Cisco live!

What You Make Possible

TOMORROW starts here.

CISCO

Cisco FabricPath Technology and Design

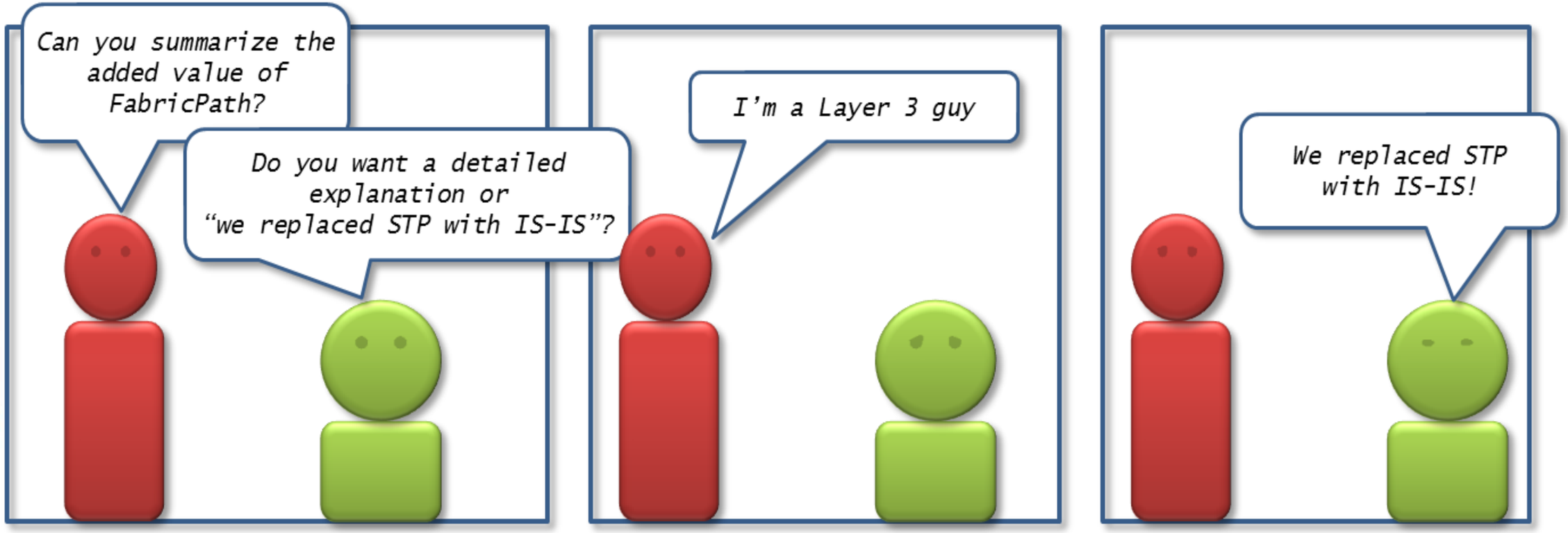
BRKDCT-2081

Agenda

- Introduction to FabricPath
- FabricPath Concepts
- FabricPath Technology
- FabricPath vs Trill
- FabricPath Designs
- Conclusion

© 2013 Cisco and/or its affiliates. All rights reserved.

Introduction to FabricPath



By Francois Tallet

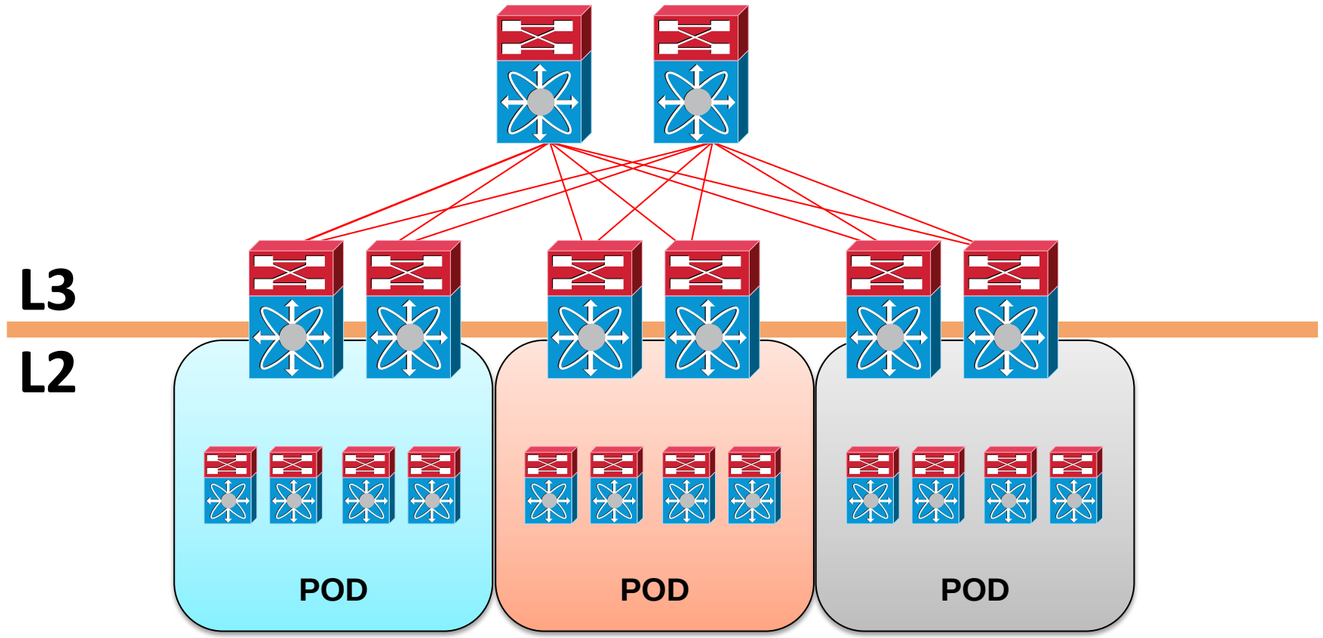
© 2013 Cisco and/or its affiliates. All rights reserved.

Why Layer 2 in the Data Centre?

- Some Applications / Protocols rely on the functionality
- Simple, plug and play
- Addressing constraints
- Allows easy server provisioning
- Allows virtual machine mobility

© 2013 Cisco and/or its affiliates. All rights reserved.

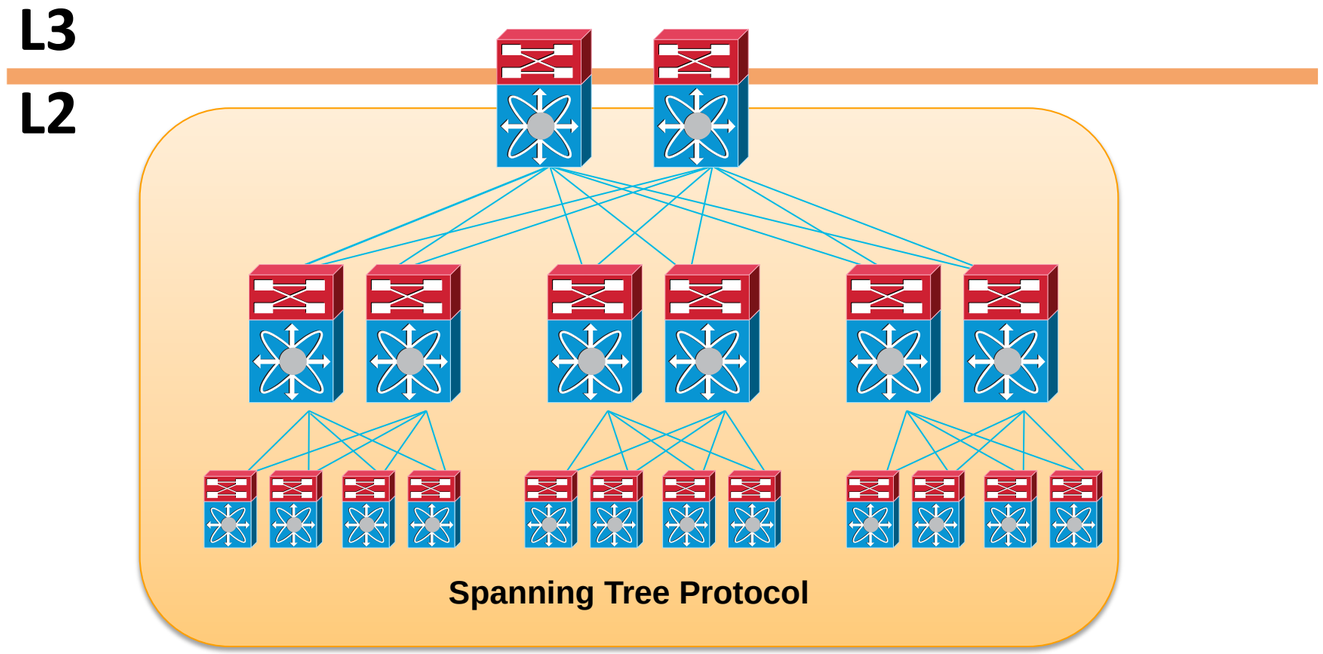
Current Data Centre Design



- L2 benefits are limited to a single POD

© 2013 Cisco and/or its affiliates. All rights reserved.

Current Data Centre Design



- We could extend STP to the whole network

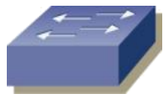
© 2013 Cisco and/or its affiliates. All rights reserved.

Typical Limitations of L2

- Local STP problems have network-wide impact, troubleshooting is difficult
- STP convergence is disruptive
- Flooding impacts the whole network
- STP provides limited bandwidth (no load balancing)
- Tree topologies introduce sub-optimal paths
- MAC address tables don’t scale

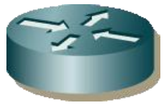
© 2013 Cisco and/or its affiliates. All rights reserved.

Cisco FabricPath Goal



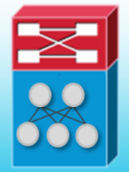
Switching

- Easy Configuration
- Plug & Play
- Provisioning Flexibility



Routing

- Multi-pathing (ECMP)
- Fast Convergence
- Highly Scalable



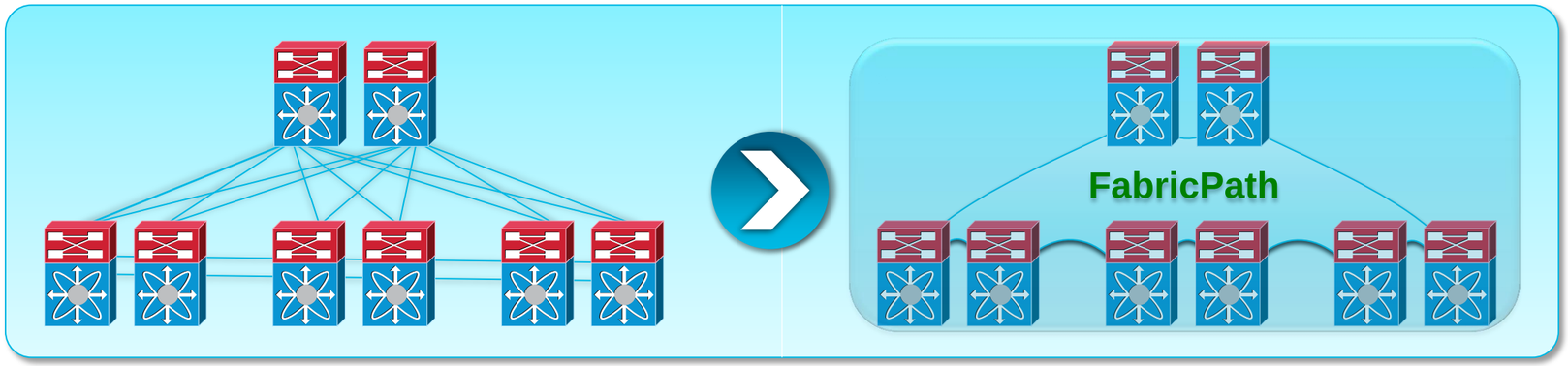
FabricPath

"FabricPath brings Layer 3 routing benefits to flexible Layer 2 bridged Ethernet networks"

© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath: An Ethernet Fabric

Turn the Network into a Fabric



- Connect a group of switches using an **arbitrary** topology
- With a simple CLI, aggregate them into a Fabric:

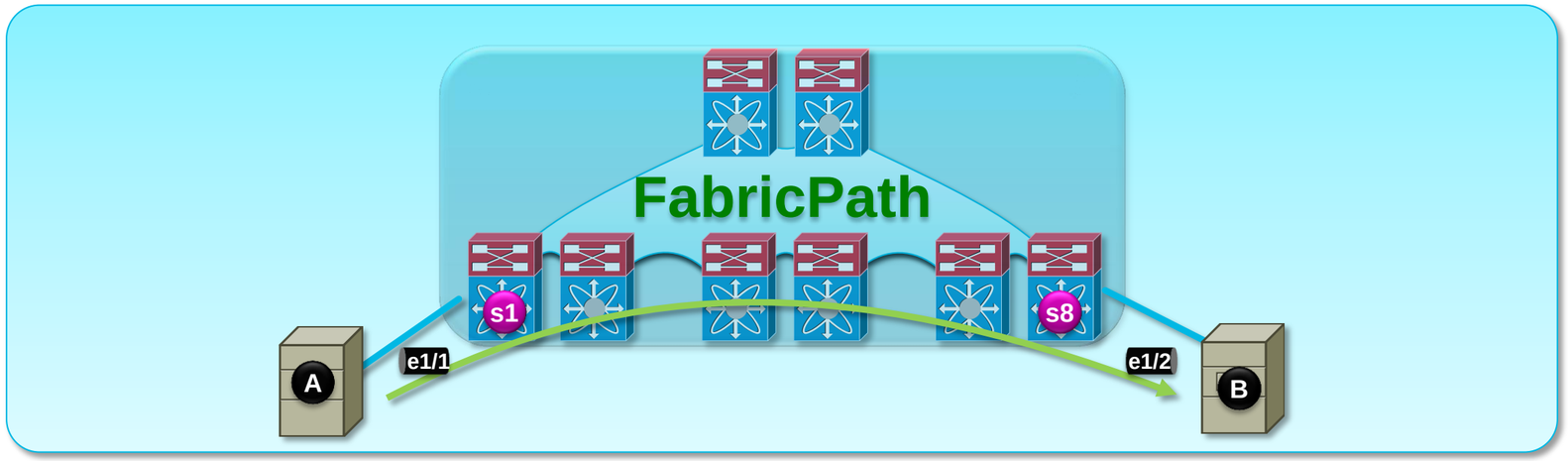
```
N7K(config)# interface ethernet 1/1
N7K(config-if)# switchport mode fabricpath
```

- No STP inside. An open protocol based on L3 technology provides Fabric-wide intelligence and ties the elements together.

© 2013 Cisco and/or its affiliates. All rights reserved.

Optimal, Low Latency Switching

Shortest Path Any-to-Any

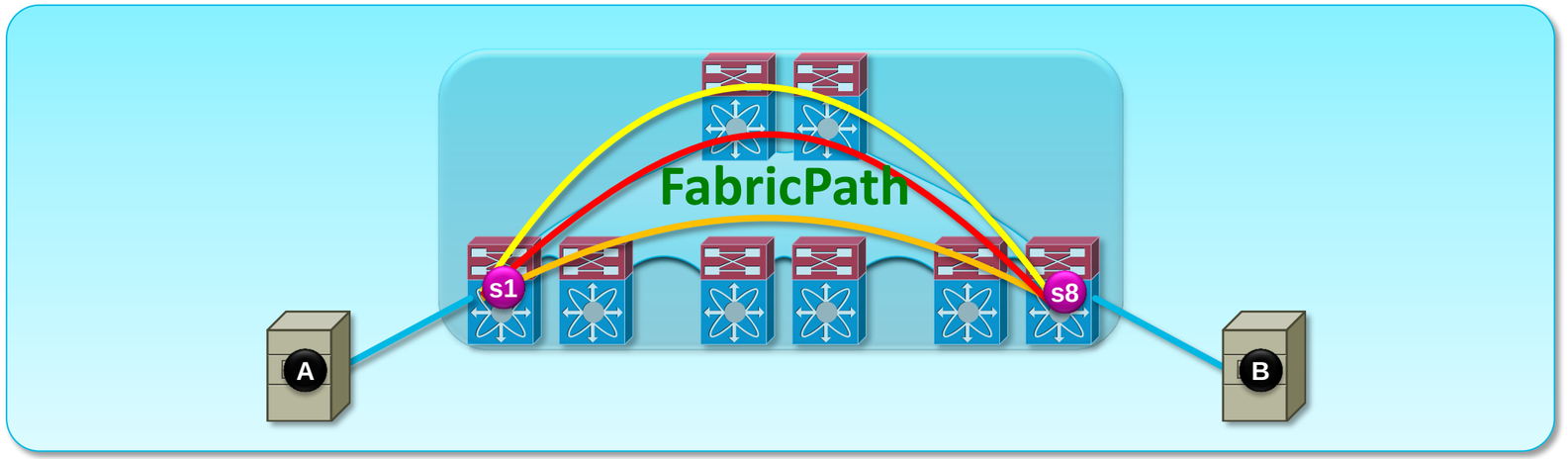


- Single address lookup at the ingress edge identifies the exit port across the fabric
- Traffic is then switched using the shortest path available
- Reliable L2 and L3 connectivity any to any
  (L2 as if it was within the same switch, **no STP inside**)

© 2013 Cisco and/or its affiliates. All rights reserved.

High Bandwidth, High Resiliency

Equal Cost Multi-Pathing

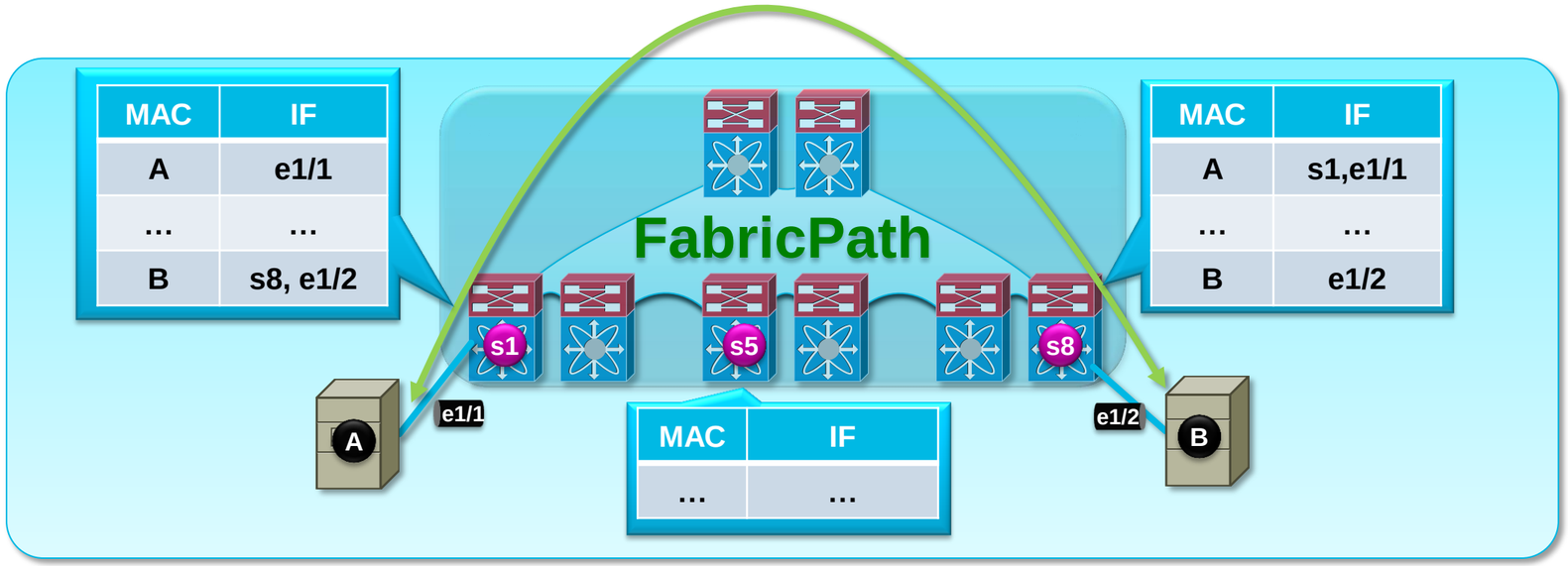


- Multi-pathing (up to 256 links active between any 2 devices)
- Traffic is redistributed across remaining links in case of failure, providing fast convergence

© 2013 Cisco and/or its affiliates. All rights reserved.

