Flexible Data Centre Fabric - FabricPath/TRILL, OTV, LISP and VXLAN

Ron Fuller– CCIE #5851 (R&S/Storage)
Technical Marketing Engineer, Nexus 7000
rfuller@cisco.com
Evolution of the Data Center Fabric

Agenda

- The Evolving Data Centre Fabric
- FabricPath
- VXLAN
- LISP
- LISP Host Mobility
- OTV LAN Extension
- Mobility with Extended Subnets
- Nexus Fabric
Goals of the Fabric
Addressing Concurrent Workloads, Mobility and Latency

Port Density
Adequate Buffer Capacity
Adequate Table Sizes
Low Latency Switching
Cut-through Switching

Architecture is evolving Rapidly – in the next 24 months
L2/L3 Boundary becomes less relevant
Clos Topologies dominate new implementations
HA models shift
Server Edge becomes more intelligent
DC Fabric becomes more scalable

Priority Flow Control
Early Congestion Notification
FabricPath Multiple Trees
ECMP L2 & L3
Multi-tenancy
Goals of the Fabric
Addressing High Availability and Fate Sharing

East-West traffic – Fate Sharing Domain
STP is the protocol of choice
1+1 redundancy – limited forwarding paths

East-West across L3 boundaries
OSPF/EIGRP are protocols of choice
N+1 redundancy – Broad forwarding Paths

North-South traffic
OSPF/EIGRP are protocols of choice
N+1 redundancy – Broad forwarding paths

Larger POD East-West Traffic – Fate Sharing Domain
N+1 redundancy
IS-IS is the protocol of choice
Broad forwarding paths
Broader Adjacency Support

Same number of physical boxes and links
Protocol behavior is L3-like
Multi-pathing over L2 and L3
More flexible L2 adjacency, better scale capacity
Better latency consistency within POD
The traditional L2 vs. L3 debate has been based on a number of issues:

- Scalability
- Availability

Requirements for the scalable design moving forward is a scalable, highly available switching fabric with the advantages of both L2 and L3.
Advantages and Drawbacks of Layer 2
“Plug-and-Play” and Mobility vs. Availability and Scaling

Advantages of Layer 2

- Practically “plug-n-play” – No user configuration is required to build forwarding database
- It makes it simple to support teaming or L2 multicast for clusters
- Easy to segment traffic with VLANs
- Very fast movement of end station addresses (ability to update MAC address tables after a vMotion-type event)

Disadvantages of Layer 2

- MAC address consumption
- BPDU generation is CPU intensive with increasing number of VLANs
- VLAN sprawl causes flooding and broadcasts to propagate even where they are not needed
- Half of the links in the topology are blocking
- Misconfigurations can cause Layer 2 loops which may make switches unmanageable
Advantages and Drawbacks of Layer 3
Availability and Scaling vs. Restricted Workload Flexibility

- Layer 3 Routed Topologies alleviate the consumption of L2 tables via route summarization
- Layer 3 Routed topologies provide for a degree of fault isolation and
- “Routed Access” provides the logical extension of the design philosophy
- “Scaling Up” of the Access Switch via such mechanism as the FEX provide a degree of workload mobility
- “L2” domain extension of some form is required for most workload mobility requirements

[Diagram: Workload Domain for most Hypervisor and Clustering based solutions is restricted by the Traditional Layer 2/3 boundary]
Fabric Technologies
Segment-ID: Scaling Logical Groupings of Connectivity

Web Server
App Server
Database Server

S1 S2 S3

802.1Q VLAN ID 12-bits
802.1Q VLAN ID 12-bits
802.1ad standardized frame format

SegmentId 24-bits

© 2010 Cisco and/or its affiliates. All rights reserved.
### Fabric Technologies

#### Location Identity Separation

- Location reachability determined by traditional routing mechanisms in the Fabric
- Identity is mapped to location addresses
- All these technologies leverage Location/Identity Mapping

<table>
<thead>
<tr>
<th></th>
<th>FabricPath / TRILL</th>
<th>VXLAN</th>
<th>OTV</th>
<th>LISP</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Location</strong></td>
<td>Switch-ID (IS-IS)</td>
<td>IP address (IP protocols)</td>
<td>IP address (IP protocols)</td>
<td>IP address (IP protocols)</td>
</tr>
<tr>
<td><strong>Identity</strong></td>
<td>Client MAC (Flooding)</td>
<td>Client MAC (Flooding)</td>
<td>Client MAC (IS-IS)</td>
<td>Client IP/MAC (Mapping DB)</td>
</tr>
<tr>
<td><strong>Multi-tenancy</strong></td>
<td></td>
<td></td>
<td></td>
<td>24-bit Segment Identifier</td>
</tr>
</tbody>
</table>
Fabric Technologies
FabricPath, LISP, VXLAN & OTV

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Intra-DC</th>
<th>Inter-DC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer 2 connectivity</td>
<td>FabricPath/TRILL/VXLAN</td>
<td>OTV/VPLS</td>
</tr>
<tr>
<td>IP Mobility</td>
<td>LISP</td>
<td>LISP</td>
</tr>
<tr>
<td>Secure Segmentation</td>
<td>VXLAN / Segment-ID</td>
<td>VPNs (LISP/MPLS)</td>
</tr>
</tbody>
</table>

Scale

IP Network

Fabric Path (Intra-DC L2)

VXLAN/OTV (Inter-DC x-L3)

OTV/VPLS (Inter-DC x-L3)

Fabric Path (Intra-DC L2)

VXLAN/OTV (Intra-DC x-L3)
Evolution of the Data Center Fabric

Agenda

- The Evolving Data Centre Fabric
- FabricPath
- VXLAN
- LISP
- LISP Host Mobility
- OTV LAN Extension
- Mobility with Extended Subnets
- Nexus Fabric
Cisco FabricPath
NX-OS Innovation Enhancing L2 with L3

“FabricPath brings Layer 3 routing benefits to flexible Layer 2 bridged Ethernet networks”
### MAC-in-MAC
- Creates hierarchical layer 2 address scheme with additional MAC header
- Source and destination Switch_ID written into outer MAC header at L2MP edge
- Forwarding inside L2MP core network is based on destination Switch_ID
- Embedded path selector (FTAG) provides multi-pathing for even broadcast and multicast
- Built-in protections (TTL and multicast RPF) minimize impact of transient network issues

