-Request to authoritative ETR
7. ETR sends Map-Reply with requested mapping to ITR
LISP-DDT Map Retrieval (regular case)

1. ITR sends Map-Request to its configured MR
2. MR sends Map-Request to cached MS
3. MS forwards Map-Request to authoritative ETR
4. ETR sends Map-Reply with requested mapping to ITR
LISP-DDT

Map Retrieval with Proxy Reply (cold start)

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 sends back Map-Reply with requested mapping to MR
7. MR sends Map-Reply with requested mapping to ITR
Map Retrieval with Proxy Reply (regular case) start

1. ITR sends Map-Request to its configured MR
2. MR sends Map-Request to cache MS
3. MS sends back Map-Reply with requested mapping to MR
4. MR sends Map-Reply with requested mapping to ITR
# LISP DDT Topology (early 2013)

<table>
<thead>
<tr>
<th>Level</th>
<th>Node 1</th>
<th>Node 2</th>
<th>Node 3</th>
</tr>
</thead>
<tbody>
<tr>
<td>DDT ROOT</td>
<td>ARIN-DDT</td>
<td>RIPE-DDT</td>
<td>VXNET-DDT</td>
</tr>
<tr>
<td>LISP BETA TLD</td>
<td>SJ-DDT</td>
<td>UNINETT-DDT</td>
<td>MSN-DDT</td>
</tr>
<tr>
<td>ARIN</td>
<td>CISCO-SJC-MR-MS-1</td>
<td>ASP-MR-MS</td>
<td>EQX-ASH-PXTR</td>
</tr>
<tr>
<td></td>
<td>ISC-MR-MS</td>
<td>EQX-ASH-MR-MS</td>
<td>ISC-PXTR</td>
</tr>
<tr>
<td></td>
<td>ISC-MN-RTR</td>
<td>ASP-MN-RTR</td>
<td>ASP-PXTR</td>
</tr>
<tr>
<td>RIPE</td>
<td>L3-LONDON-MR-MS</td>
<td>INTOUCH-MR-MS-1</td>
<td>INTOUCH-PXTR-1</td>
</tr>
<tr>
<td></td>
<td>TDC-MR-MS</td>
<td>INTOUCH-MR-MS-2</td>
<td>INTOUCH-PXTR-2</td>
</tr>
<tr>
<td></td>
<td>INTOUCH-MN-RTR</td>
<td></td>
<td>SG-NUS-PXTR</td>
</tr>
<tr>
<td>APNIC</td>
<td>APNIC-MR-MS</td>
<td></td>
<td>APAN-PXTR</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>TBD-PXTR</td>
</tr>
<tr>
<td>LACNIC</td>
<td>LACNIC-MR-MS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
LISP-DDT Reliability

The diagram illustrates the percent of answered Map-Repplies across different resolvers. The x-axis represents the range of resolvers, while the y-axis shows the percent of answered replies. There are two markers: 'first week' and 'last week,' indicating the performance comparison between these two periods.

The number of requests for the first and last week is also shown on the right side of the diagram, with the y-axis denoting the number of requests ranging from 1e+06 to 1e+00.
LISP-DDT RTT

EUROPE

US

PlanetLab

NOT PlanetLab

Luigi Iannone
Telecom ParisTech
LISP - JCSA 2013
27
The Big Picture....

Does LISP finally solve (all of) the issues previously mentioned???
What the Internet looks like with LISP

- Simulations
  - Synthetic topologies generated with GHITTLLE
  - Hierarchical with business relationship
  - 14965 ASes
  - Gain: 1 order of magnitude

This work has been carried out in collaboration with Prof. B. Quoitin @ UMons
What the Internet looks like with LISP

Less than 20% of Today’s BGP RIB (roughly ~ 100,000 entries)
Large Scale LISP Emulation

- LISP Emulation
  - Based on PCAP Traces (> 20,000 DSL customers)
  - Used BGP Granularity for mappings

Cache Entries Evolution for Vanilla LISP

24h Cache Size Evolution

Cache Miss Classification

Cache Entries Lifetime
What the Internet looks like with LISP

- Depends on Cache Management policy (but we can easily go for less than 50,000 entries)

- Less than 20% of Today’s BGP RIB (roughly ~ 100,000 entries)

Currently lisp4.net has a DFZ potential reduction of 93.8%!
LISP main benefits

- FIB (Forwarding Information Base)
  - shrunk since stub AS do not inject anything in the DFZ (Default Free Zone)
- PI Addresses
  - Just a matter of changing RLOCs
- Multi-Homing
  - Just a list of RLOCs
- Traffic Engineering
  - Just a matter of giving priorities to RLOCs
- Churn
  - Flapping Edge AS are not anymore in the DFZ
- Additional Benefits
  - Path Diversity
But there is much more

- Mobility in Wireless Mesh Networks
- Smartphones seamless connectivity
- WiFi Offloading
- MaaS: Mobility as a Service
- Cloud Networking
  - VM migration
  - Load Balancing
  - VPN
  - Content Distribution
- Real-Time Traffic Optimization

- ... and more to come....
Thank You!

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

- Coming soon...

Lisp-Lab
Building Today the Internet of Tomorrow

Lisp-View