Scalable

Conversational MAC Learning

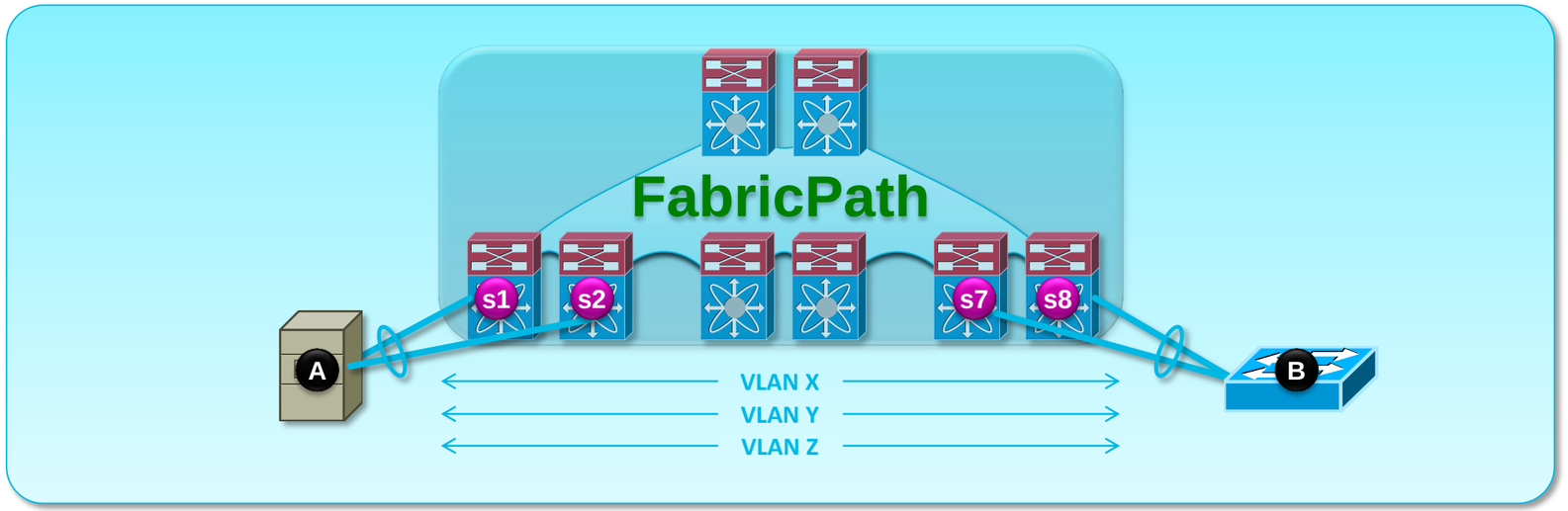


- Per-port MAC address table only needs to learn the peers that are reached across the fabric
- A larger number of hosts can be attached to the fabric

© 2013 Cisco and/or its affiliates. All rights reserved.

Layer 2 Integration

VPC+

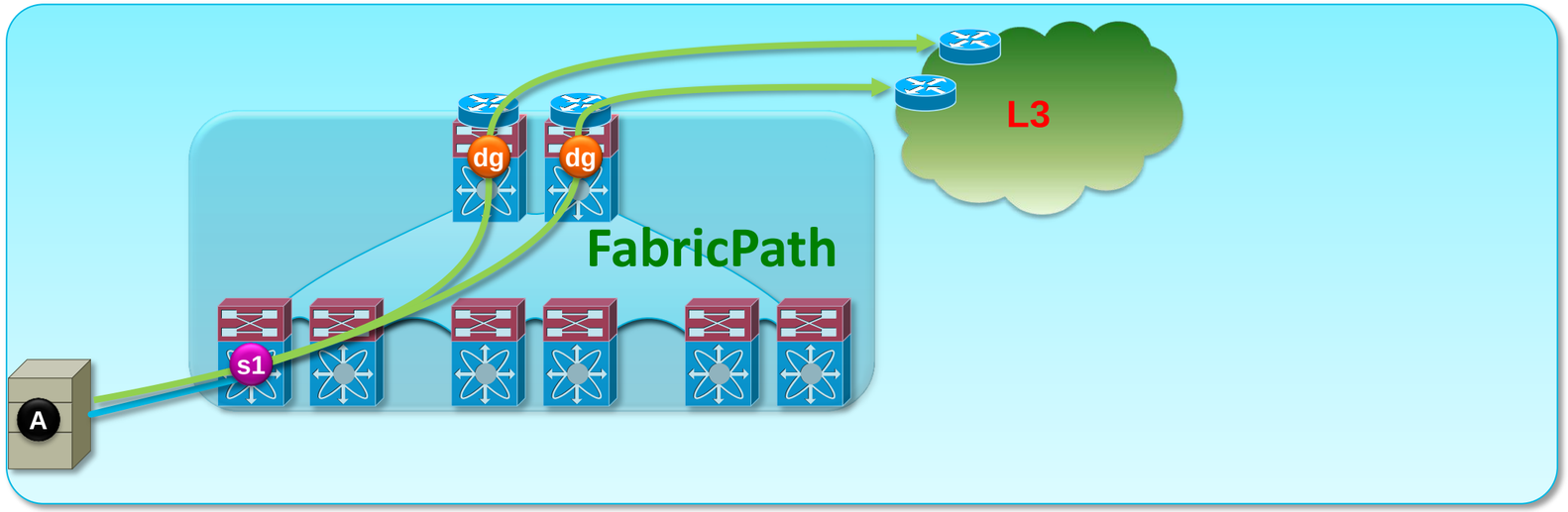


- Allows extending VLANs with no limitation (no risks of loop)
- Devices can be attached active/active to the fabric using IEEE standard port channels and without resorting to STP

© 2013 Cisco and/or its affiliates. All rights reserved.

Edge Device Integration

Hosts Can Leverage Multiple L3 Default Gateways

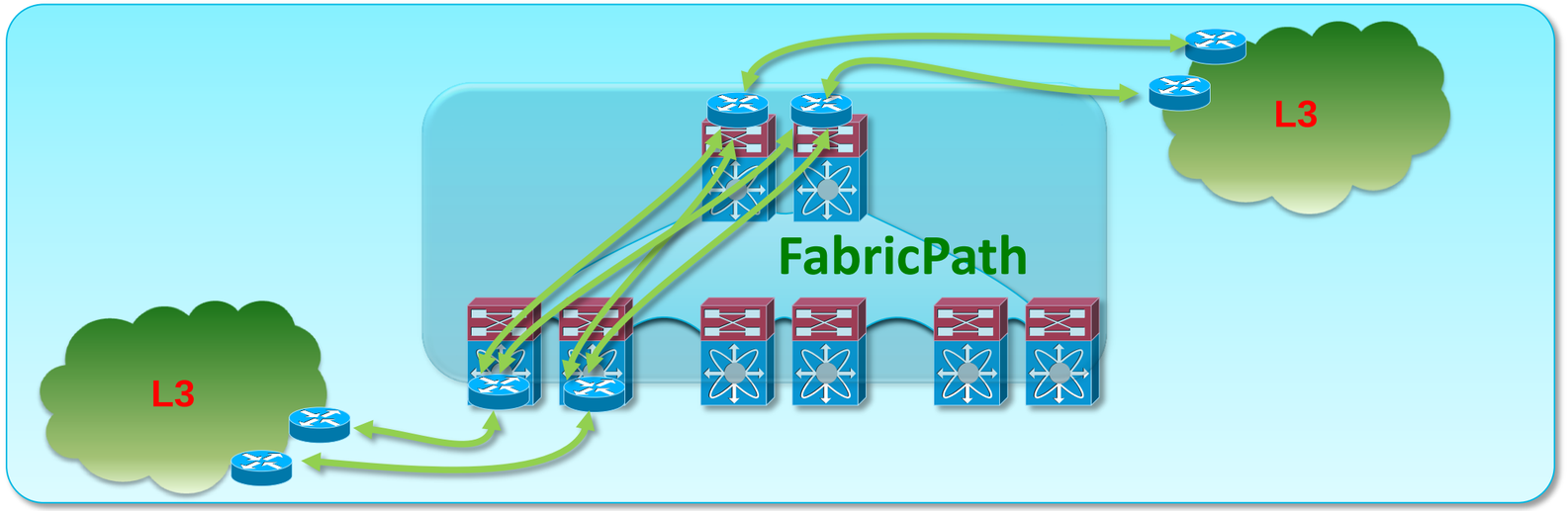


- Hosts see a single default gateway
- The fabric provide them transparently with multiple simultaneously active default gateways dg
- Allows extending the multipathing from the inside of the fabric to the L3 domain outside the fabric

© 2013 Cisco and/or its affiliates. All rights reserved.

Layer 3 Integration

XL Tables, SVIs Anywhere

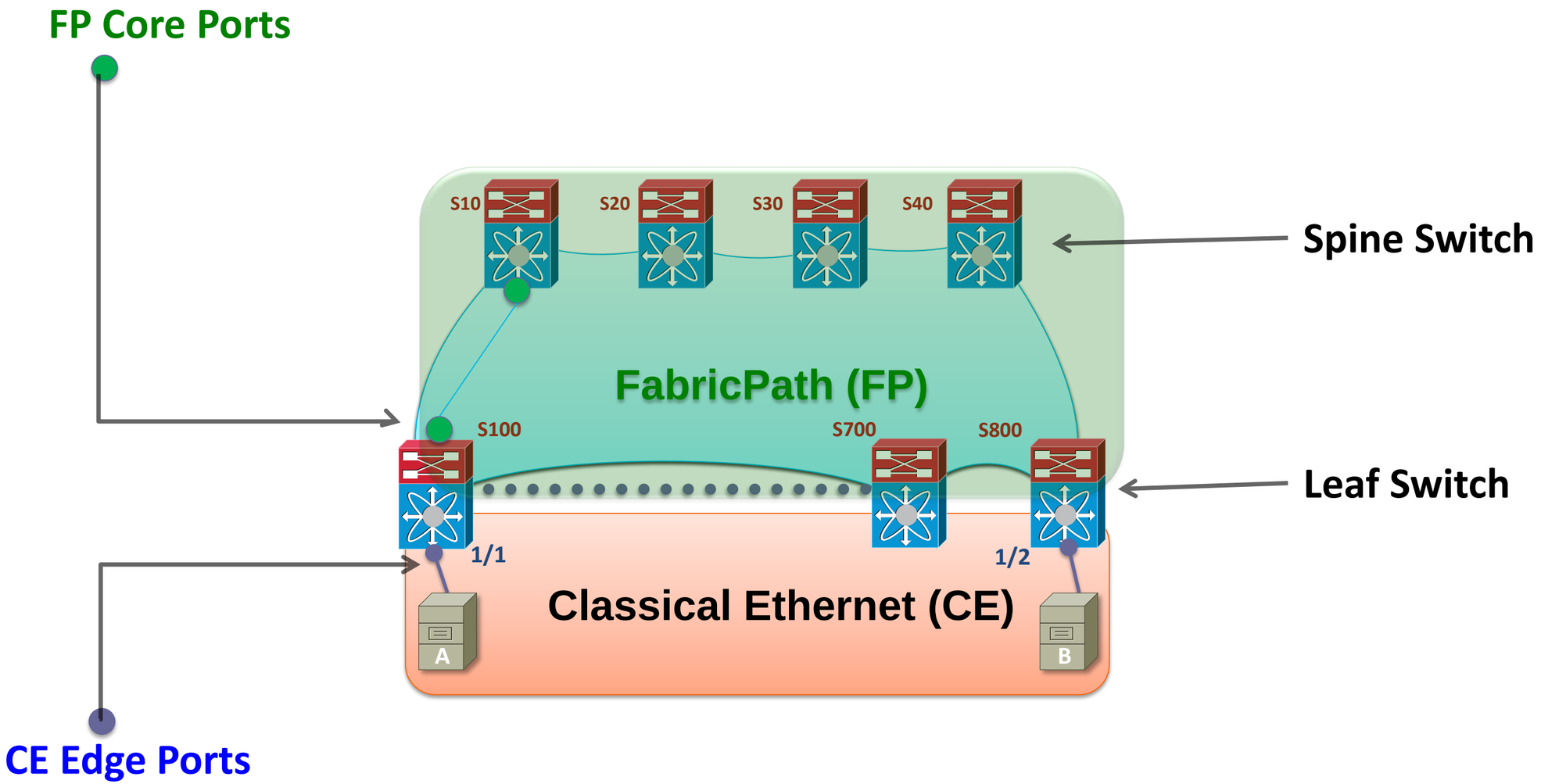


- The fabric provides seamless L3 integration
- An arbitrary number of routed interfaces can be created at the edge or within the fabric
- Attached L3 devices can peer with those interfaces
- The hardware is capable of handling million of routes

© 2013 Cisco and/or its affiliates. All rights reserved.

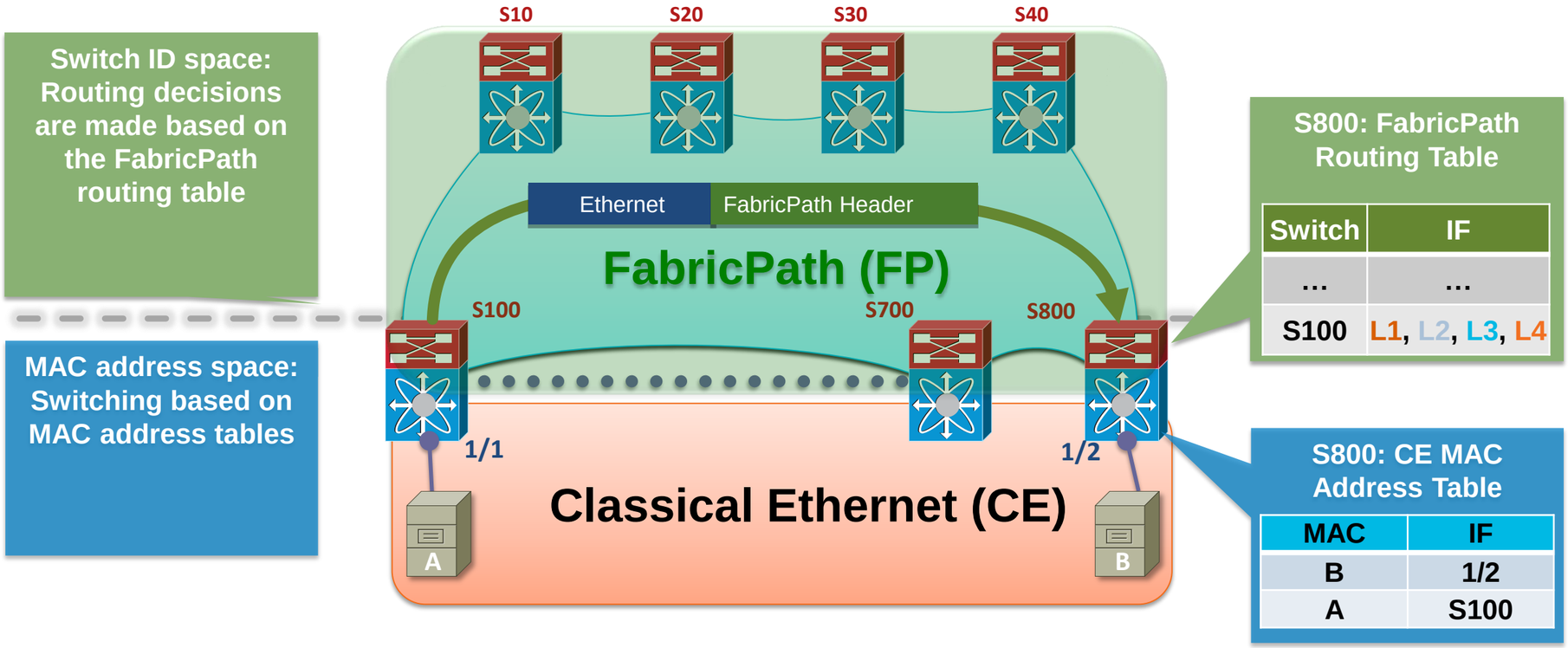
FabricPath Concepts

FabricPath Terminology



© 2013 Cisco and/or its affiliates. All rights reserved.

New Data Plane



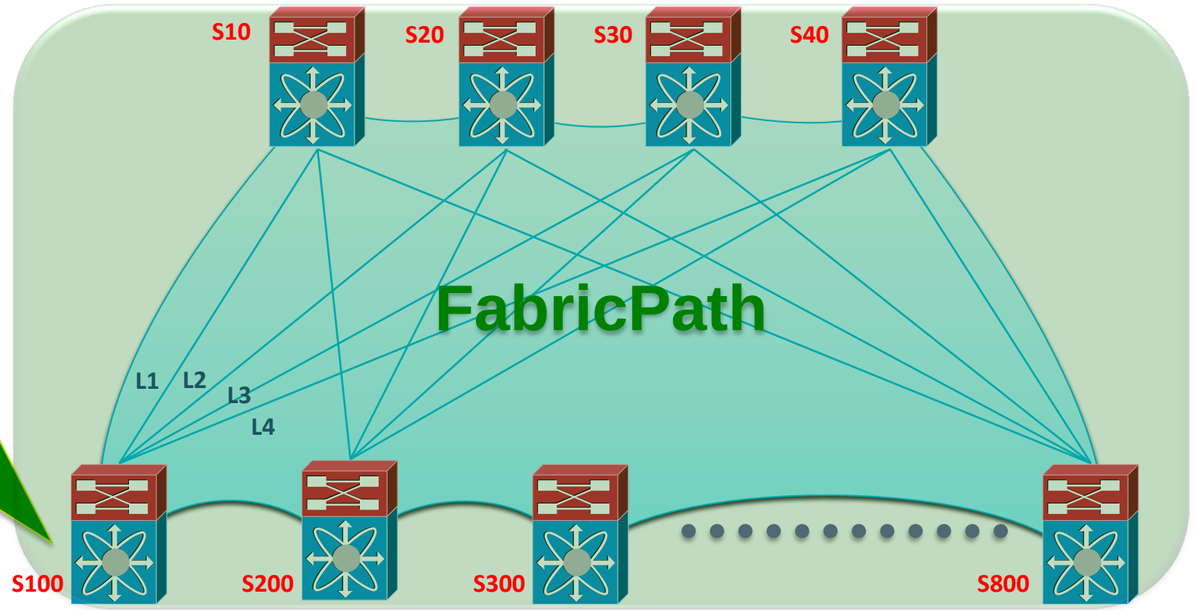
- The association MAC address/Switch ID is maintained at the edge
- Traffic is encapsulated across the Fabric

© 2013 Cisco and/or its affiliates. All rights reserved.

New Control Plane

FabricPath Routing Table

Switch	IF
S10	L1
S20	L2
S30	L3
S40	L4
S200	L1, L2, L3, L4
…	…
S800	L1, L2, L3, L4



- MAC addresses are not carried or redistributed into the Control Plane
- The Control plane determines fabric topology and switch reachability

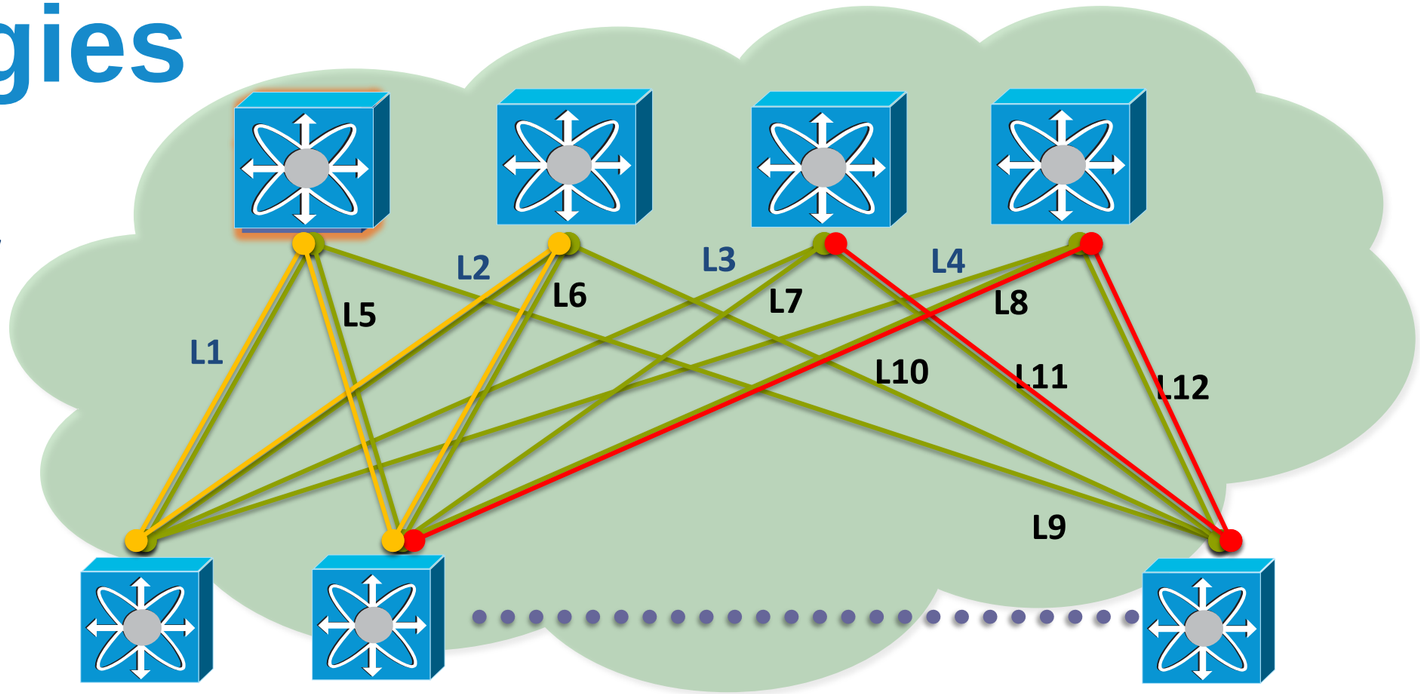
© 2013 Cisco and/or its affiliates. All rights reserved.