### Optimal MAC Learning
- Prevent potential MAC table overflow in large scale L2 domain
- Traditional source-learning only on Edge port for locally connected MAC addresses
- Learning is disabled on Core port to reduce MAC table utilization
- Non-local source-MAC only learned if destination-MAC is already learned as local entry

### IS-IS
- Scalable routing protocol with proven implementation for fast convergence upon network changes
- Link-state protocol ensures optimal path between any 2 nodes
- Built-in authentication mechanism enhances network security and stability
- Inherent support for ECMP and multi-topology maximize link utilization
FabricPath
New Control Plane

Plug-n-Play L2 IS-IS manages forwarding topology

- IS-IS assigns addresses to all FabricPath switches automatically
- Compute shortest, pair-wise paths
- Support equal-cost paths between any FabricPath switch pairs

FabricPath Routing Table

<table>
<thead>
<tr>
<th>Switch</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>S10</td>
<td>L1</td>
</tr>
<tr>
<td>S20</td>
<td>L2</td>
</tr>
<tr>
<td>S30</td>
<td>L3</td>
</tr>
<tr>
<td>S40</td>
<td>L4</td>
</tr>
<tr>
<td>S200</td>
<td>L1, L2, L3, L4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>S400</td>
<td>L1, L2, L3, L4</td>
</tr>
</tbody>
</table>

© 2010 Cisco and/or its affiliates. All rights reserved.
Classical Ethernet (CE)

- The association MAC address/Switch ID is maintained at the edge
- Core fabric leverages an independent routing topology from the edge
- Scales MAC learning
- Scales Core topology state

Switch ID space: Routing decisions are made based on the FabricPath routing table

MAC address space: Switching based on MAC address tables

FabricPath (FP)

S300: FabricPath Routing Table

<table>
<thead>
<tr>
<th>Switch</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>S100</td>
<td>L1, L2, L3, L4</td>
</tr>
</tbody>
</table>

S300: CE MAC Address Table

<table>
<thead>
<tr>
<th>MAC</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>B</td>
<td>1/2</td>
</tr>
<tr>
<td>A</td>
<td>S100</td>
</tr>
</tbody>
</table>
FabricPath
New Control and Data Plane

- Edge switches maintain both MAC address table and Switch ID table
- Ingress switch uses MAC table to determine destination Switch ID
- Egress switch uses MAC table (optionally) to determine output switchport

<table>
<thead>
<tr>
<th>MAC</th>
<th>IF/SID</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>e1/1</td>
</tr>
<tr>
<td>B</td>
<td>e1/2</td>
</tr>
<tr>
<td>C</td>
<td>S101</td>
</tr>
<tr>
<td>D</td>
<td>S200</td>
</tr>
</tbody>
</table>

Local MACs point to switchports
Remote MACs point to Switch IDs
FabricPath
New Control and Data Plane

- FabricPath IS-IS manages Switch ID (routing) table
- All FabricPath-enabled switches automatically assigned Switch ID (no user configuration required)
- Algorithm computes shortest (best) paths to each Switch ID based on link metrics
- Equal-cost paths supported between FabricPath switches

<table>
<thead>
<tr>
<th>Switch</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>S10</td>
<td>L1</td>
</tr>
<tr>
<td>S20</td>
<td>L2</td>
</tr>
<tr>
<td>S30</td>
<td>L3</td>
</tr>
<tr>
<td>S40</td>
<td>L4</td>
</tr>
<tr>
<td>S101</td>
<td>L1, L2, L3, L4</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>S200</td>
<td>L1, L2, L3, L4</td>
</tr>
</tbody>
</table>

One ‘best’ path to S10 (via L1)
Four equal-cost paths to S101
FabricPath
Scaling – Conversational Learning

- Edge switch only learn the MAC of remote hosts when there are two way communications between remote hosts and local hosts
- Unknown unicast flooding alone won’t have all switches within VLAN learn the source MAC
- Intermediate switches don’t learn the MAC
- Hardware based MAC learning
Cisco FabricPath Terminology

- **CE Edge Ports**: Interface connected to traditional network device, sends/receives traffic in standard 802.3 Ethernet frame format, participates in STP domain, forwarding based on MAC table.

- **FP Core Ports**: Interface connected to another FabricPath device, sends/receives traffic with FabricPath header, does not run spanning tree, does not perform MAC learning, exchanges topology info through L2 ISIS adjacency, forwarding based on ‘Switch ID Table’.

- **Spine Switch**

- **Leaf Switch**

- **Classical Ethernet (CE)**

- **FabricPath (FP)**
VLAN Pruning
Configuration Simplicity

Automatically handled by IS-IS
FabricPath Multidestination Trees

- Multidestination traffic constrained to loop-free trees touching all FabricPath switches
- Root switch assigned for each multidestination tree in FabricPath domain
- Loop-free tree built from each Root and assigned a network-wide identifier (Ftag)
- Support for multiple multidestination trees provides multipathing for multi-destination traffic

Two trees supported in NX-OS release 5.1
Scaling the Segments within the Fabric Multi-Topology Support

- Extending FabricPath to the edge switches without requiring a redesign of the VLAN topology.
- Each FP switch can have up to 2 Topology ID’s defined (Topology ID’s does not have to be unique).
- Each Topology will have 2 Multi-Destination Trees defined.
FabricPath
Mac-in-Mac Header

Classical Ethernet Frame

Cisco FabricPath Frame

- **Switch ID** – Unique number identifying each FabricPath switch
- **Sub-Switch ID** – Identifies devices/hosts connected via VPC+
- **LID** – Local ID, identifies the destination or source interface
- **Ftag** (Forwarding tag) – Unique number identifying topology and/or distribution tree
- **TTL** – Decremented at each switch hop to prevent frames looping infinitely
Putting it all together – Host A to Host B
(1) Broadcast ARP Request

**FabricPath MAC Table on S100**

<table>
<thead>
<tr>
<th>MAC</th>
<th>IF/SID</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>e1/13 (local)</td>
</tr>
</tbody>
</table>

**Multidestination Trees on Switch 100**

<table>
<thead>
<tr>
<th>Tree</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>po10</td>
</tr>
<tr>
<td>2</td>
<td>po10, po20, po30, po40</td>
</tr>
</tbody>
</table>

**FabricPath MAC Table on S200**

<table>
<thead>
<tr>
<th>MAC</th>
<th>IF/SID</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>e1/13 (local)</td>
</tr>
</tbody>
</table>

**Multidestination Trees on Switch 10**

<table>
<thead>
<tr>
<th>Tree</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>po100, po200, po300</td>
</tr>
<tr>
<td>2</td>
<td>po10</td>
</tr>
</tbody>
</table>

**Multidestination Trees on Switch 300**

<table>
<thead>
<tr>
<th>Tree</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>po10, po20, po30, po40</td>
</tr>
<tr>
<td>2</td>
<td>po40</td>
</tr>
</tbody>
</table>

**Putting it all together – Host A to Host B**

(1) Broadcast ARP Request
Putting it all together – Host A to Host B

(1) Broadcast ARP Request

• S100:

S100# `sh mac address-table dynamic`

Legend:

* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age - seconds since last seen, + - primary entry using vPC Peer-Link

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC Address</th>
<th>Type</th>
<th>age</th>
<th>Secure</th>
<th>NTFY Ports/SWID.SSID.LID</th>
</tr>
</thead>
<tbody>
<tr>
<td>* 10</td>
<td>0000.0000.000a</td>
<td>dynamic</td>
<td>0</td>
<td>F</td>
<td>F Eth1/13</td>
</tr>
</tbody>
</table>

S100#

• S10 (and S20, S30, S40, S200, S300):

S10# `sh mac address-table dynamic`

Legend:

* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age - seconds since last seen, + - primary entry using vPC Peer-Link

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC Address</th>
<th>Type</th>
<th>age</th>
<th>Secure</th>
<th>NTFY Ports/SWID.SSID.LID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MAC A learned as local entry on e1/13

MAC A not learned on other switches
(2) Broadcast ARP Reply

Multidestination Trees on Switch 10
- **Tree 1**
  - **IF**: po100, po200, po300
- **Tree 2**
  - **IF**: po10

Multidestination Trees on Switch 100
- **Tree 1**
  - **IF**: po10, po20, po30, po40
- **Tree 2**
  - **IF**: po100, po200, po300

Multidestination Trees on Switch 300
- **Tree 1**
  - **IF**: po10, po20, po30, po40
- **Tree 2**
  - **IF**: po40

FabricPath MAC Table on S100
- **MAC A**
  - **IF/SID**: e1/13 (local)
- **MAC B**
  - **IF/SID**: 300.0.64 (remote)

FabricPath MAC Table on S300
- **MISS**
- **MAC B**
  - **IF/SID**: e2/29 (local)
Putting it all together – Host A to Host B
MAC Address Table after the first ARP frame

• **S100:**

S100# `sh mac address-table dynamic`

Legend:

* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age - seconds since last seen, + - primary entry using vPC Peer-Link

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC Address</th>
<th>Type</th>
<th>age</th>
<th>Secure NTFY Ports/SWID.SSID.LID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>* 10</td>
<td>0000.0000.000a</td>
<td>dynamic</td>
<td>90</td>
<td>F  F  Eth1/13</td>
</tr>
<tr>
<td>10</td>
<td>0000.0000.000b</td>
<td>dynamic</td>
<td>60</td>
<td>F  F  300.0.64</td>
</tr>
</tbody>
</table>

• **S300:**

S300# `sh mac address-table dynamic`

Legend:

* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age - seconds since last seen, + - primary entry using vPC Peer-Link

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC Address</th>
<th>Type</th>
<th>age</th>
<th>Secure NTFY Ports/SWID.SSID.LID</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>0000.0000.000b</td>
<td>dynamic</td>
<td>0</td>
<td>F  F  Eth2/29</td>
</tr>
</tbody>
</table>

S100 learns MAC B as remote entry reached through S300

MAC B learned as local entry on e2/29
If DMAC is known, then learn remote MAC

FabricPath Routing Table on S100

<table>
<thead>
<tr>
<th>Switch</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>S10</td>
<td>po10</td>
</tr>
<tr>
<td>S20</td>
<td>po20</td>
</tr>
<tr>
<td>S30</td>
<td>po30</td>
</tr>
<tr>
<td>S40</td>
<td>po40</td>
</tr>
<tr>
<td>S200</td>
<td>po10, po20, po30, po40</td>
</tr>
<tr>
<td>S300</td>
<td>po10, po20, po30, po40</td>
</tr>
</tbody>
</table>

FabricPath MAC Table on S100

<table>
<thead>
<tr>
<th>MAC</th>
<th>IF/SID</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>e1/13 (local)</td>
</tr>
<tr>
<td>B</td>
<td>300.0.64 (remote)</td>
</tr>
</tbody>
</table>

FabricPath Routing Table on S300

<table>
<thead>
<tr>
<th>Switch</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>S10</td>
<td>...</td>
</tr>
<tr>
<td>S20</td>
<td>...</td>
</tr>
<tr>
<td>S30</td>
<td>...</td>
</tr>
<tr>
<td>S40</td>
<td>...</td>
</tr>
<tr>
<td>S200</td>
<td>po10, po20, po30, po40</td>
</tr>
<tr>
<td>S300</td>
<td>po10, po20, po30, po40</td>
</tr>
</tbody>
</table>

FabricPath Routing Table on S300

<table>
<thead>
<tr>
<th>Switch</th>
<th>IF</th>
</tr>
</thead>
<tbody>
<tr>
<td>S10</td>
<td>...</td>
</tr>
<tr>
<td>S20</td>
<td>...</td>
</tr>
<tr>
<td>S30</td>
<td>...</td>
</tr>
<tr>
<td>S40</td>
<td>...</td>
</tr>
<tr>
<td>S200</td>
<td>po10, po20, po30, po40</td>
</tr>
<tr>
<td>S300</td>
<td>Use LID (64)</td>
</tr>
</tbody>
</table>
Putting it all together – Host A to Host B
Unicast forwarding

S100# sh mac address-table dynamic

Legend:
* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age - seconds since last seen, + - primary entry using vPC Peer-Link