Multiple Topologies

Topology 0 – Default Topology

Topology 1

Topology 2



Topology: A group of links in the Fabric.
By default, all the links are part of topology 0.

- A VLAN is mapped to a unique topology
- Other topologies can be created by assigning a subset of the links to them.
- A link can belong to several topologies

Topologies can be used for migration designs (i.e. VLAN localisation, VLAN re-use), traffic engineering, security etc…

© 2013 Cisco and/or its affiliates. All rights reserved.

Equal Cost Multipathing

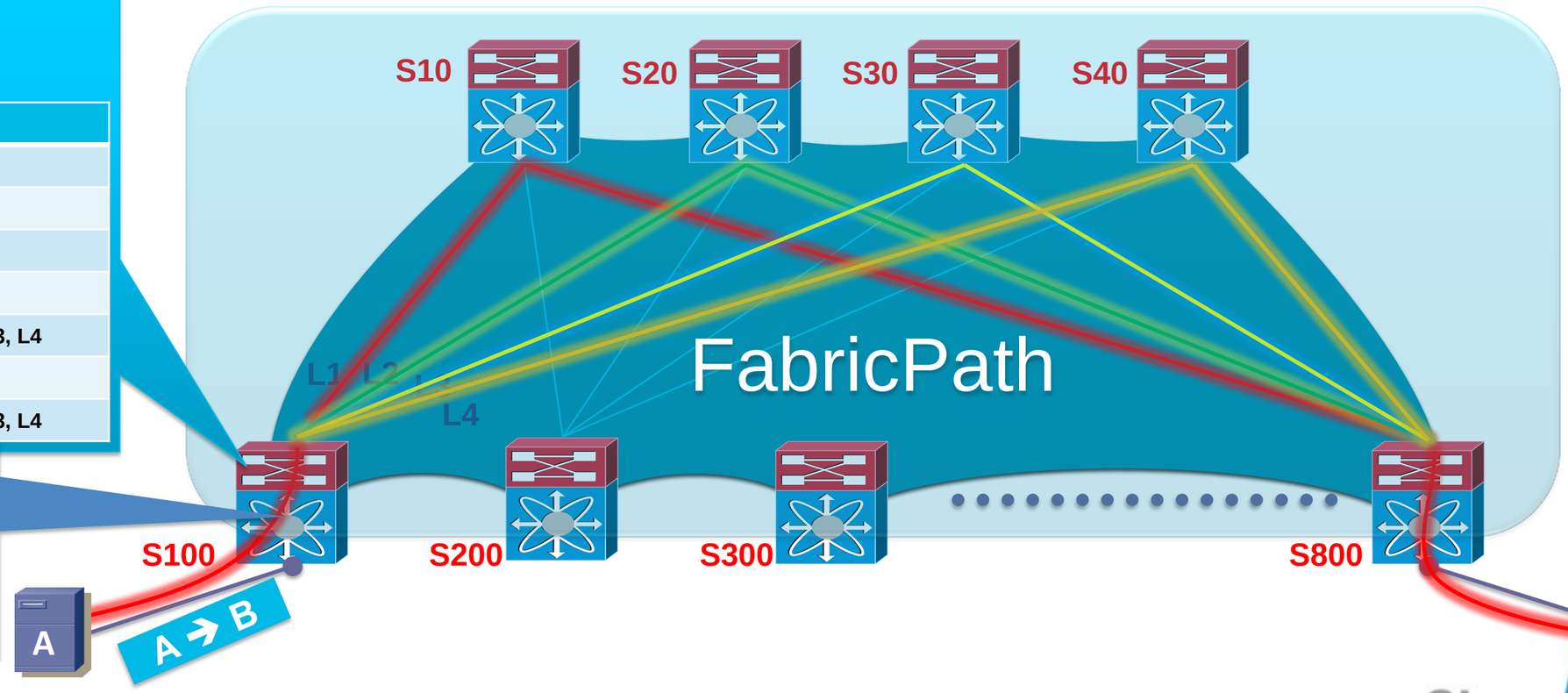
Traffic Forwarded Based on a Routing Table

FabricPath Routing Table

Switch	IF
S10	L1
S20	L2
S30	L3
S40	L4
S200	L1, L2, L3, L4
…	…
S800	L1, L2, L3, L4

S100: CE MAC Address Table

MAC	IF
A	1/1
B	S800

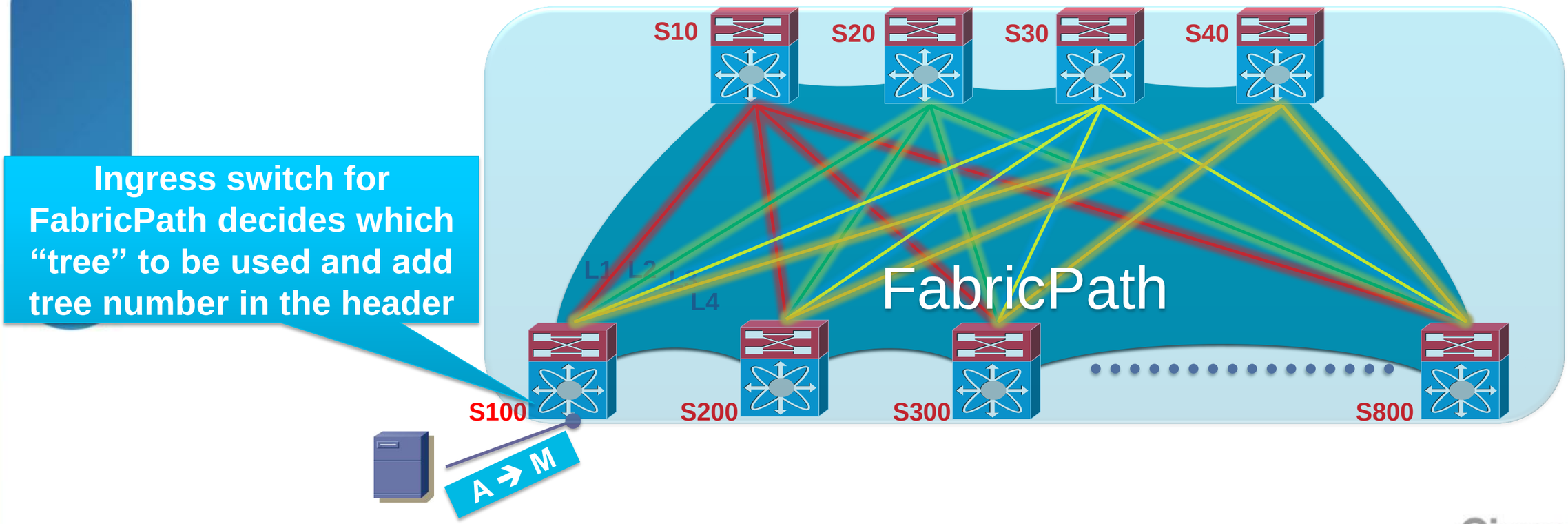


© 2013 Cisco and/or its affiliates. All rights reserved.

Multicast Traffic

Load Balancing on a Per-tree Basis

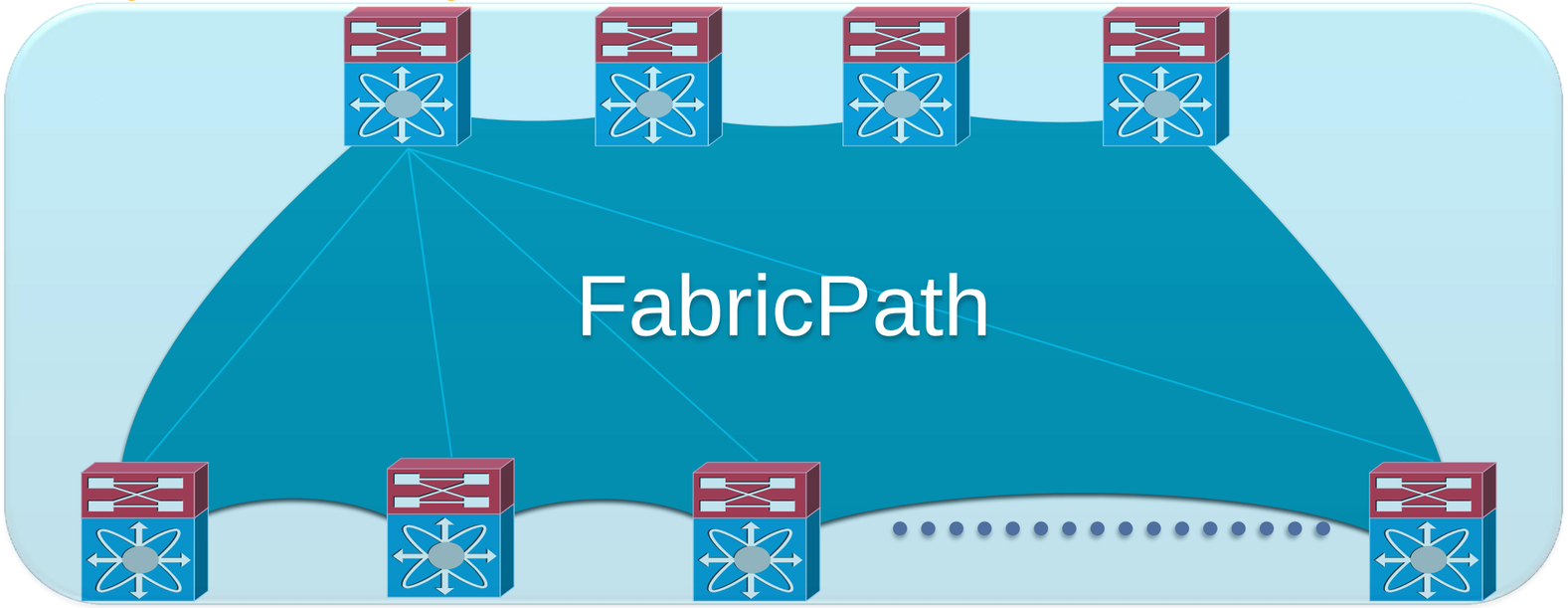
- IS-IS computes several trees automatically
- Location of the root switches can be configured
- Multicast traffic is pinned to a tree at the edge

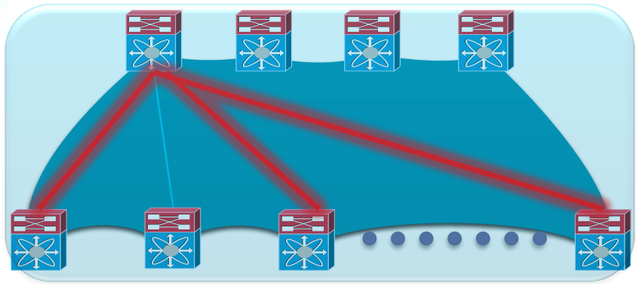


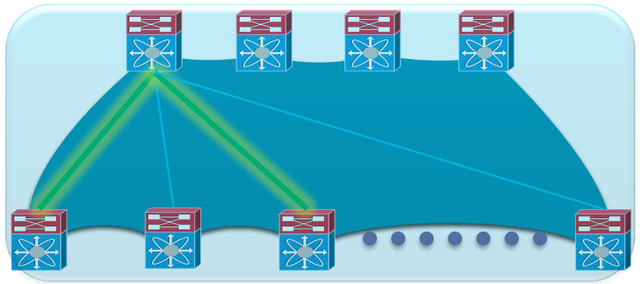
© 2013 Cisco and/or its affiliates. All rights reserved.

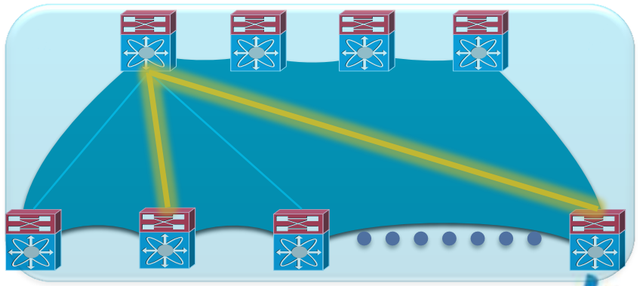
VLAN Pruning By Design

Automatically Handled by IS-IS







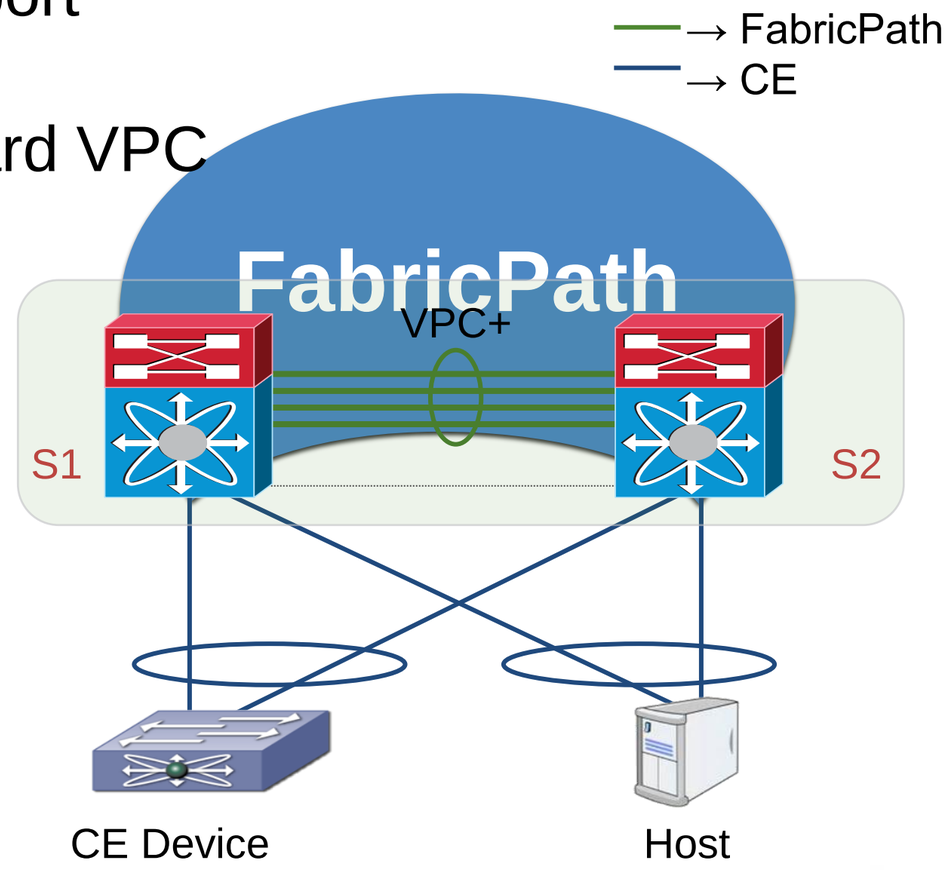


© 2013 Cisco and/or its affiliates. All rights reserved.

VPC+

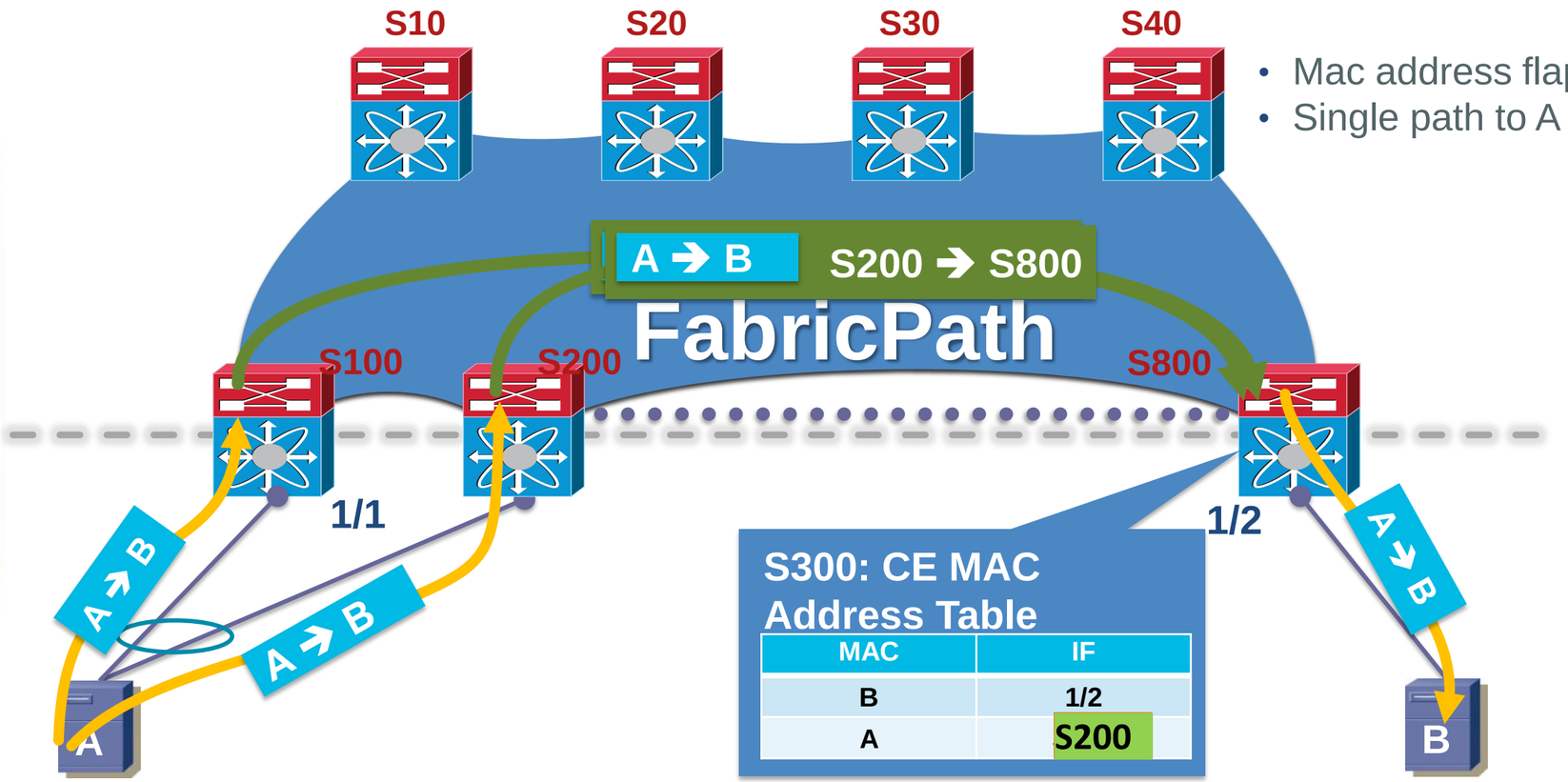
Virtual Port Channel in FabricPath Environment

- Allows non FabricPath capable devices to connect redundantly to the fabric using port channels
- Configuration virtually identical to standard VPC
- Provides active/active HSRP



© 2013 Cisco and/or its affiliates. All rights reserved.

VPC+ Technical Challenges



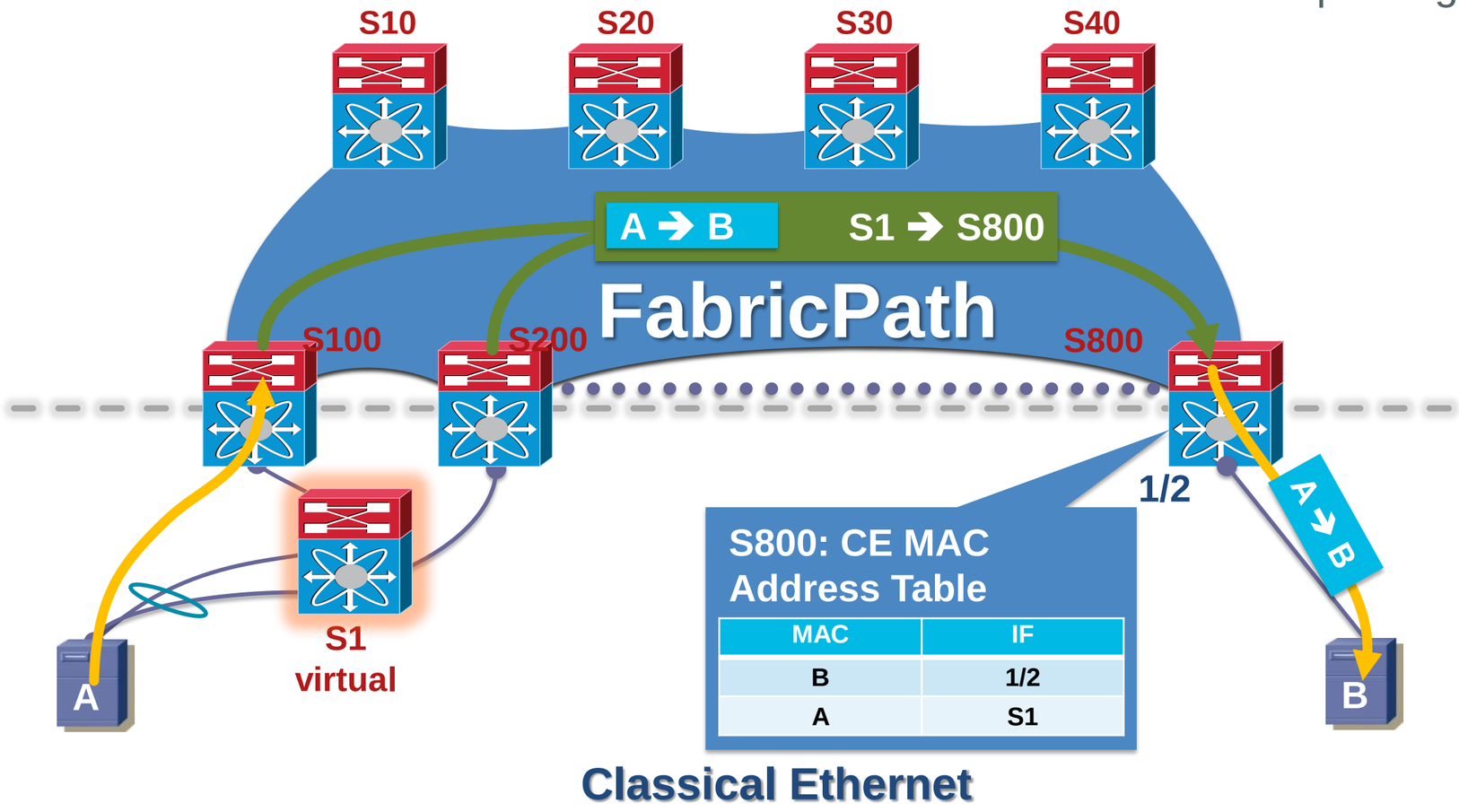
- Mac address flapping on S300
- Single path to A

Classical Ethernet

© 2013 Cisco and/or its affiliates. All rights reserved.