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC Address</th>
<th>Type</th>
<th>age</th>
<th>Secure</th>
<th>NTFY Ports/SWID.SSID.LID</th>
</tr>
</thead>
<tbody>
<tr>
<td>* 10</td>
<td>0000.0000.000a</td>
<td>dynamic</td>
<td>90</td>
<td>F</td>
<td>F Eth1/13</td>
</tr>
<tr>
<td>10</td>
<td>0000.0000.000b</td>
<td>dynamic</td>
<td>60</td>
<td>F</td>
<td>F 300.0.64</td>
</tr>
</tbody>
</table>

S300# sh mac address-table dynamic

Legend:
* - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
age - seconds since last seen, + - primary entry using vPC Peer-Link

<table>
<thead>
<tr>
<th>VLAN</th>
<th>MAC Address</th>
<th>Type</th>
<th>age</th>
<th>Secure</th>
<th>NTFY Ports/SWID.SSID.LID</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0000.0000.000a</td>
<td>dynamic</td>
<td>30</td>
<td>F</td>
<td>F 100.0.12</td>
</tr>
<tr>
<td>• 10</td>
<td>0000.0000.000b</td>
<td>dynamic</td>
<td>90</td>
<td>F</td>
<td>F Eth2/29</td>
</tr>
</tbody>
</table>

S100 learns MAC A as remote entry reached through S100
## Putting it all together – Host A to Host B

### Unicast Forwarding

S100# `sh fabricpath route`

**FabricPath Unicast Route Table**

'\a/b/c' denotes ftag/switch-id/subswitch-id

'\[x/y\]' denotes [admin distance/metric]

ftag 0 is local ftag

subswitch-id 0 is default subswitch-id

<table>
<thead>
<tr>
<th>Topology (ftag), Switch ID, Sub-Switch ID</th>
<th>Administrative distance, routing metric</th>
<th>Route age</th>
<th>Client protocol</th>
<th>Next-hop interface(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td>0/100/0, number of next hops: 0</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0/100/0, number of next hops: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0/100/0, number of next hops: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>0/100/0, number of next hops: 1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

### FabricPath Unicast Route Table for Topology-Default

0/100/0, number of next hops: 0

via Po10, [60/0], 0 day/s 04:43:51, local

1/10/0, number of next hops: 1

via Po10, [115/20], 0 day/s 02:24:02, `isis_fabricpath-default`

1/20/0, number of next hops: 1

via Po20, [115/20], 0 day/s 04:43:25, `isis_fabricpath-default`

1/30/0, number of next hops: 1

via Po30, [115/20], 0 day/s 04:43:25, `isis_fabricpath-default`

1/40/0, number of next hops: 1

via Po40, [115/20], 0 day/s 04:43:25, `isis_fabricpath-default`

1/200/0, number of next hops: 4

via Po10, [115/40], 0 day/s 02:24:02, `isis_fabricpath-default`

via Po20, [115/40], 0 day/s 04:43:06, `isis_fabricpath-default`

via Po30, [115/40], 0 day/s 04:43:06, `isis_fabricpath-default`

via Po40, [115/40], 0 day/s 04:43:06, `isis_fabricpath-default`

1/300/0, number of next hops: 4

via Po10, [115/40], 0 day/s 02:24:02, `isis_fabricpath-default`

via Po20, [115/40], 0 day/s 04:43:25, `isis_fabricpath-default`

via Po30, [115/40], 0 day/s 04:43:25, `isis_fabricpath-default`

via Po40, [115/40], 0 day/s 04:43:25, `isis_fabricpath-default`
FabricPath Design
STP Interaction

- FabricPath domain appears as single Spanning-Tree bridge
- All FabricPath bridges share a common (static) bridge ID
  - Cisco reserved MAC c84c.75fa.6000
- STP BPDUs are not carried through the FabricPath network
- Configure all FabricPath edge switches using “spanning-tree vlan <x> root primary” (or manually configure bridge priority lower than any STP bridge)
  - Each FabricPath edge switch must be the root for all connected STP domains
  - Strongly recommended to use the same bridge priority on all FabricPath edge switches
FabricPath
L2/L3 Boundary Location

Layer 3 Boundary at the Spine

- Straightforward with two spine switches
- Considerations with more than two spines:
  - HSRP: Traffic polarized to spines on a per VLAN basis (South-North)
  - GLBP to distribute servers to different default gateways
  - Anycast FHRP future solution

Layer 3 Integration at the Leaf/Edge

- Provides a “cleaner” spine design
- Traffic distributed equally across spines (no hot spot)
- Increased number of hops to reach gateway (latency)
FabricPath L2/L3 Boundary Location
Classic Two Switch Spine

- Simplest migration from most existing designs
- The spine is also used for routing with M1/F1 in the same VDC
- Consideration – MAC Learning and Scaling
- Compared to classic ethernet designs you gain:
  - Ease of configuration
  - MAC address table increased scalability and more efficient learning
  - Traffic distribution on all uplinks
  - Possibility to offload the spine by providing direct communication paths between the edge layer devices

L3 Domain

edge/spine

Switch-id based forwarding + MAC learning for routed traffic

Conversational Learning

edge

Conversational Learning
FabricPath L2/L3 Boundary Location
Leaf/Spine/Boundary Architecture

• By separating the L3 function from the spine, the F1 card in the spine performs pure switch-id forwarding

• The L3 edge will need both M1/F1 in order to connect with Fabricpath ports to the spine

• The M1/F1 L3 edge will need to perform learning for the remote mac addresses

• L3 edge and spine can be combined in the same chassis by means of VDCs
Large Scale Fabric
Nexus Edge, Core & Boundary Nodes

Large Scale Fabric 4K VLAN’s, 128K MAC Address, 512K Routes
Nexus FabricPath
Standards Based + Cisco Extensions

- Nexus 5500, F1, F2 and all future HW are capable of IETF standards TRILL
- Support for TRILL in NX-OS is pending completion of extensions to the baseline protocol
- Multi-topology, VRRP interaction, …
FabricPath
Flexibility in the Fabric - Layer 2 Routing
Evolution of the Data Center Fabric

Agenda

- The Evolving Data Centre Fabric
- FabricPath
- VXLAN
- LISP
- LISP Host Mobility
- OTV LAN Extension
- Mobility with Extended Subnets
- Nexus Fabric
Virtual eXtensible LAN – VXLAN