VPC+ Virtual Switch

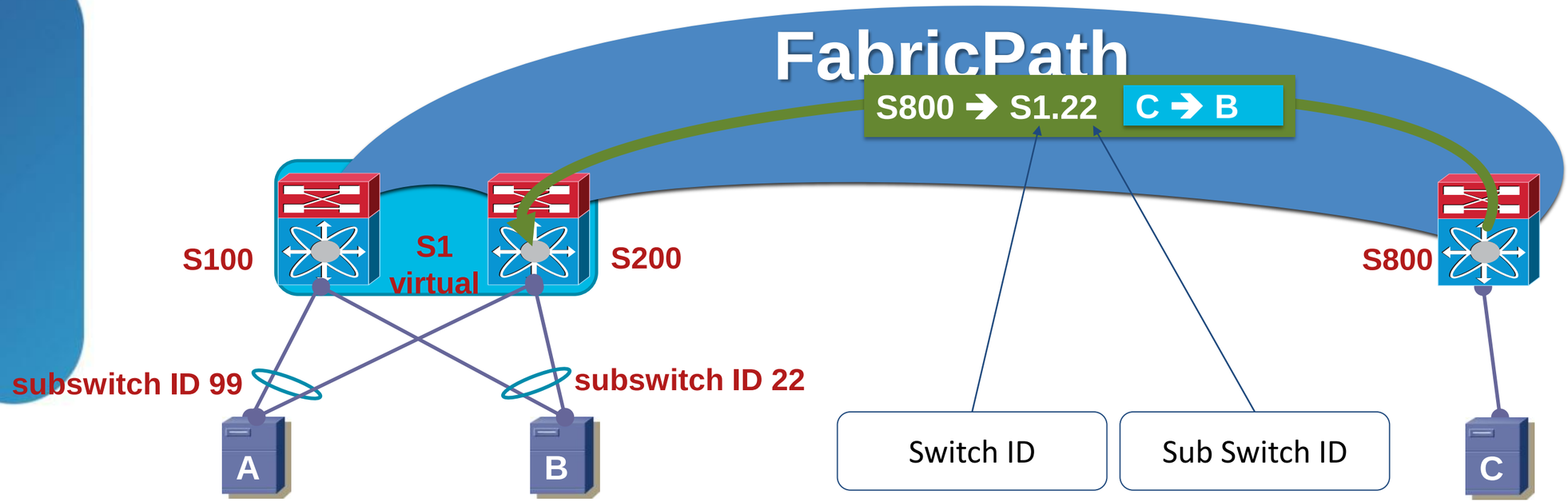
- A consistently associated to S1
- Multipathing to A



© 2013 Cisco and/or its affiliates. All rights reserved.

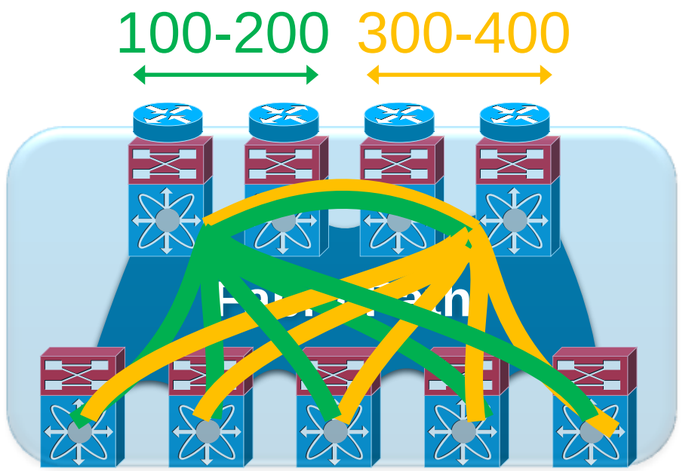
Sub-Switch ID

Identifies a VPC off a Virtual Switch



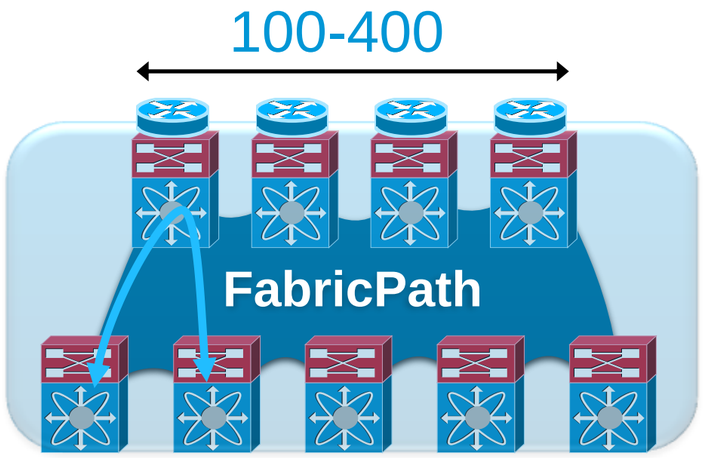
© 2013 Cisco and/or its affiliates. All rights reserved.

Anycast FHRP



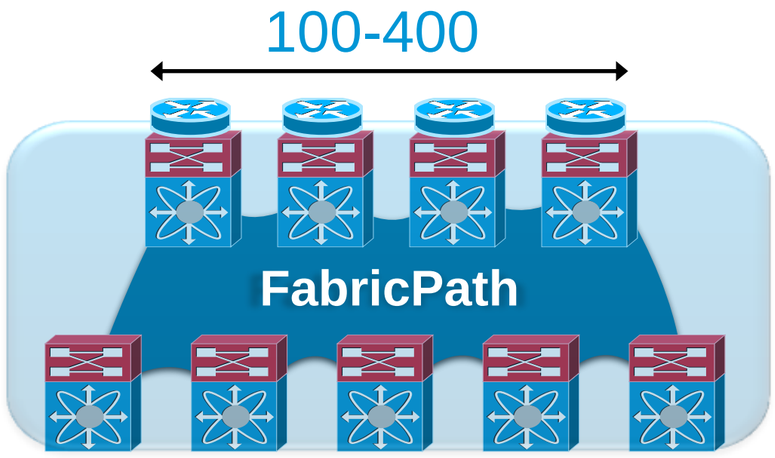
Split VLANs

- Some polarisation
- Inter-VLAN traffic can be suboptimal



GLBP

- Host is pinned to a single gateway
- sLess granular load balancing



Anycast FHRP

- All active paths
- Available in the future for routing

© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath Technology

FabricPath Control Plane

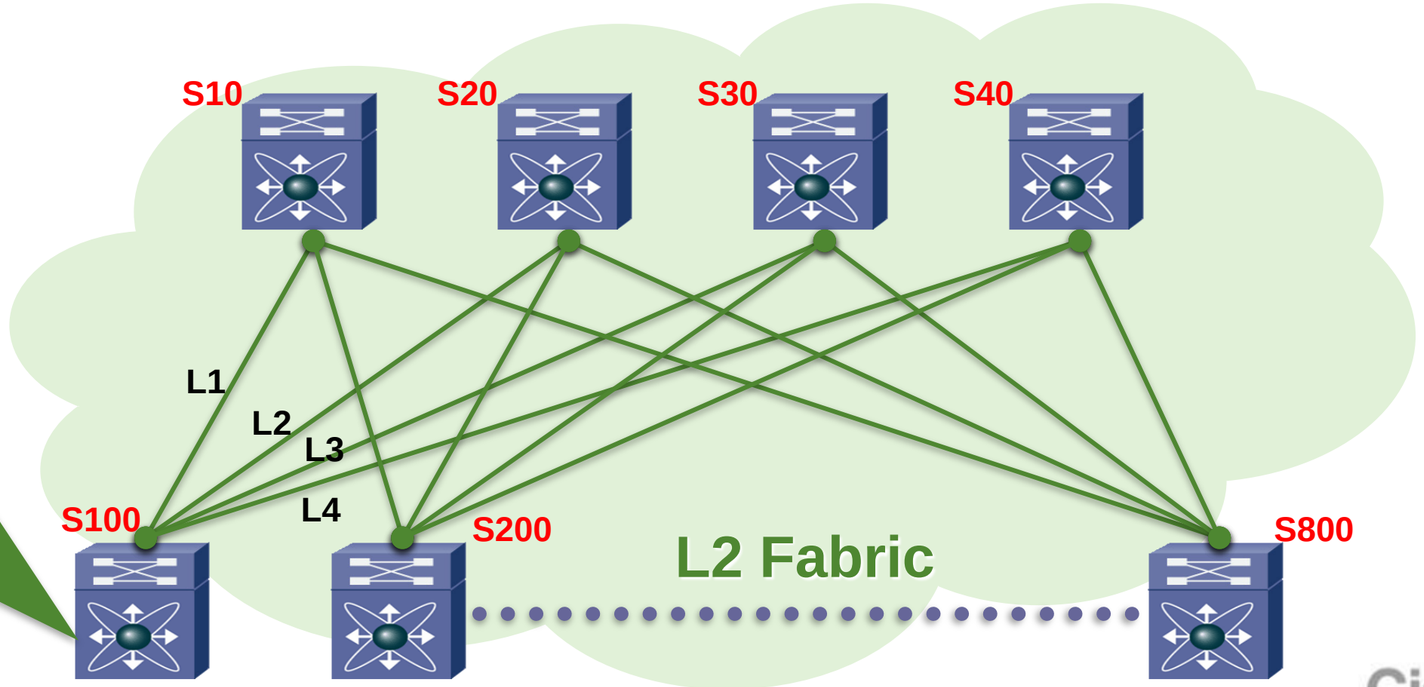
Control Plane Operation

Plug-N-Play L2 IS-IS is Used to Manage Forwarding Topology

- Assigned switch addresses to all FabricPath enabled switches automatically (no user configuration required)
- Compute shortest, pair-wise paths
- Support equal-cost paths between any FabricPath switch pairs

FabricPath Routing Table

Switch	IF
S10	L1
S20	L2
S30	L3
S40	L4
S200	L1, L2, L3, L4
…	…
S800	L1, L2, L3, L4



© 2013 Cisco and/or its affiliates. All rights reserved.

Display IS-IS View of Routing Topology

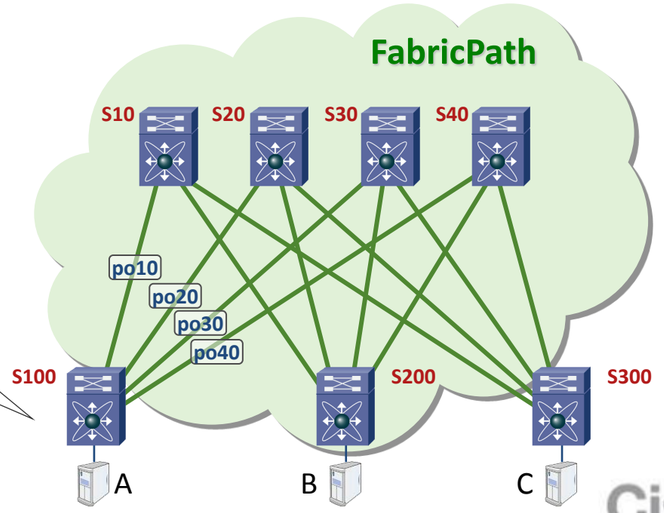
Show FabricPath Isis Route

Destination Switch ID

Next-hop interface(s)

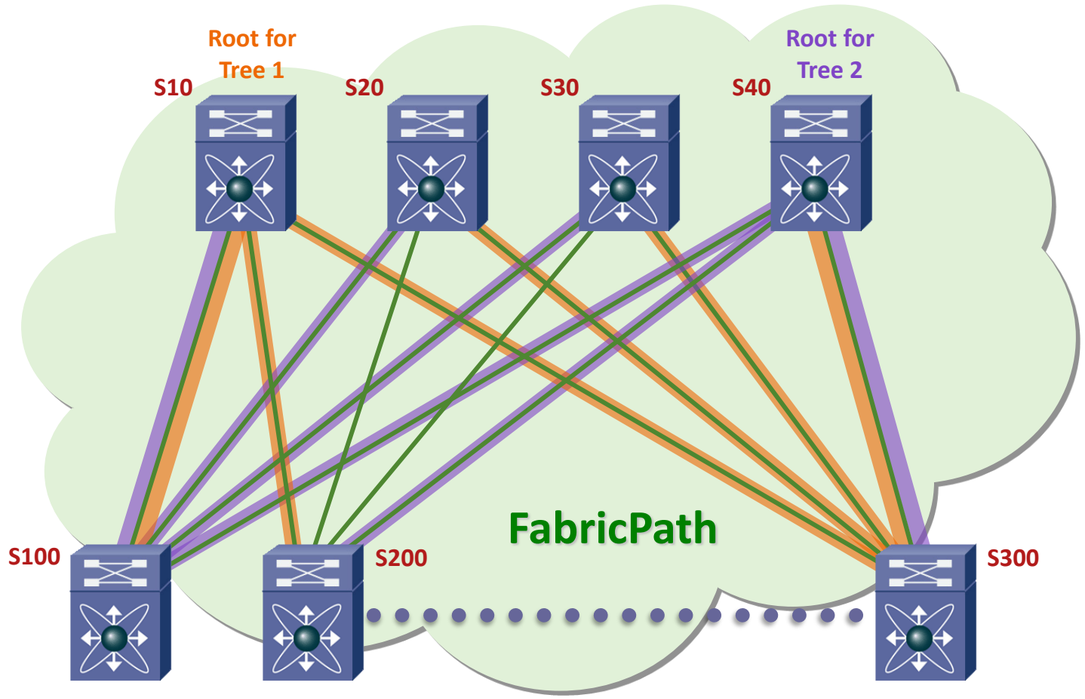
Routing metric

```
S100# sh fabricpath isis route
Fabricpath IS-IS domain: default MT-0
Topology 0, Tree 0, Swid routing table
10, L1
 via port-channel10, metric 20
20, L1
 via port-channel20, metric 20
30, L1
 via port-channel30, metric 20
40, L1
 via port-channel40, metric 20
200, L1
 via port-channel30, metric 40
 via port-channel40, metric 40
 via port-channel20, metric 40
 via port-channel10, metric 40
300, L1
 via port-channel30, metric 40
 via port-channel40, metric 40
 via port-channel20, metric 40
 via port-channel10, metric 40

S100#
```

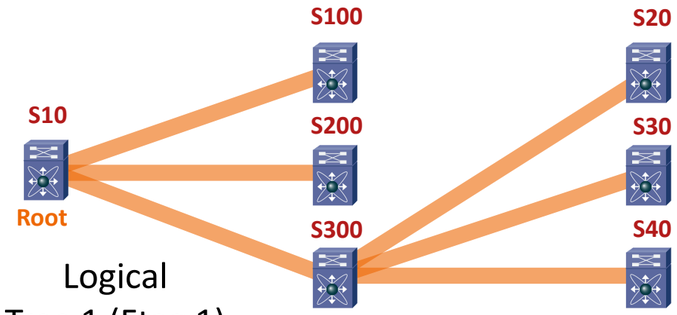


© 2013 Cisco and/or its affiliates. All rights reserved.

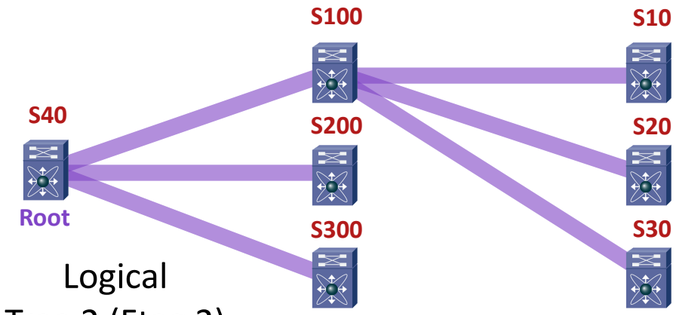
FabricPath Multidestination Trees



- Multidestination traffic constrained to loop-free trees touching all FabricPath switches
- Root switch elected for each multidestination tree in the FabricPath domain
- Network-wide identifier (Ftag) assigned toeach loop-free tree
- Support for multiple multidestination trees provides multipathing for multi-destination traffic
  - Two multidestination trees supported in NX-OS release 5.1



Logical
Tree 1 (Ftag 1)

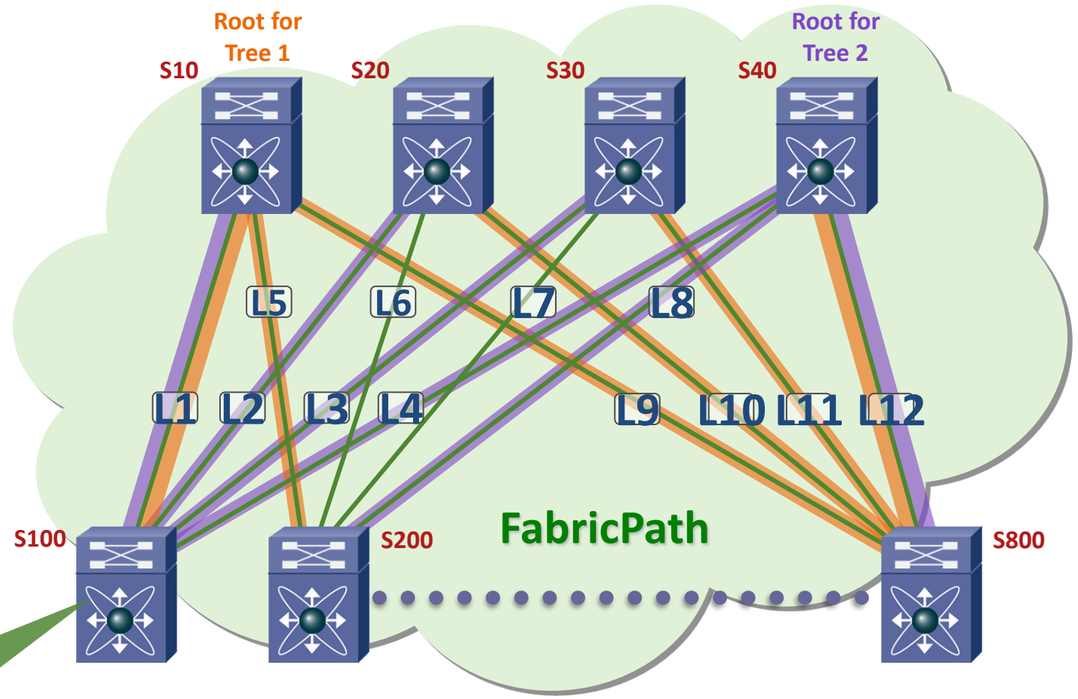


Logical
Tree 2 (Ftag 2)

© 2013 Cisco and/or its affiliates. All rights reserved.

Multidestination Trees and Role of the Ingress FabricPath Switch

- Ingress FabricPath switch determines which tree to use for each flow
- Other FabricPath switches forward based on tree selected by ingress switch
- Broadcast and unknown unicast typically use first tree
- Hash-based tree selection for IP multicast, with several configurable hash options



Multidestination Trees on Switch 100

Tree	IF
1	L1
2	L1,L2,L3,L4

© 2013 C eserved.

Display IS-IS View of Multidestination Trees

Show FabricPath Isis Trees

Destination Switch ID

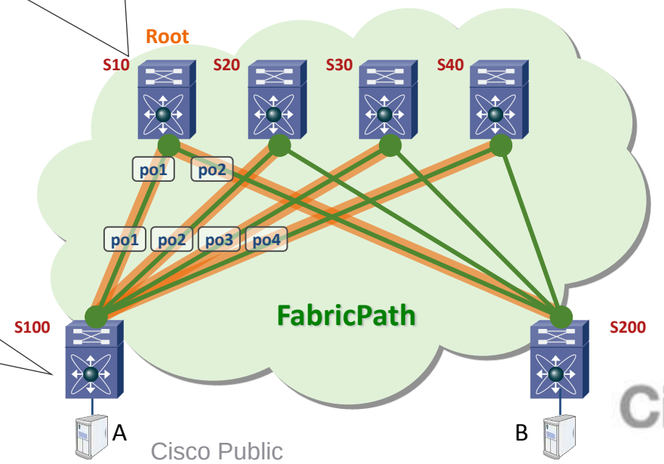
OIF to reach this switch on this tree

Metric from root to this switch

Multidestination tree

```
S10# sh fabricpath isis trees multidestination 1
Fabricpath IS-IS domain: default
Note: The metric mentioned for multidestination tree is
from the root of that tree to that switch-id

MT-0
Topology 0, Tree 1, Swid routing table
20, L1
 via port-channel1, metric 20
30, L1
 via port-channel1, metric 20
40, L1
 via port-channel1, metric 20
100, L1
 via port-channel1, metric 10
200, L1
 via port-channel2, metric 10

S10#
```

```
S100# sh fabricpath isis trees multidestination 1
Fabricpath IS-IS domain: default
Note: The metric mentioned for multidestination tree is
from the root of that tree to that switch-id

MT-0
Topology 0, Tree 1, Swid routing table
10, L1
 via port-channel1, metric 0
20, L1
 via port-channel2, metric 20
30, L1
 via port-channel3, metric 20
40, L1
 via port-channel4, metric 20
200, L1
 via port-channel1, metric 10

S100#
```



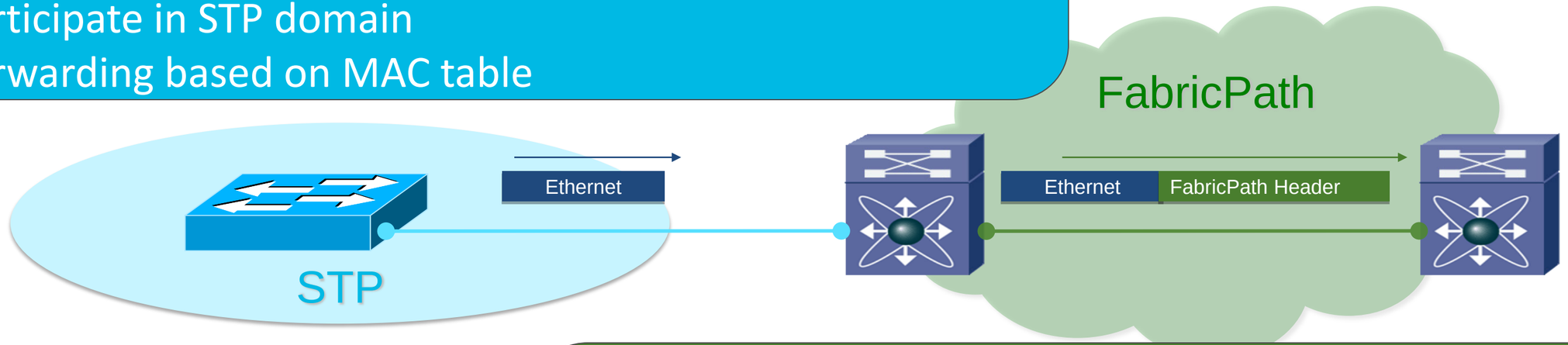
© 2013 Cisco and/or its affiliates. All rights reserved.
Cisco Public

FabricPath Data Plane

FabricPath versus Classic Ethernet Interfaces

Classic Ethernet (CE) Interface

- Interfaces connected to existing NICs and traditional network devices
- Send/receive traffic in 802.3 Ethernet frame format
- Participate in STP domain
- Forwarding based on MAC table



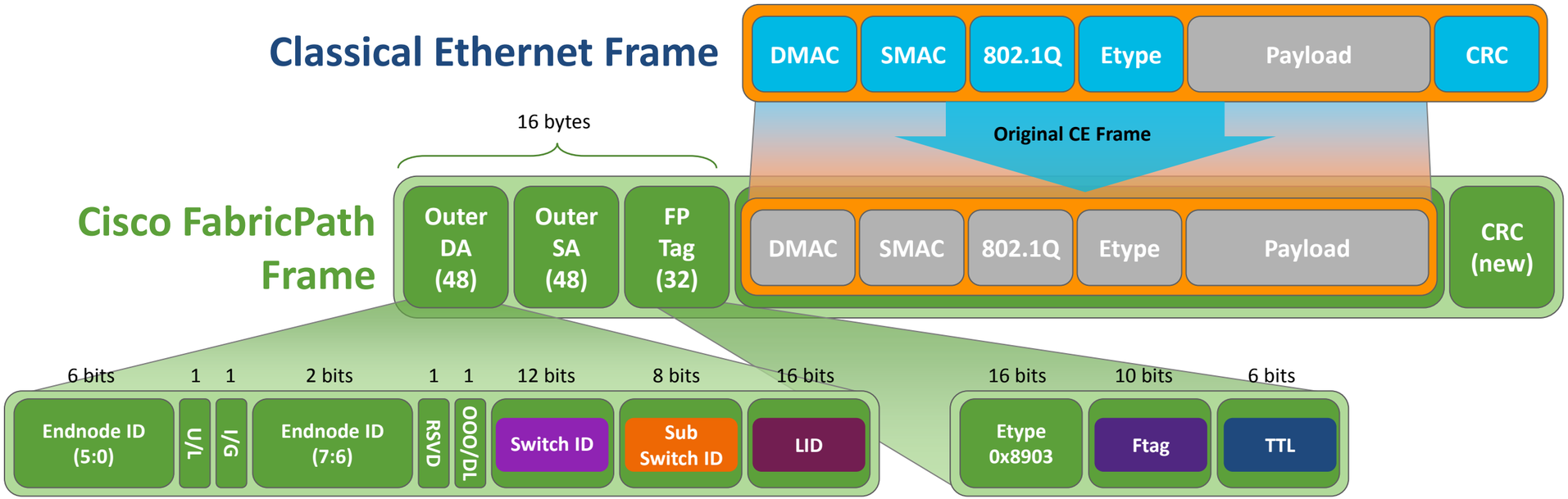
FabricPath Interface

- Interfaces connected to another FabricPath device
- Send/receive traffic with FabricPath header
- No spanning tree!!!
- No MAC learning
- Exchange topology info through L2 ISIS adjacency
- Forwarding based on ‘Switch ID Table’

© 2013 Cisco and/or its

FabricPath Encapsulation

16-Byte MAC-in-MAC Header



- **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

© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath Unicast Forwarding

Control plane:

- Routing table – FabricPath IS-IS learns switch IDs (SIDs) and builds routing table
- Multidestination trees – FabricPath IS-IS elects roots and builds multidestination forwarding trees

Data plane:

- MAC table – Hardware performs MAC table lookups to determine destination FabricPath switch (unicast) or to identify multidestination frames
- Switch table – Hardware performs destination SID lookups to forward unicast frames to other switches
- Multidestination table – Hardware selects multidestination tree to forward multidestination frames through network fabric

© 2013 Cisco and/or its affiliates. All rights reserved.

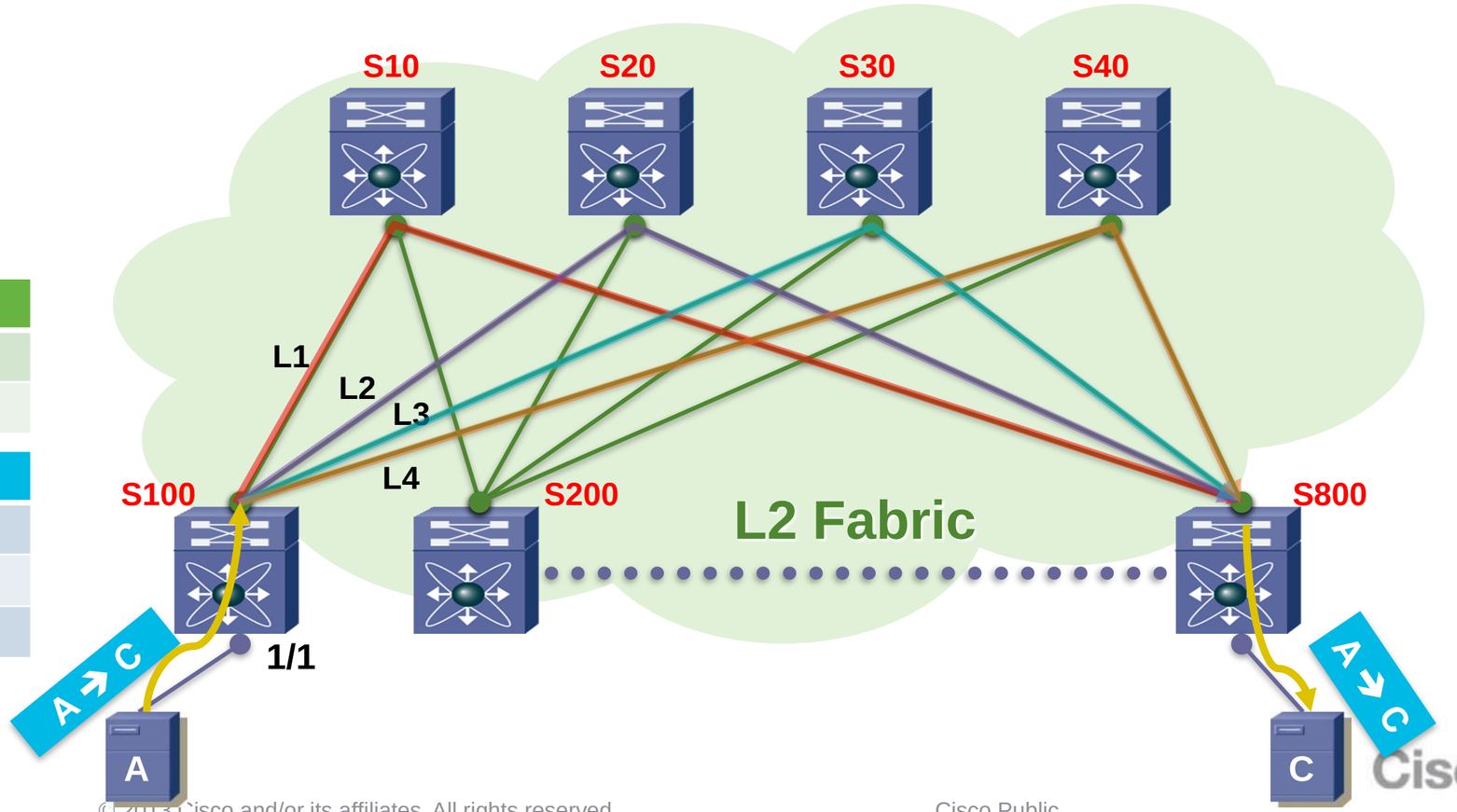
Unicast with FabricPath

Forwarding Decision Based on ‘FabricPath Routing Table’

- Support more than 2 active paths (up to 16) across the Fabric
- Increase bi-sectional bandwidth beyond port-channel
- High availability with N+1 path redundancy

Switch	IF
…	…
S800	L1, L2, L3, L4

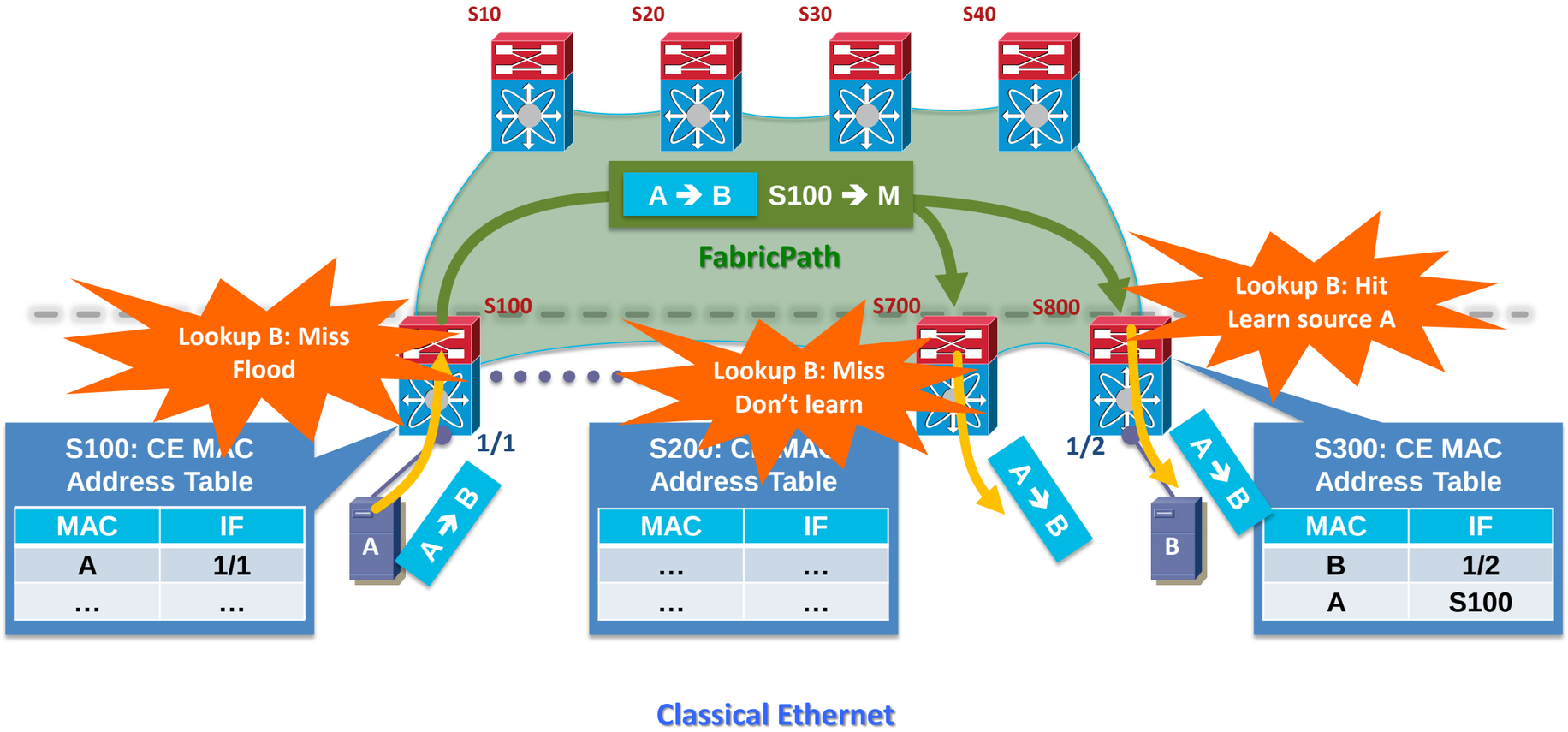
MAC	IF
A	1/1
…	…
C	S800



© 2013 Cisco and/or its affiliates. All rights reserved.

Conversational Learning

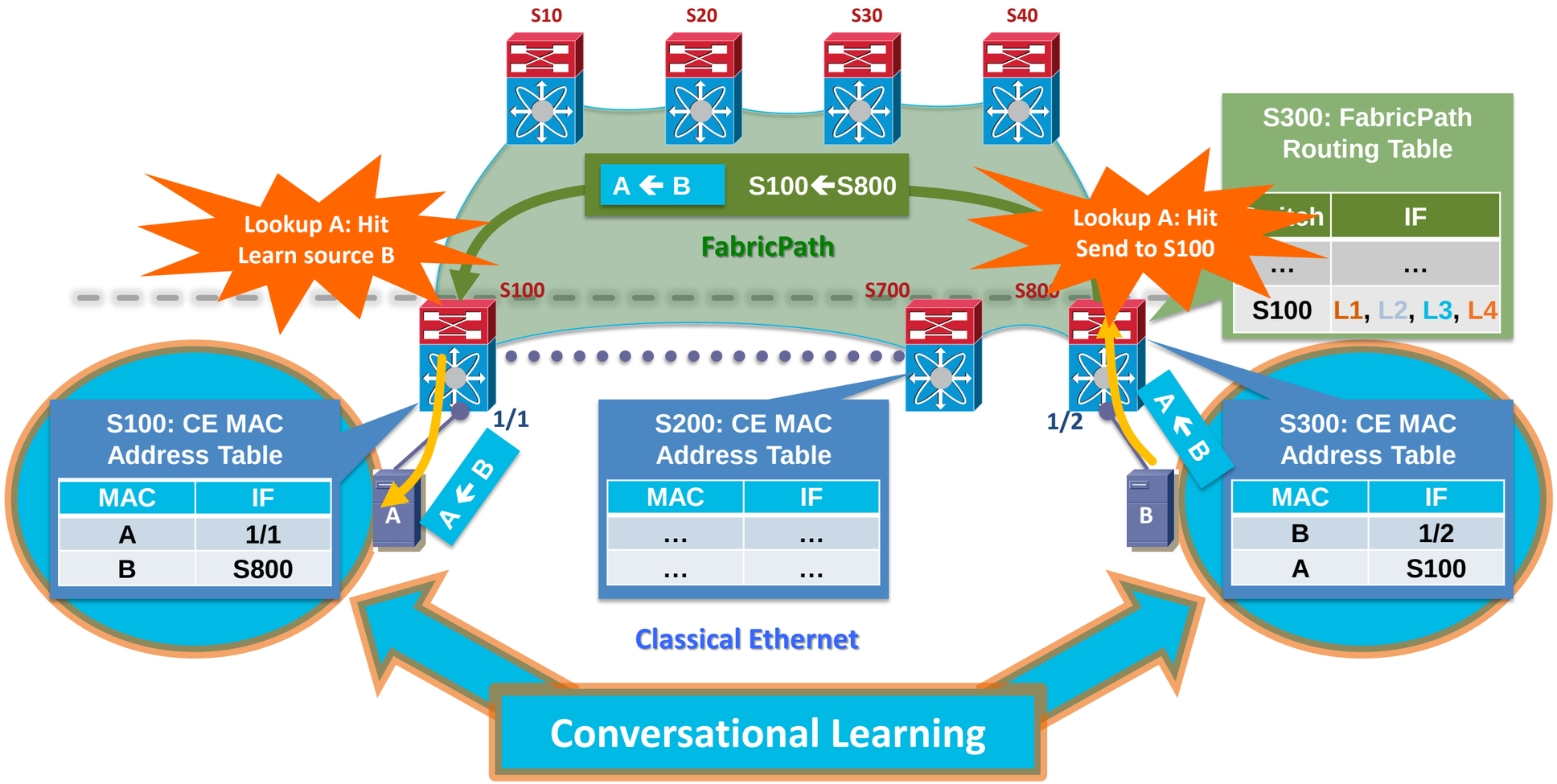
Unknown Unicast



© 2013 Cisco and/or its affiliates. All rights reserved.