- **Customer Requirement**
  Secure movement of vApps across cloud infrastructure

- **Solution: VXLAN**
  Millions of dedicated LAN segments
  Security at Scale
  vApp mobility across data centers & clouds

- **VXLAN is network friendly**
  Efficient load sharing of links (port channel)
  Supports NAT; better security controls

Virtual Extensible Local Area Network (VXLAN)

- Ethernet in IP overlay network
  - Entire L2 frame encapsulated in UDP
  - 50 bytes of overhead
- Include 24 bit VXLAN Identifier
  - 16 Million logical networks
- VXLAN can cross Layer 3 (IPv4 currently)

- Tunnel between VEMs
  - VMs do NOT see VXLAN ID
- IP multicast used for L2 broadcast/multicast, unknown unicast
- Technology submitted to IETF for standardization (Cisco, VMware, Citrix, Red Hat, Broadcom, Arista, and Others)
VXLAN Overview

- The Nexus 1000V VEMs act as the VXLAN Tunnel Endpoints (VTEP)
- Nexus 1000V uses a VMKNIC to terminate VTEP traffic
- VM to VM traffic on different access switches is encapsulated in a VXLAN header + UDP + IP
- VTEPs use multicast to deliver unknown destination VM MAC addresses to all VTEPs participating in a given VXLANs
- VM MAC to VTEP IP address mappings are gleaned from encapsulated packets
  Similar to Ethernet bridge flood and learn behavior
- Known destination VM MAC addresses are carried over point to point tunnels between VTEPs
VXLAN Data Plane Model

VTEP = VXLAN Tunnel End Point
VNI = VXLAN Network Identifier

- **IP Multicast Enabled Underlying Network**
  - **Access Switch**
  - **Bridge Domain Switch**
  - **VTEP**

**VXLAN’s IP Any Source Multicast Group (*,G)** acts as a bus for delivery to all relevant VTEPs for a given VNI (Carries unknown/broadcast/multicast frames)

- **Direct Unicast tunnels between VTEPs** (Carries known unicast frames)

**VTEP**

- **Access Switch**
- **Bridge Domain Switch**
- **End System**
VXLAN: VM to VM communication

1. IP = 192.168.0.100
   MAC = 00:0C:29:2F:32:A0

2. IP = 165.193.123.45
   MAC = 00:AA:00:23:45:67

3. S-MAC = 00:AA:00:23:45:67
   D-MAC = 00:10:11:FE:D8:D2
   S-IP = 165.193.123.45
   D-IP = 12.123.45.67

4. IP = 192.168.0.101
   MAC = 00:0C:29:F2:23:0A

S-MAC = 00:0C:29:F2:32:A0
D-MAC = 00:0C:29:F2:23:0A
S-IP = 192.168.0.100
D-IP = 192.168.0.101
VTEP Use Of IGMP

IGMP Used to Join Each VXLANs Assigned Multicast Group on Demand
VXLAN Example Data Flow
VM1 Communicating with VM2 in a VXLAN
VXLAN Example Data Flow
VM1 Communicating with VM2 in a VXLAN

VM1: VM Source MAC
VM2: Remote Host
Remote Host VXLAN IP

VM1: abc
1.1.1.1
VXLAN Example Data Flow
VM1 Communicating with VM2 in a VXLAN

<table>
<thead>
<tr>
<th>MAC Table: VEM 1</th>
<th>MAC Table: VEM 2</th>
</tr>
</thead>
<tbody>
<tr>
<td>VM Source MAC</td>
<td>Remote Host</td>
</tr>
<tr>
<td>VM1:abc</td>
<td>VXLAN IP</td>
</tr>
<tr>
<td>VM2:xyz</td>
<td>2.2.2.2</td>
</tr>
<tr>
<td>VM3:xyz</td>
<td>3.3.3.3</td>
</tr>
<tr>
<td>MAC: abc</td>
<td>MAC: xyz</td>
</tr>
</tbody>
</table>

VXLAN VMKNIC: 1.1.1.1, 2.2.2.2, 3.3.3.3
VXLAN Example Data Flow

VM1 Communicating with VM2 in a VXLAN

<table>
<thead>
<tr>
<th>VM Source MAC</th>
<th>Remote Host VXLAN IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>VM2:xyz</td>
<td>2.2.2.2</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>VM Source MAC</th>
<th>Remote Host VXLAN IP</th>
</tr>
</thead>
<tbody>
<tr>
<td>VM1:abc</td>
<td>1.1.1.1</td>
</tr>
</tbody>
</table>
Multiple VXLANs Can Share One Multicast Group
Blue & Red VXLANs Share The 239.1.1.1 Multicast Group

- Encapsulate with Blue VXLAN ID
- Multicast to Servers Registered for 239.1.1.1 Multicast Group

VM Broadcast Frames Sent to More Servers
But Broadcast Domain Respected Within VXLAN Segment

VEM Discards Since No VM with Blue VXLAN ID
Evolution of the Data Center Fabric

Agenda

- The Evolving Data Centre Fabric
- FabricPath
- VXLAN
- LISP
- LISP Host Mobility
- OTV LAN Extension
- Mobility with Extended Subnets
- Nexus Fabric
Location ID/Separation Protocol (LISP) 
Next Generation Networking Architecture

Single Network Architecture Delivers:
- VM **Mobility** (topology independent addressing)
- **Security**: VPNs/Multi-tenancy
- Route **Scalability** (on demand routing)
- IPv6 enablement,
- Routing Policy simplification

**Benefits**
- Services integrated in a single architecture
- Services can be offered across organizational boundaries (multiple providers)
- Very large scale
- Open model to integrate with cloud orchestrators

**Use-Cases**
- DCI route optimization/mobility
- Workload Portability to Cloud
- Secure Multi-tenancy across organizations
- Rapid IPv6 Deployment
- Route scaling
**LISP Use Cases**

**Consolidated Architecture with Multiple Applications**

---

**Efficient Multi-Homing**
- IP Portability
- Ingress Traffic Engineering without BGP

---

**IPv6 Transition Support**
- v6-over-v4, v6-over-v6
- v4-over-v6, v4-over-v4

---

**Multi-Tenancy and VPNs**
- Reduced CapEx/OpEx
- Large scale Segmentation

---

**Host-Mobility**
- Cloud / Layer 3 VM moves
- Segmentation
Location Identity Separation Protocol
What Do We Mean by “Location” and “Identity”? 