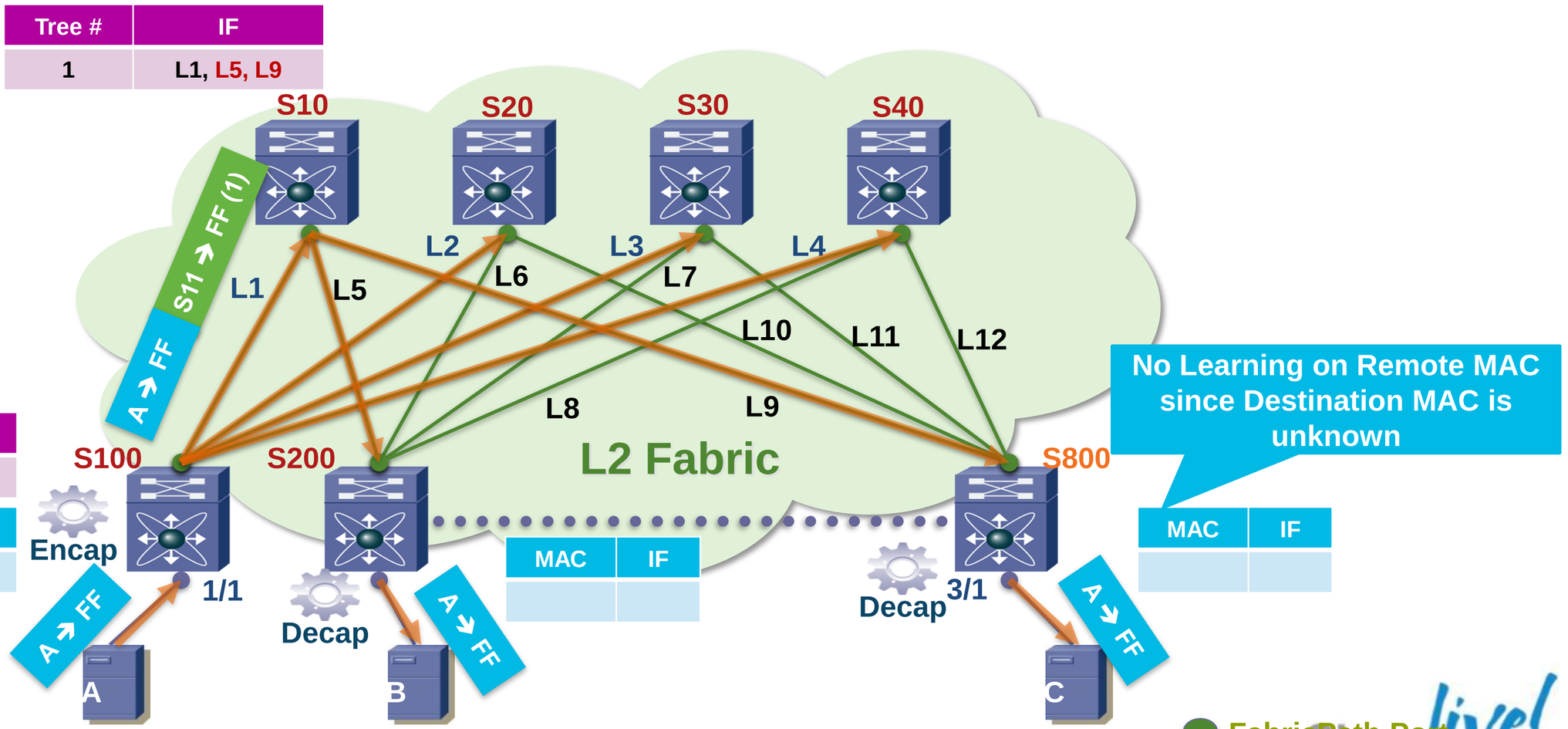
Conversational Learning

Unknown Unicast



© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath Forwarding: Broadcast



© 2013 Cisco and/or its affiliates. All rights reserved.

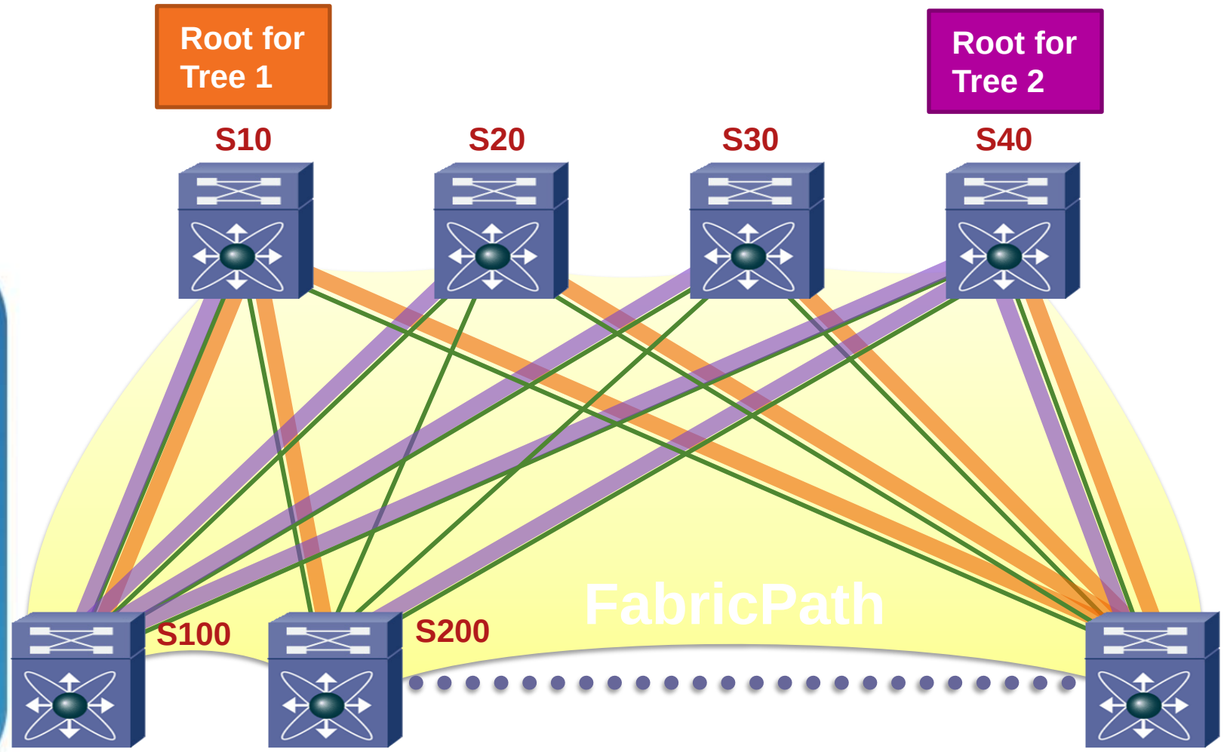
Multicast Forwarding

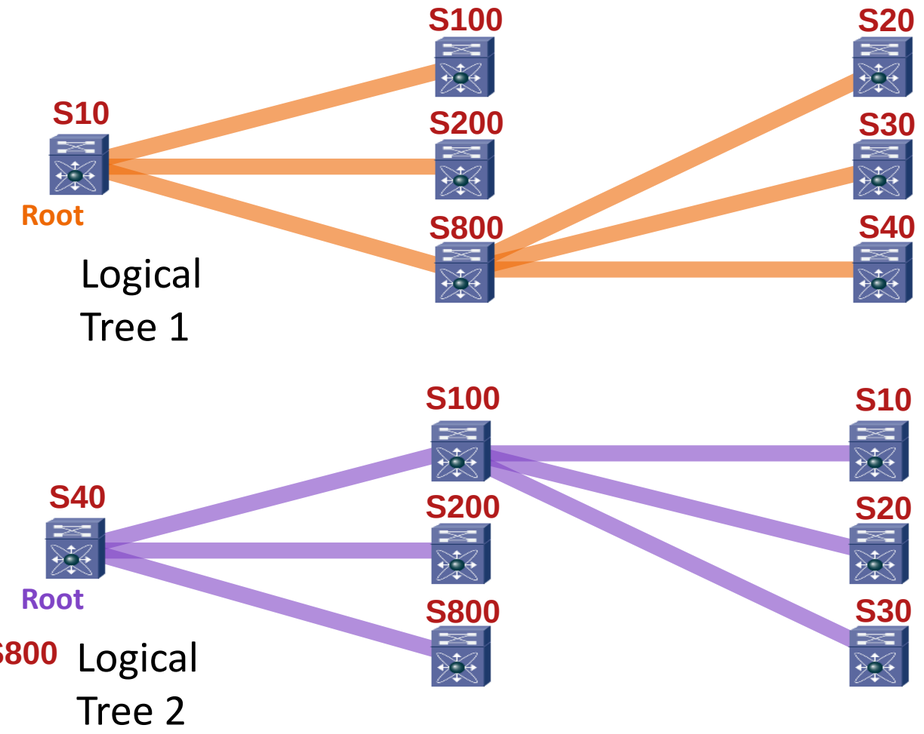
FabricPath IP Multicast

- **Control plane:**
  - Build several multidestination trees
  - Run IGMP snooping on FabricPath edge switches
  - Advertise receivers location with dedicated LSPs
- **Data plane (hardware):**
  - Selects which multidestination tree for each flow based on hash function
  - Forward traffic along selected tree

© 2013 Cisco and/or its affiliates. All rights reserved.

Multicast Trees Determination

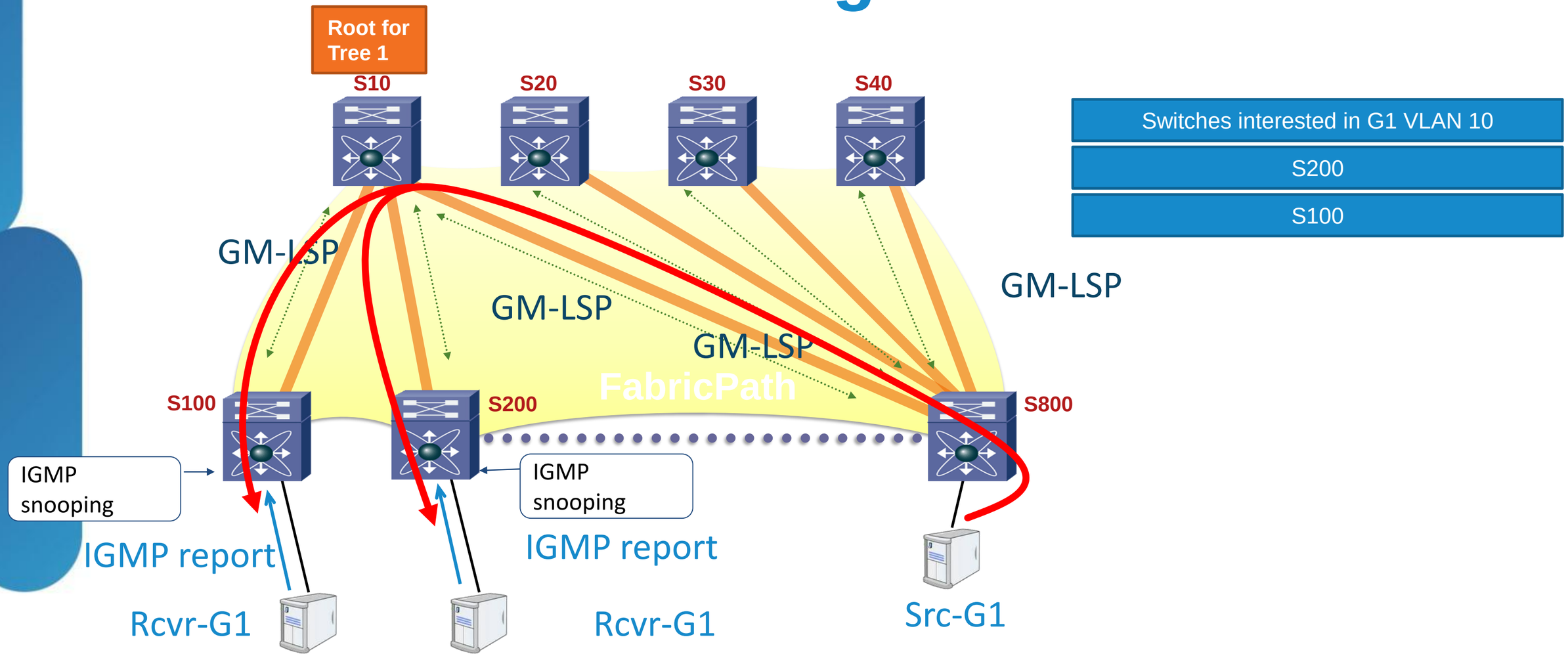




- Switch with highest priority value becomes root for primary tree
  - Highest system ID, then highest Switch ID value, in case of a tie
- Primary root designates different secondary root(s) ensuring path variety.

© 2013 Cisco and/or its affiliates. All rights reserved.

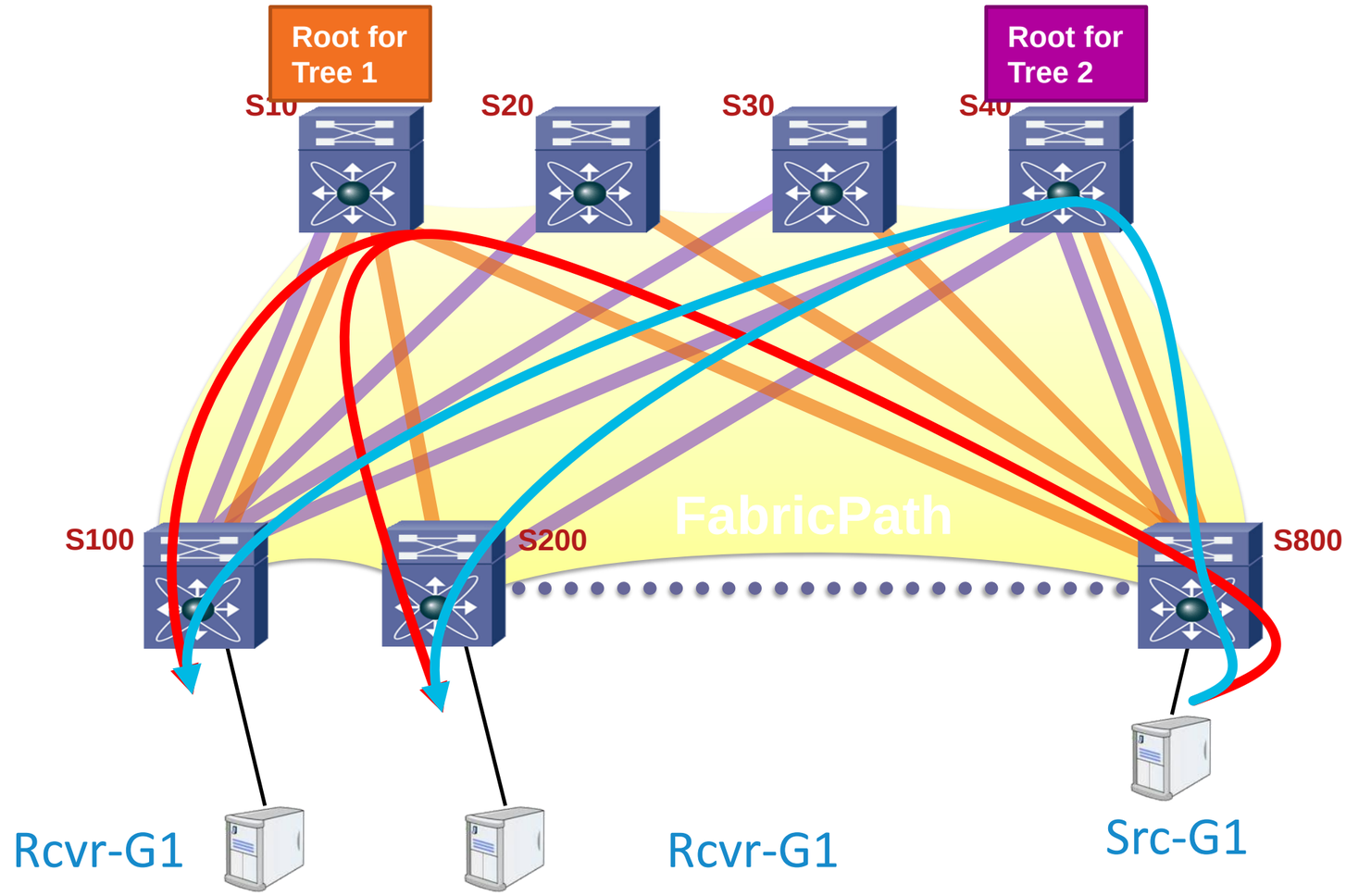
Multicast Tree Pruning



- IS-IS Group Membership LSPs contain multicast forwarding information

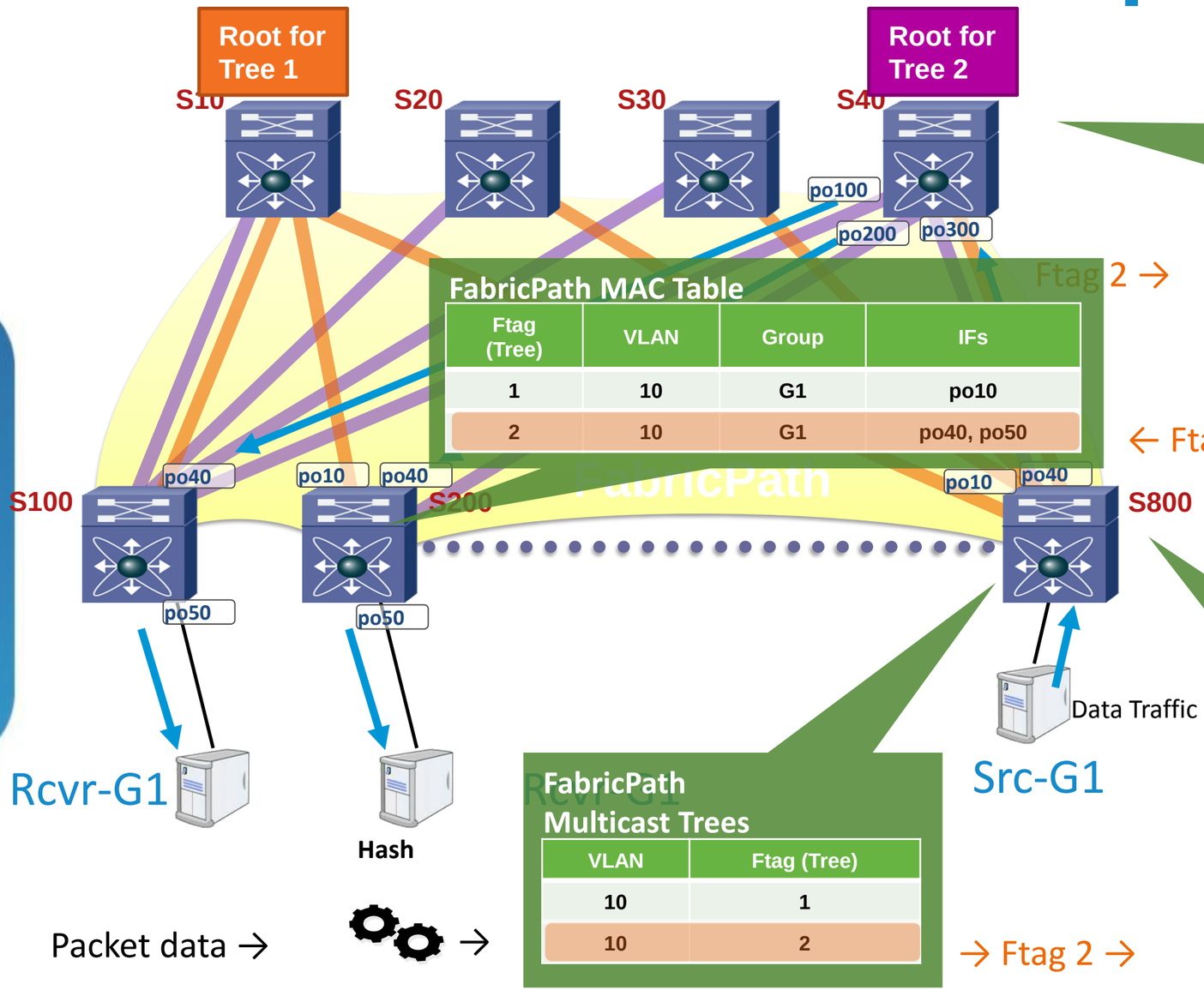
© 2013 Cisco and/or its affiliates. All rights reserved.

Multicast Load Balancing



© 2013 Cisco and/or its affiliates. All rights reserved.

Multicast Data Plane Step by Step



- **Ftag: Forwarding tag** – Unique 10-bit number identifying topology and/or distribution tree

© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath v.s TRILL

Transparent Interconnection of Lots of Links (TRILL)

- IETF standard for Layer 2 multipathing
- Driven by multiple vendors, including Cisco
- TRILL now officially moved from Draft to Proposed Standard in IETF
- Proposed Standard status means vendors can confidently begin developing TRILL-compliant software implementations
- Cisco FabricPath capable hardware is also TRILL capable

http://datatracker.ietf.org/wg/trill/

© 2013 Cisco and/or its affiliates. All rights reserved.