Today’s IP Behavior
Loc/ID “Overloaded” Semantic

When the Device Moves, It Gets a New IPv4 or IPv6 Address for Its New Identity and Location

LISP Behavior
Loc/ID “Split”

When the Device Moves, Keeps Its IPv4 or IPv6 Address. It Has the Same Identity

Device IPv4 or IPv6 Address Represents
Identity and Location

Device IPv4 or IPv6 Address Represents
Identity Only. Its Location Is Here!

Only the Location Changes
A LISP Packet Walk
How Does LISP Operate?

1. DNS Entry: D.abc.com A 10.2.0.1

2. 10.1.0.1 -> 10.2.0.1

3. Mapping Entry
   - EID-prefix: 10.2.0.0/24
   - Locator-set:
     - 2.1.1.1, priority: 1, weight: 50 (D1)
     - 2.1.2.1, priority: 1, weight: 50 (D2)

4. 1.1.1.1 -> 2.1.1.1
   - 10.1.0.1 -> 10.2.0.1

5. 10.1.0.1 -> 10.2.0.1

This Policy Controlled by Destination Site

IP Network

- EID-to-RLOC mapping
- 5.4.4.4
- 5.3.3.3
- 5.1.1.1
- 5.2.2.2

LISP Site

- ITR
- PITR

ET 1

West-DC

- 2.1.1.1
- 2.1.2.1

10.2.0.0/24

10.3.0.0/24

East-DC

- 3.1.1.1
- 3.1.2.1

10.1.0.0/24

© 2010 Cisco and/or its affiliates. All rights reserved.
A LISP Packet Walk
How About Non-LISP Sites?

1. DNS Entry: D.abc.com A 10.2.0.1

2. 192.3.0.1 -> 10.2.0.1

3. Mapping Entry
   EID-Prefix: 10.2.0.0/24
   Locator-Set:
   - 2.1.1.1, priority: 1, weight: 50 (D1)
   - 2.1.2.1, priority: 1, weight: 50 (D2)

4. 4.4.4.4 -> 2.1.2.1
   192.3.0.1 -> 10.2.0.1

5. 192.3.0.1 -> 10.2.0.1

IP Network

West-DC

East-DC

ETR

PITR

Non-LISP Site

Non-LISP Site

S

© 2010 Cisco and/or its affiliates. All rights reserved.
LISP Roles and Address Spaces
What Are the Different Components Involved?

LISP Roles

- **Tunnel Routers - xTRs**
  - Edge devices in charge of encap/decap
  - Ingress/Egress Tunnel Routers (ITR/ETR)

- **EID to RLOC Mapping DB**
  - Contains RLOC to EID mappings
  - Distributed across multiple Map Servers (MS)
  - MS may connect over an ALT network

- **Proxy Tunnel Routers - PxTR**
  - Coexistence between LISP and non-LISP sites
  - Ingress/Egress: PITR, PETR

Address Spaces

- **EID = End-point Identifier**
  - Host IP or prefix
- **RLOC = Routing Locator**
  - IP address of routers in the backbone
LISP Mapping Database
The Basics – Registration and Resolution

Database Mapping Entry (on ETR):
10.2.0.0/16 -> (2.1.1.1, 2.1.2.1)

Database Mapping Entry (on ETR):
10.3.0.0/16 -> (3.1.1.1, 3.1.2.1)

Mapping Cache Entry (on ITR):
10.2.0.0/16 -> (2.1.1.1, 2.1.2.1)

Map Server / Resolver: 5.1.1.1

Map-Request
10.2.0.1

Map-Reply
10.2.0.0/16 -> (2.1.1.1, 2.1.2.1)

Map-Register
10.2.0.1

2.1.1.1 2.1.2.1 3.1.1.1 3.1.2.1

ETR ETR ETR ETR

West-DC 10.2.0.0 /16

X Y

10.2.0.2

East-DC 10.3.0.0/16

Y Z
Basic LISP Configuration

Border Routers Between Backbones
- ip lisp proxy-itr
- ip lisp ITR map-resolver 5.3.3.3

Branch Routers
- ip lisp itr-etr
- ip lisp ITR map-resolver 5.3.3.3

DC Aggregation Routers
- ip lisp itr-etr
- ip lisp database-mapping 10.2.0.0/24 2.1.1.1 p1 w50
- ip lisp database-mapping 10.2.0.0/24 2.1.2.1 p1 w50
- ip lisp ETR map-server 5.1.1.1 key s3cr3t
- ip lisp ETR map-server 5.2.2.2 key s3cr3t

Servers
- ip lisp map-resolver
- ip lisp map-server
- lisp site west-DC
- authentication-key 0 s3cr3t
- eid-prefix 10.2.0.0/24

Non-LISP Sites
- IP Network
- EID
- LISP Encap/Decap

USUALLY DEVICES WILL BE CONFIGURED AS ITRs AND ETRs TO HANDLE TRAFFIC IN BOTH DIRECTIONS; WE ILLUSTRATE ONLY ONE DIRECTION FOR SIMPLICITY.
Evolution of the Data Center Fabric

Agenda

- The Evolving Data Centre Fabric
- FabricPath
- VXLAN
- LISP
- LISP Host Mobility
- OTV LAN Extension
- Mobility with Extended Subnets
- Nexus Fabric
LISP Host-Mobility

**Needs:**
- Global IP-Mobility across subnets
- Optimized routing across extended subnet sites

**LISP Solution:**
- Automated *move detection* on xTRs
- Dynamically update EID-to-RLOC mappings
- *Traffic Redirection* on ITRs or PITRs

**Benefits:**
- Direct Path (no triangulation)
- Connections maintained across move
- No routing re-convergence
- No DNS updates required
- Transparent to the hosts
- Global Scalability (cloud bursting)
- IPv4/IPv6 Support
Host-Mobility Scenarios

Moves Without LAN Extension

- Internet or Shared WAN
- LISP Site
- xTR
- Mapping DB
- LISP-VM (xTR)
- West-DC
- East-DC
- DR Location or Cloud Provider DC