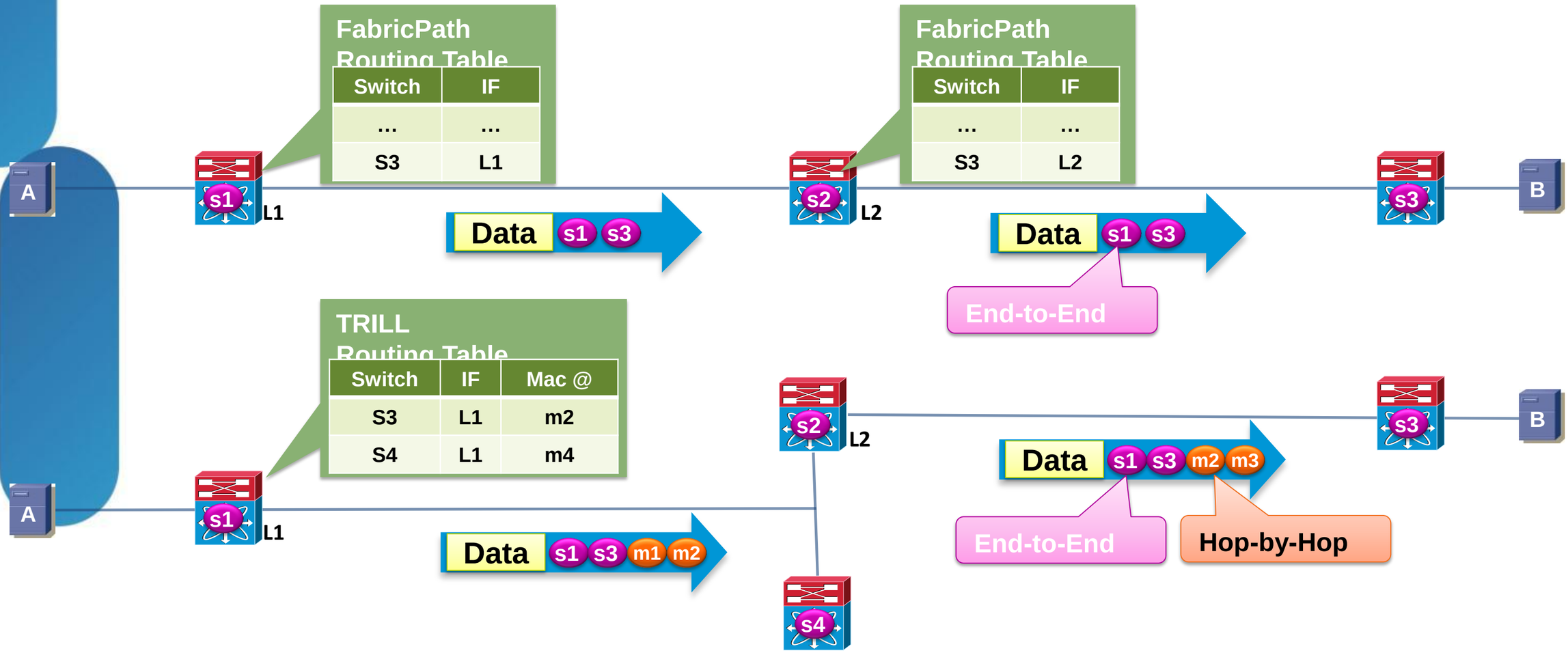
FabricPath vs. TRILL Overview

	FabricPath	TRILL
Frame routing (ECMP, TTL, RPFC etc…)	Yes	Yes
vPC+	Yes	No
FHRP active/active	Yes	No
Multiple topologies	Yes	No
Conversational learning	Yes	No
Inter-switch links	Point-to-point only	Point-to-point OR shared

- FabricPath will provide a TRILL mode with a software upgrade (hardware is already TRILL capable)
- Cisco will push FabricPath specific enhancements to TRILL

© 2013 Cisco and/or its affiliates. All rights reserved.

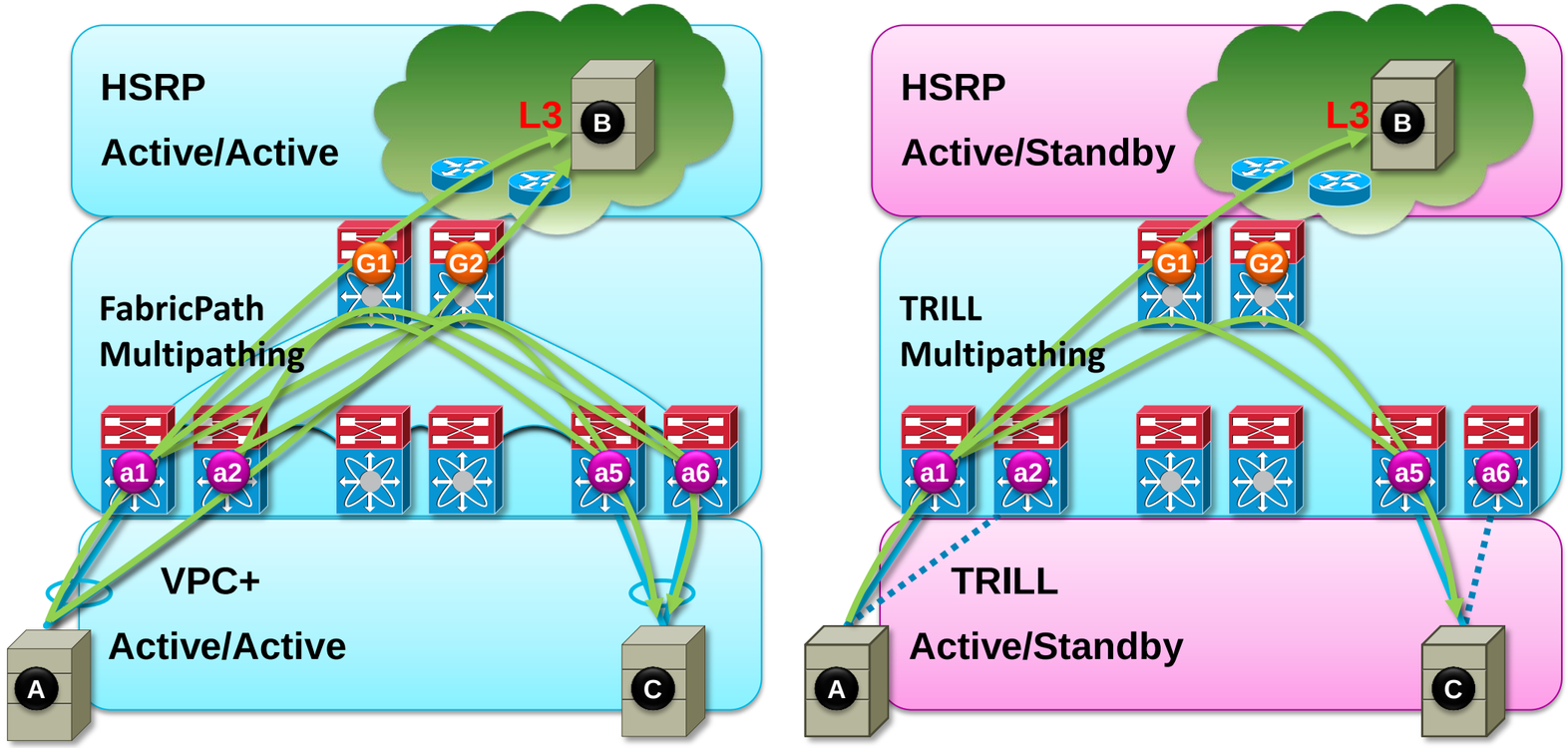
FabricPath vs. TRILL: Encapsulation



© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath vs. TRILL:

Multipathing

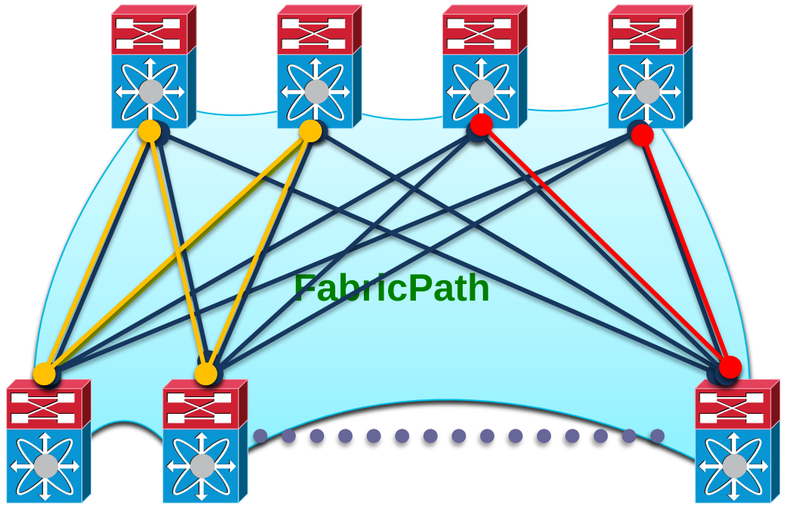


- End-to-end multipathing (L2 edge, Fabric, L3 edge) provides resiliency and fast convergence

© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath vs. TRILL:

FabricPath Multiple Topologies



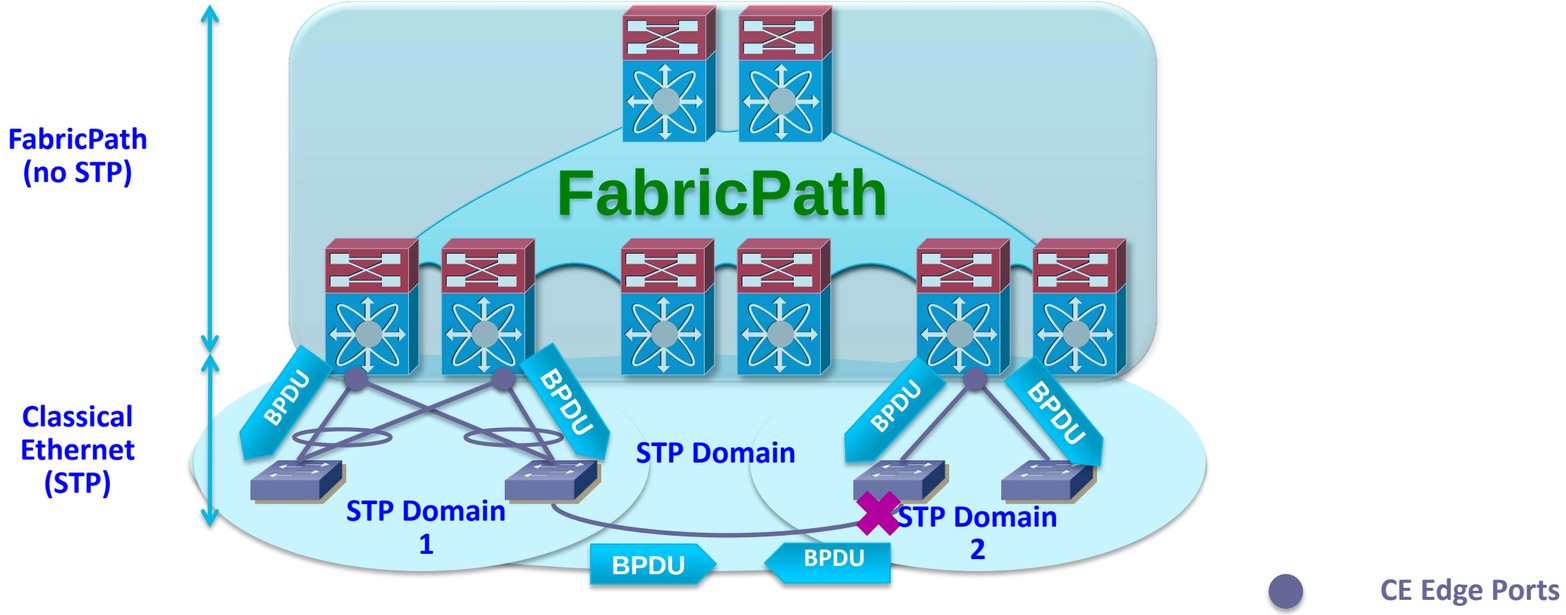
- **Topology**: A group of links in the Fabric.

- By default, all the links are part of topology 0.
- Other topologies are created by assigning a subset of the links to them.
- A link can belong to several topologies
- A VLAN is mapped to a unique topology
- Topologies are used for VLAN pruning, security, traffic engineering etc…

© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath vs. TRILL:

FabricPath Simple STP Interaction



- The Fabric looks like a single bridge:
  - It sends the same STP information on all edge ports
  - It expects to be the root of the STP for now (edge ports will block if they receive better information)
- **No BPDUs** are forwarded across the fabric
- An optional mechanism allows propagating TCNs if needed

© 2013 Cisco and/or its affiliates. All rights reserved.

Summary – FabricPath Technology

- Bandwidth
  - Multi pathing (ECMP)
  - Optimal paths
- Scale
  - Conversational MAC Learning
  - Efficient Flooding
  - Multiple Topologie
- Transparent to Existing Systems
  - VPC+
  - Edge Routing Integration

© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath Network Designs

Benefits of FabricPath Designs

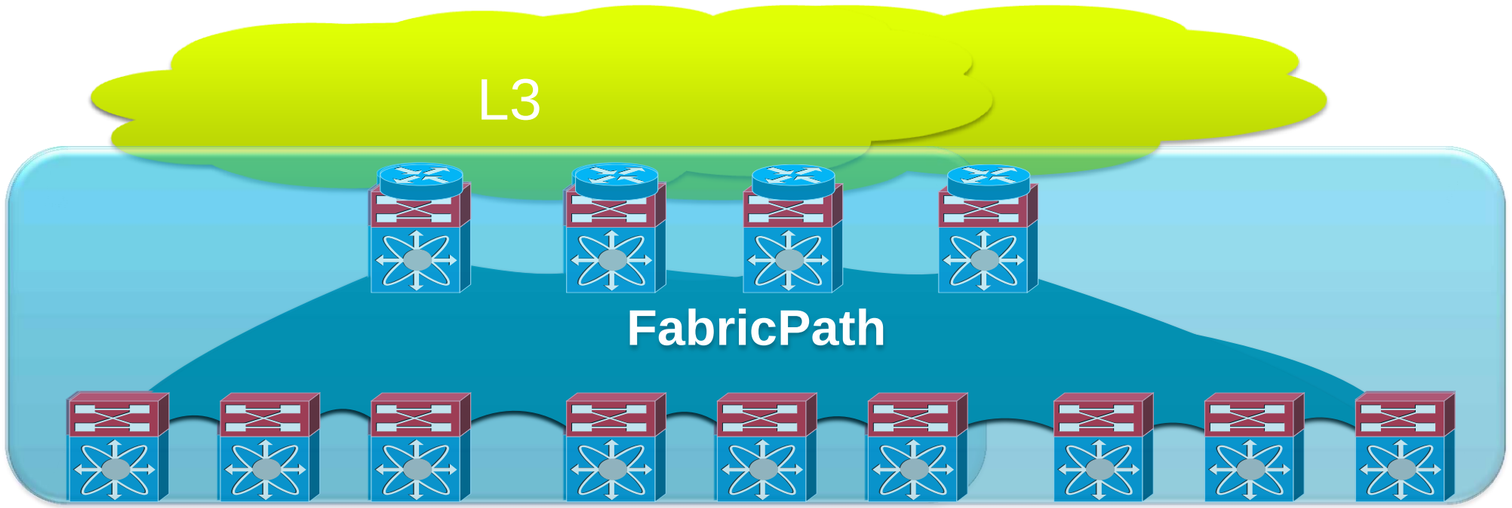
- Configuration simplicity
- Independence from / elimination of Spanning-Tree Protocol
- Deterministic throughput and latency
- Multi-way load-sharing for unicast and multicast at Layer 2
- Direct/optimised communication paths
- “VLAN anywhere” providing flexibility, L2 adjacency, and VM mobility
- Layer 2 domain scalability (ARP, MAC)
- Loop mitigation (TTL, RPF checks)
- Better stability and convergence characteristics

© 2013 Cisco and/or its affiliates. All rights reserved.

FabricPath Flexibility

The Network Can Evolve with no Disruption

- Need more edge ports? → Add more leaf switches
- Need more bandwidth? → Add more links and spines



© 2013 Cisco and/or its affiliates. All rights reserved.

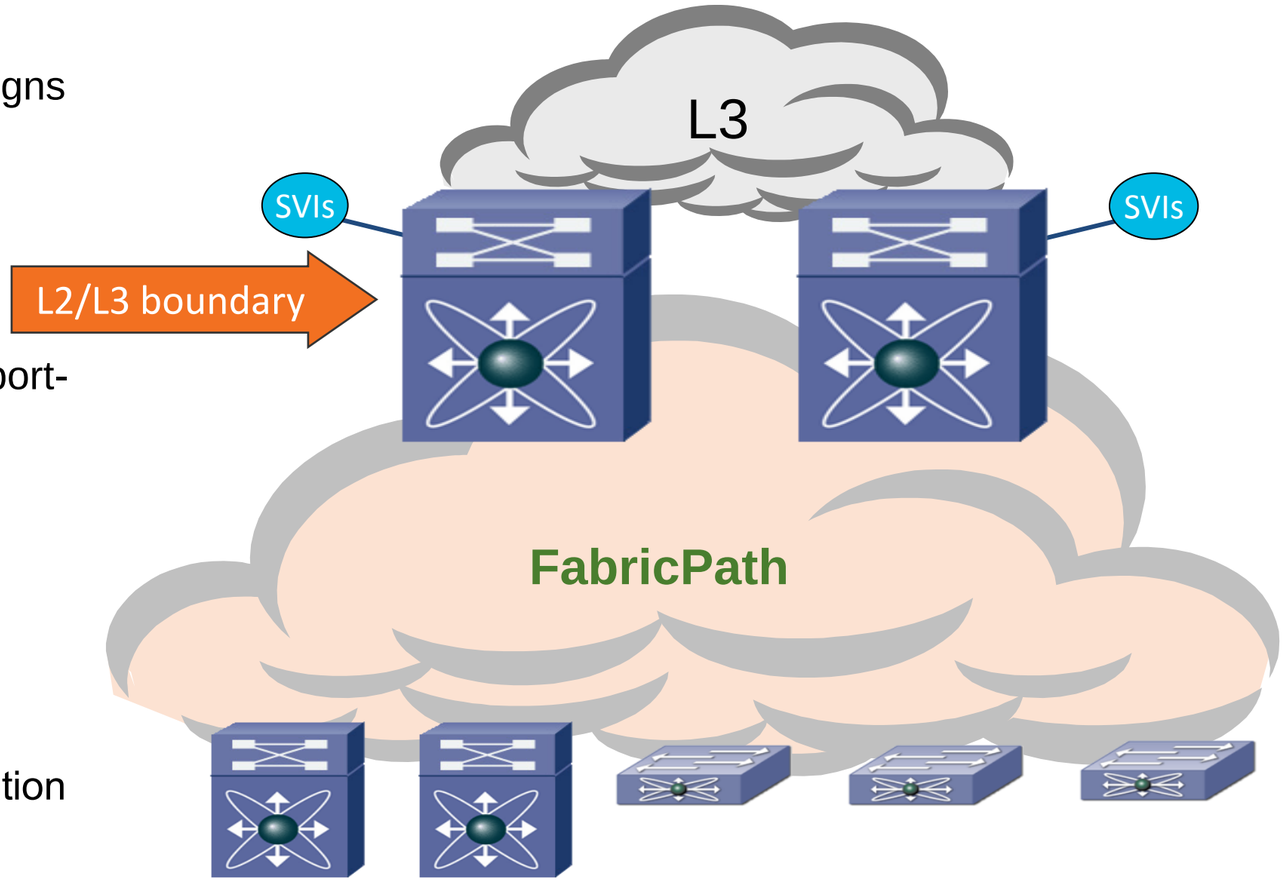
Routing at Aggregation

Two Spine Design

- Simplest design option
- Extension of traditional aggregation/access designs

Immediate benefits:

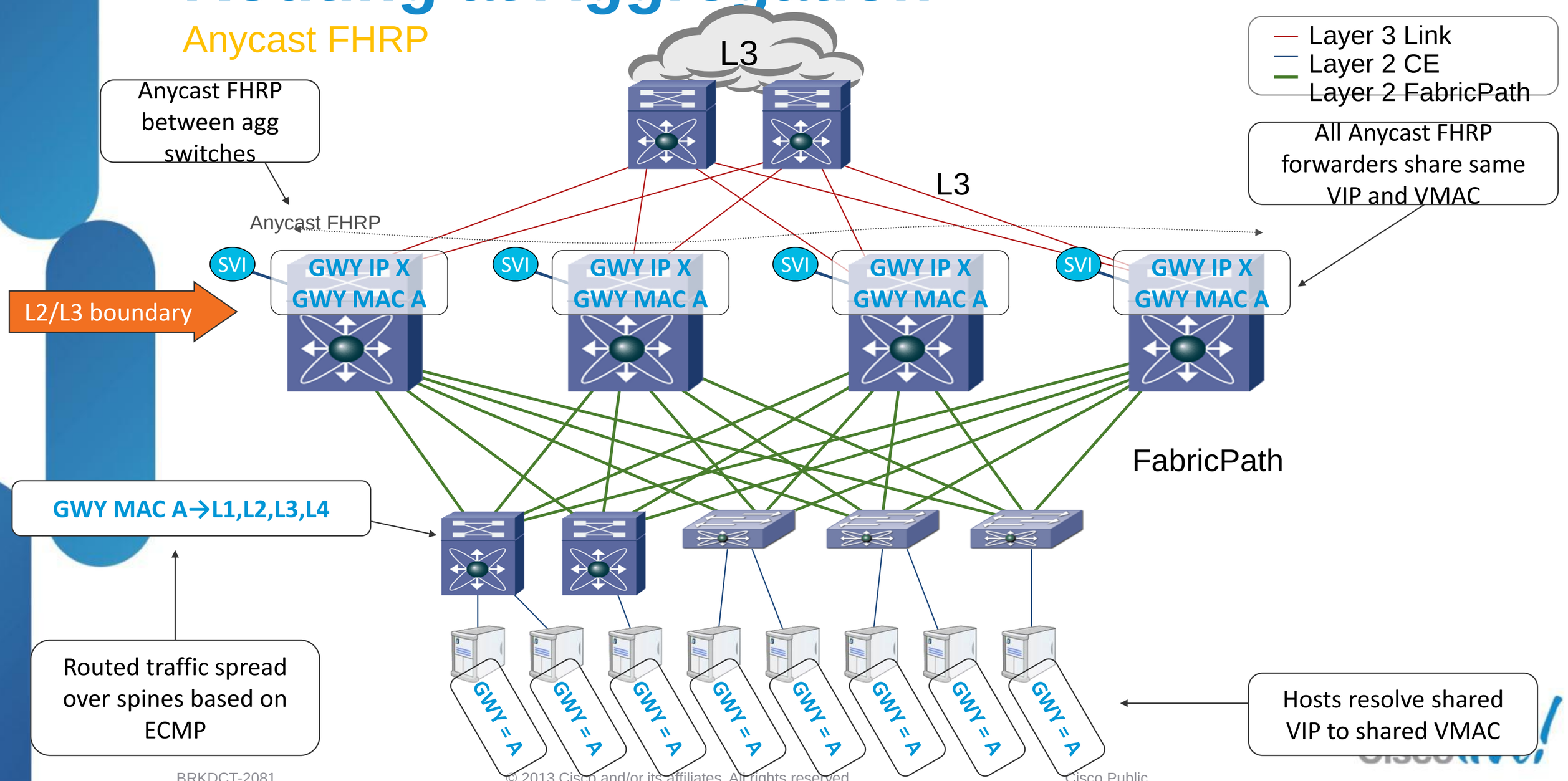
- Simplified configuration
- Removal of STP
- Traffic distribution over all uplinks without VPC port-channels
- Active/active gateways
- “VLAN anywhere” at access layer
- Topological flexibility
  - Direct-path forwarding option
  - Easily provision additional access↔aggregation bandwidth
  - Easily deploy L4-7 services
  - Option for VPC+ for legacy access switches



© 2013 Cisco and/or its affiliates. All rights reserved.