Moves With LAN Extension

- Non-LISP Site
- Internet or Shared WAN
- LISP Site
- xTR
- Mapping DB
- LISP-VM (xTR)
- West-DC
- East-DC
- IP Network
- LAN Extension

IP Mobility Across Subnets
- Disaster Recovery
- Cloud Bursting

Routing for Extended Subnets
- Active-Active Data Centers
- Distributed Clusters

Application Members Distributed (Broadcasts across sites)

Application Members in One Location
LISP Host-Mobility - Move Detection
Monitor the Source of Received Traffic

- The new xTR checks the source of received traffic
- Configured dynamic-EIDs define which prefixes may roam

```
lisp dynamic-eid roamer
database-mapping 10.2.0.0/24 <RLOC-C> p1 w50
database-mapping 10.2.0.0/24 <RLOC-D> p1 w50
map-server 5.1.1.1 key abcd
interface vlan 100
lisp mobility roamer
```

Received a Packet …
… It’s from a “New” Host
… It’s in the Dynamic-EID Allowed Range
…It’s a Move!
Register the /32 with LISP

```
Received packet 10.2.0.2
```

```
West-DC
10.2.0.0/16

X

Y

A

B

C

LISP-VM (xTR)

10.2.0.2

10.3.0.0/16

Y

Z

East-DC

```
LISP Host-Mobility - Traffic Redirection
Update Location Mappings for the Host System Wide

• When a host move is detected, updates are triggered:
  The host-to-location mapping in the Database is updated to reflect the new location
  The old ETR is notified of the move
  ITRs are notified to update their Map-caches

• Ingress routers (ITRs or PITRs) now send traffic to the new location
LISP Host-Mobility - First Hop Routing Across Different Subnets

- SVI (Interface VLAN x) and HSRP configured as usual (Consistent GWY-MAC configured across all dynamic subnets)

- The lisp mobility <dyn-eid-map> command enables proxy-arp functionality on the SVI
  
The LISP-VM router services first hop routing requests for both local and roaming subnets

- Hosts can move anywhere and always talk to a local gateway with the same MAC
Host-Mobility and Multi-homing
ETR updates – across LISP sites

Null0 host routes indicate the host is “away”
Refreshing the map caches

1. ITRs and PITRs with cached mappings continue to send traffic to the old locators
   The old xTR knows the host has moved (Null0 route).

2. Old xTR sends Solicit Map Request (SMR) messages to any encapsulators sending traffic to the moved host

3. The ITR then initiates a new map request process

4. An updated map-reply is issued from the new location

5. The ITR Map Cache is updated
   - Traffic is now re-directed
   - SMRs are an important integrity measure to avoid unsolicited map responses and spoofing
LISP Host-Mobility Configuration Across Subnets (No LAN Extensions)

```plaintext
ip lisp ITR-ETR
ip lisp database-mapping 10.2.0.0/16 <RLOC-A>
ip lisp database-mapping 10.2.0.0/16 <RLOC-B>

lisp dynamic-eid roamer
database-mapping 10.2.0.0/24 <RLOC-A>
database-mapping 10.2.0.0/24 <RLOC-B>
map-server 1.1.1.1 key abcd
map-notify-group 239.1.1.1
interface vlan 100
ip address 10.2.0.10 /16
lisp mobility roamer
ip proxy-arp
hsrp 101
  mac-address 0000.01e1.010c
  ip 10.2.0.1

Mapping DB
```
Evolution of the Data Center Fabric

Agenda

- The Evolving Data Centre Fabric
- FabricPath
- VXLAN
- LISP
- LISP Host Mobility
- OTV LAN Extension
- Mobility with Extended Subnets
- Nexus Fabric
Overlay Transport Virtualization (OTV)
Simplifying LAN Extensions

- Ethernet LAN Extension over any Network
  Works over dark fiber, MPLS, or IP
  Multi-data center scalability

- Simplified Configuration & Operation
  Seamless overlay - No network re-design
  Single touch site configuration

- High Resiliency
  Failure domain isolation
  Seamless Multi-homing

- Maximizes available bandwidth
  Automated multi-pathing
  Optimal multicast replication

Many Physical Sites – One Logical Data Center

Any Workload, Anytime, Anywhere
Unleashing the Full Potential of Compute Virtualization
OTV Data Plane
Inter-Site Packet Flow

1. Layer 2 lookup on the destination MAC. MAC 3 is reachable through IP B
2. The Edge Device encapsulates the frame
3. The transport delivers the packet to the Edge Device on site East
4. The Edge Device on site East receives and decapsulates the packet
5. Layer 2 lookup on the original frame. MAC 3 is a local MAC
6. The frame is delivered to the destination
Building the MAC Tables
The OTV Control Plane

- OTV **proactively advertises** MAC reachability (control-plane learning)
- MAC addresses advertised in the background once OTV has been configured
- IS-IS is the OTV Control Protocol running between the Edge Devices
- **No specific configuration is required**
Evolution of the Data Center Fabric

Agenda

- The Evolving Data Centre Fabric
- FabricPath
- VXLAN
- LISP
- LISP Host Mobility
- OTV LAN Extension
- Mobility with Extended Subnets
- Nexus Fabric
Ingress Routing Challenge in DCI
Extending Subnets Creates a Routing Challenge

- A subnet usually implies location
- Yet we use LAN extensions to stretch subnets across locations
  Location semantics of subnets are lost
- Traditional routing relies on the location semantics of the subnet
  Can’t tell if a server is at the East or West location of the subnet
- More granular (host level) information is required
  LISP provides host level location semantics
Host-Mobility and Multi-homing
ETR updates – Extended Subnets

Null0 host routes indicate the host is “away”
Refreshing the map caches

1. ITRs and PITRs with cached mappings continue to send traffic to the old locators
   1. The old xTR knows the host has moved (Null0 route).

2. Old xTR sends Solicit Map Request (SMR) messages to any encapsulators sending traffic to the moved host

3. The ITR then initiates a new map request process

4. An updated map-reply is issued from the new location

5. The ITR Map Cache is updated
   - Traffic is now re-directed
   - SMRs are an important integrity measure to avoid unsolicited map responses and spoofing
LISP Host-Mobility - First Hop Routing With Extended Subnets

- Consistent GWY-IP and GWY-MAC configured across all sites
  - Consistent HSRP group number across sites ➔ consistent GWY-MAC
- Servers can move anywhere and always talk to a local gateway with the same IP/MAC

```plaintext
interface vlan 100
  ip address 10.2.0.5/24
  lisp mobility roamer
  lisp extended-subnet-mode
  hsrp 101
  ip 10.2.0.1

interface vlan 100
  ip address 10.2.0.7/24
  lisp mobility roamer
  lisp extended-subnet-mode
  hsrp 101
  ip 10.2.0.1
```

```plaintext
interface vlan 200
  ip address 10.2.0.8/24
```
LISP VM-Mobility Configuration
With Extended Subnets → Use “Extended-Subnet-Mode”

```
ip lisp ITR-ETR
ip lisp database-mapping 10.2.0.0/16 <RLOC-A>
```

```
lisp dynamic-eid roamer
database-mapping 10.2.0.0/24 <RLOC-A> ...
database-mapping 10.2.0.0/24 <RLOC-B>
map-server 1.1.1.1 key abcd
map-notify-group 239.10.10.10
```

```
interface vlan 100
ip address 10.2.0.10 /16
lisp mobility roamer
lisp extended-subnet-mode
hsrp 101
   ip 10.2.0.1
```

```
map-notify-group 239.10.10.10
```

```
interface vlan 100
ip address 10.2.0.11 /16
lisp mobility roamer
lisp extended-subnet-mode
hsrp 101
   ip 10.2.0.1
```

```
lisp dynamic-eid roamer
database-mapping 10.2.0.0/24 <RLOC-C>
database-mapping 10.2.0.0/24 <RLOC-D>
map-server 1.1.1.1 key abcd
```

© 2010 Cisco and/or its affiliates. All rights reserved.
Off-Subnet Client-Server Traffic
All Off-Subnet/Off-Site Traffic Is LISP Encapsulated

- Clients (192.168.0.1 & 192.168.2.1) communicate with Server 10.2.0.2

- Client-server traffic is LISP encapsulated at the ITRs or PITRs
  - Client-to-server:
    - to ETRs C or D
  - Server-to-client:
    - to ETR (F) for LISP sites
    - to PETR (G) for non-LISP sites

- Server-Server off-subnet traffic across sites is also LISP encapsulated
On-Subnet Server-Server Traffic

On Subnet Traffic Across L3 boundaries

With LAN Extension
- Live moves and cluster member dispersion
- Traffic between X & Y uses the LAN Extension
- Link-local-multicast handled by the LAN Extension

Without LAN Extensions
- Cold moves, no application dispersion
- X- Y traffic is sent to the LISP-VM router & LISP encapsulated
- Need LAN extensions for link-local multicast traffic
Evolution of the Data Center Fabric

Agenda

- The Evolving Data Centre Fabric
- FabricPath
- VXLAN
- LISP
- LISP Host Mobility
- OTV LAN Extension
- Mobility with Extended Subnets
- Nexus Fabric
**Fabric Use Case – Distributed Clusters**

- Enhance application availability by distributing Cluster members across PODs and across locations.
- Distance limited by application latency budget and storage replication.
- Two types of traffic specific to the cluster:
  - Non-routable heartbeats: FabricPath (Intra-DC) & OTV (Inter-DC) provide LAN connectivity.
  - Front-end IP connectivity: LISP provides mobility for cluster virtual-IP failover.

![Diagram showing Fabirc Path (Intra-DC) and OTV (Inter-DC) connecting DC-west and DC-east with LISP IP mobility.](image)
Fabric Use Case – Elastic Capacity

- VXLAN & FP provide elasticity within the DC within a L2 POD and across PODs
- OTV extends the LAN across DC sites without compromising network stability
- LISP integrates with SLBs and balances traffic across the SLBs (Future)

<table>
<thead>
<tr>
<th>Intra-DC</th>
<th>Inter-DC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Virtual Machines</td>
<td>VXLAN (x-L3), FabricPath (L2)</td>
</tr>
<tr>
<td>Physical Machines</td>
<td>FabricPath (L2), VXLAN GWY (future)</td>
</tr>
</tbody>
</table>
Fabric Use Case – Disaster Recovery

- Reduce Disaster Recovery Bring-up times - Less Network Changes/Operations = Faster recovery times
- Preserve IP addressing with LISP host mobility
  - No reconfiguration of applications or network service policies
  - No routing re-convergence
  - Automatic routing re-localization (upon application bring-up at DR)
- VXLAN segments move along with the applications (vApps)
Fabric Use Case – Private Cloud

- Move virtual Applications (vApps) to private cloud PODs
  - Move VMs and virtual Segments (VXLANs)
- LISP host mobility allows the vApp GWY to roam
  - Maintain GWY IP address and optimal reachability
- VXLAN segments move along with the applications (vApps)
  - Very large scale of virtual segments can move and extend across L3 boundaries

vApp = Collection of VMs and segments with a GWY
Complimentary Capabilities
FabricPath, VXLAN, LISP

<table>
<thead>
<tr>
<th>Requirement</th>
<th>Intra-DC</th>
<th>Inter-DC</th>
</tr>
</thead>
<tbody>
<tr>
<td>Layer 2 connectivity</td>
<td>FabricPath/TRILL/VXLAN</td>
<td>OTV/VPLS</td>
</tr>
<tr>
<td>IP Mobility</td>
<td>LISP</td>
<td>LISP</td>
</tr>
<tr>
<td>Secure Segmentation</td>
<td>VXLAN / Segment-ID</td>
<td>VPNs (LISP/MPLS)</td>
</tr>
</tbody>
</table>

- **Fabric Path (Intra-DC L2)**
- **VXLAN/OTV (Intra-DC x-L3)**
- **OTV/VPLS (Inter-DC x-L3)**
- **Fabric Path (Intra-DC L2)**
- **VXLAN/OTV (Intra-DC x-L3)**

**Scale**

**IP Network**

**LISP IP mobility**

**App OS**
We value your feedback.
Please be sure to complete the Evaluation Form for this session.

Access today’s presentations at cisco.com/ca/plus

Follow @CiscoCanada and join the #CiscoPlusCA conversation