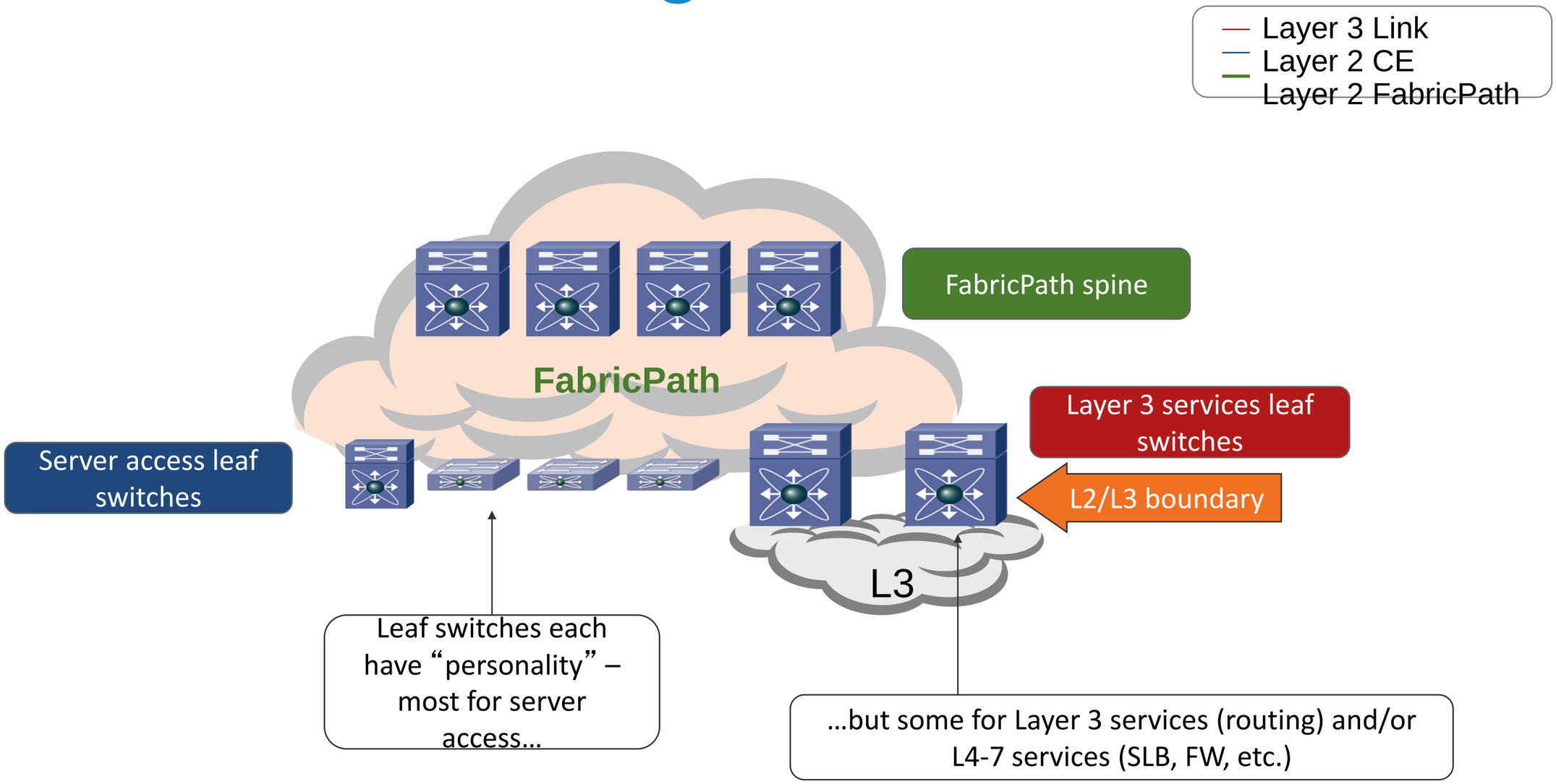
Routing at Aggregation

Anycast FHRP



© 2013 Cisco and/or its affiliates. All rights reserved.

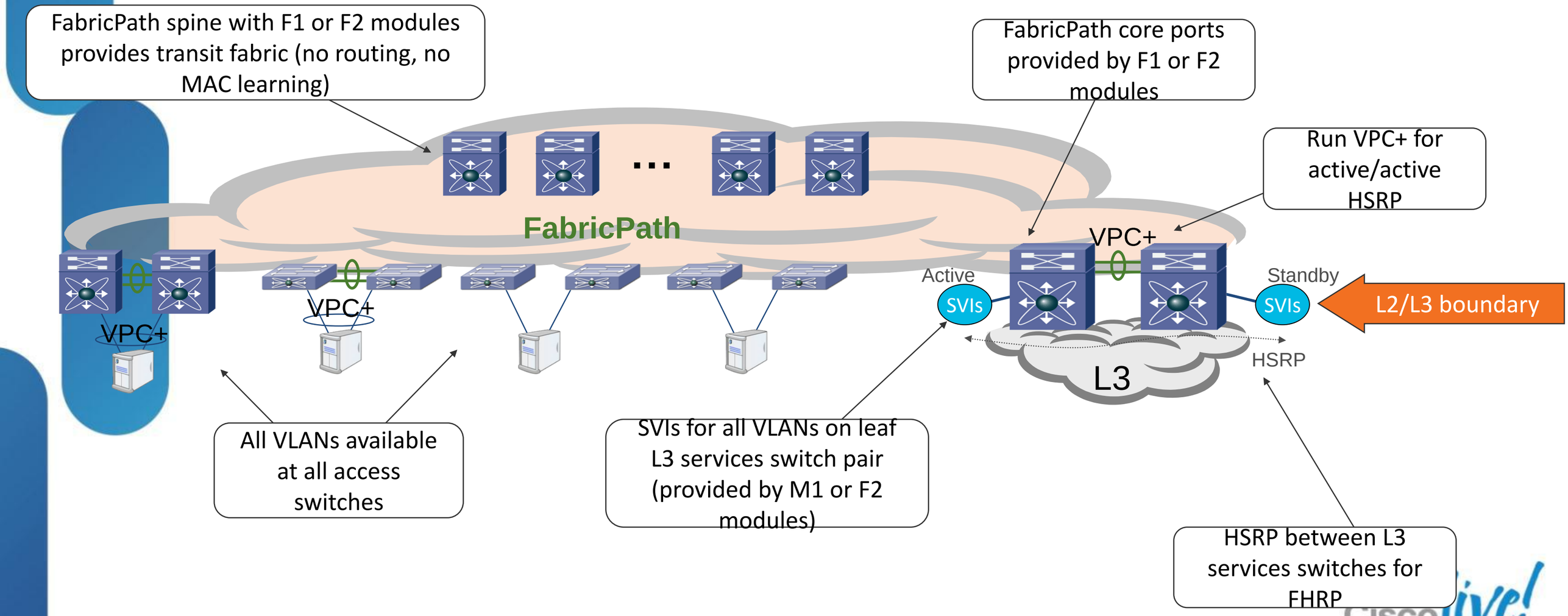
Centralised Routing



© 2013 Cisco and/or its affiliates. All rights reserved.

Centralised Routing

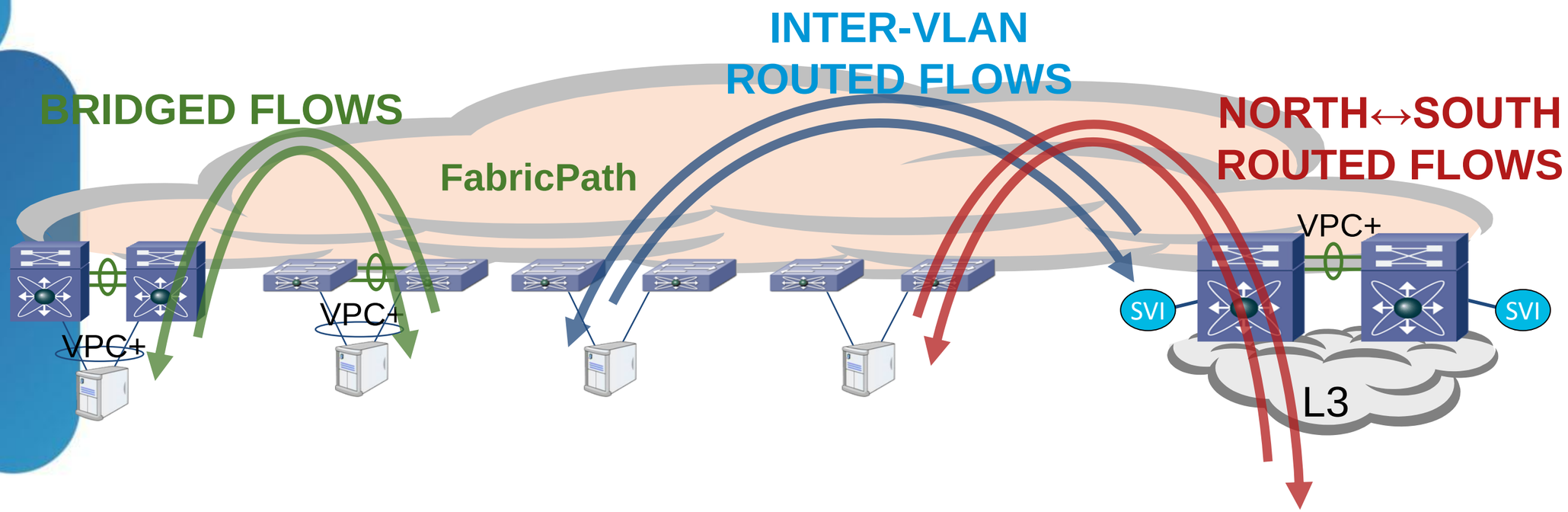
FabricPath-Connected Leaf



© 2013 Cisco and/or its affiliates. All rights reserved.

Centralised Routing

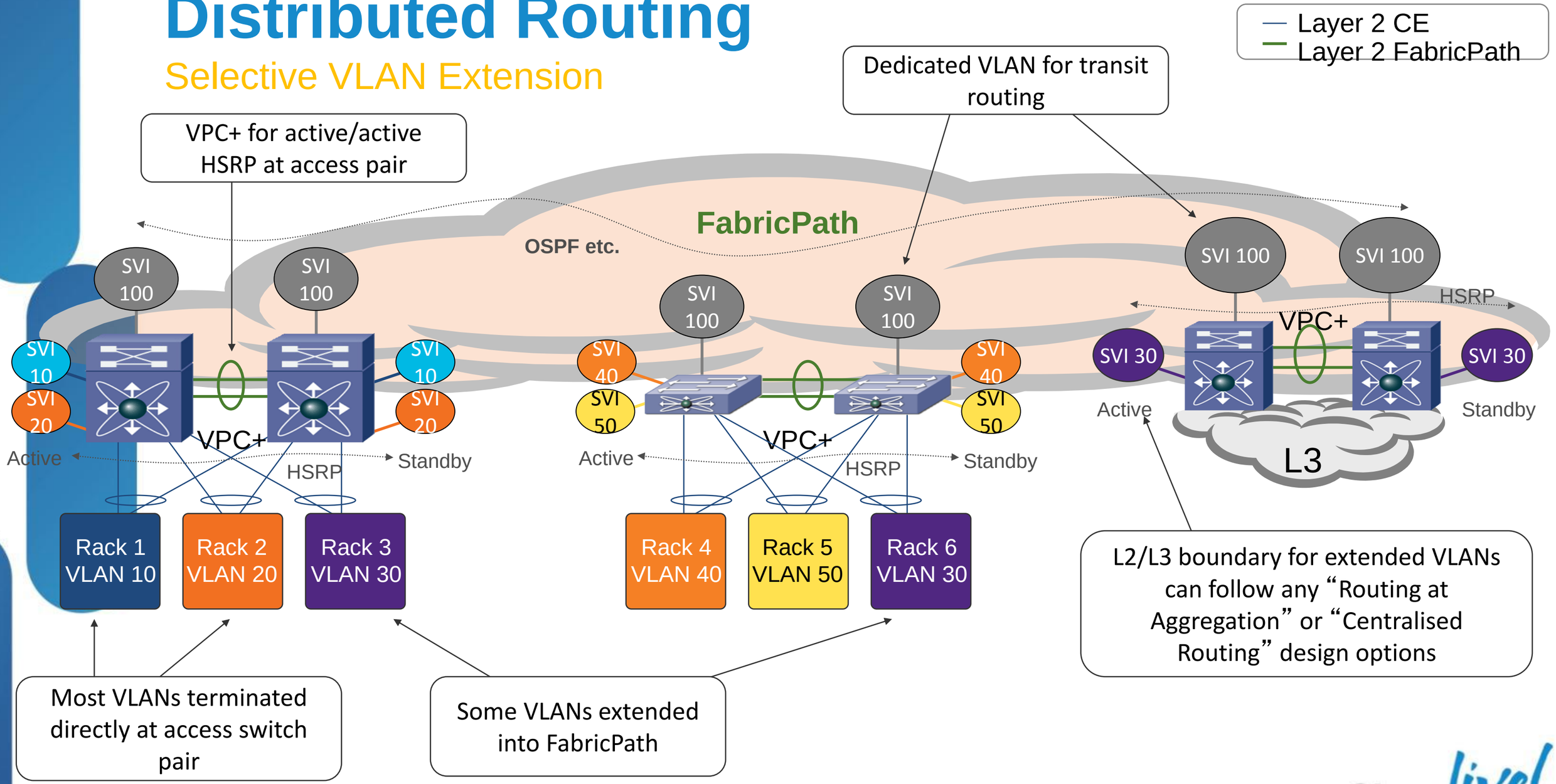
FabricPath-Connected Leaf



© 2013 Cisco and/or its affiliates. All rights reserved.

Distributed Routing

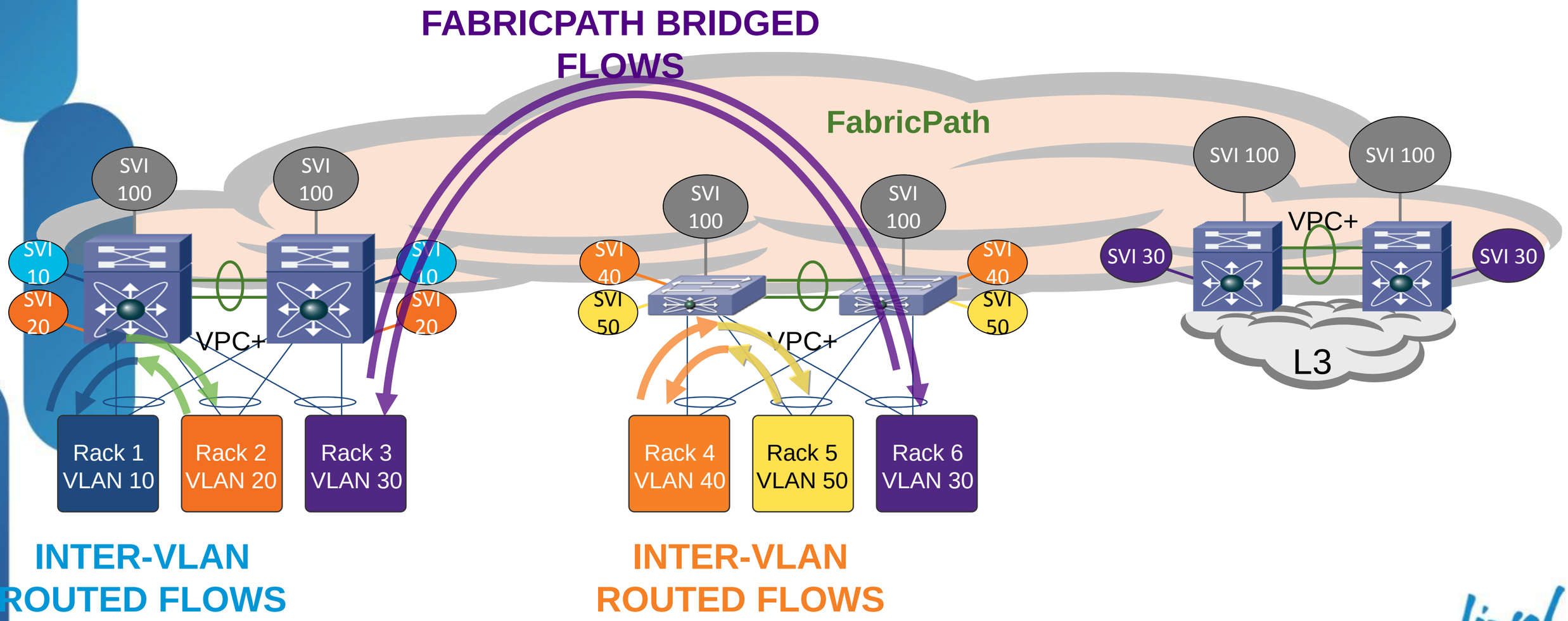
Selective VLAN Extension



© 2013 Cisco and/or its affiliates. All rights reserved.

Distributed Routing

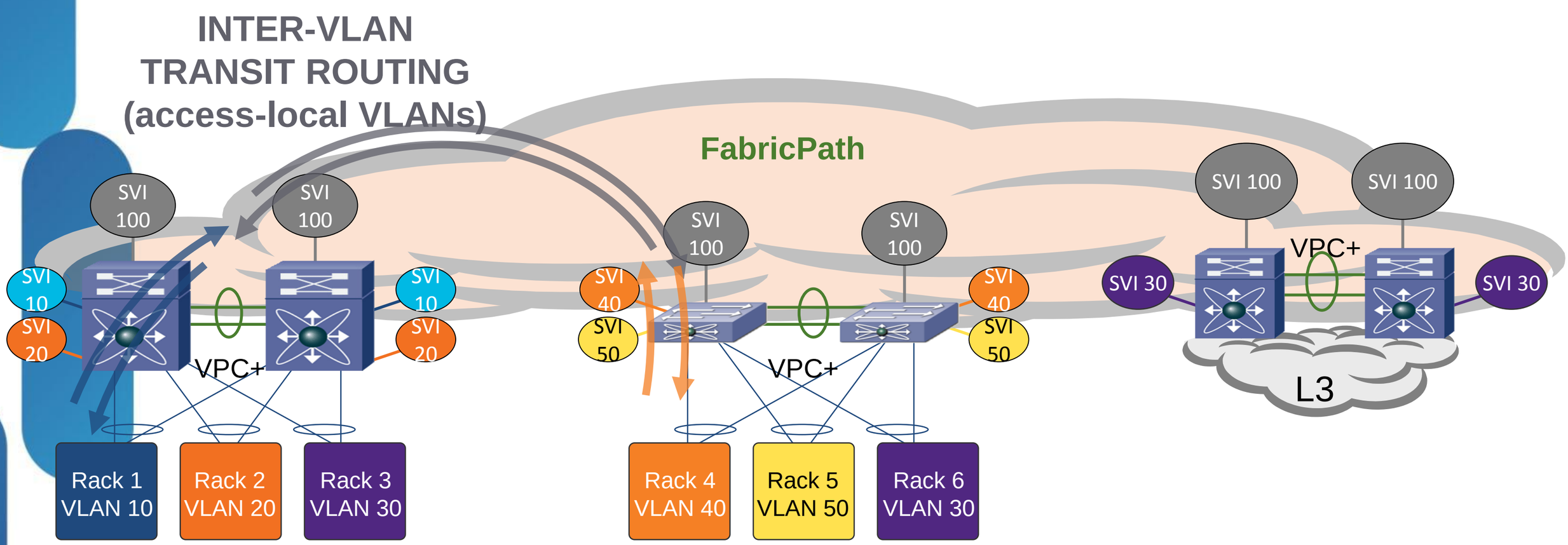
Selective VLAN Extension



© 2013 Cisco and/or its affiliates. All rights reserved.

Distributed Routing

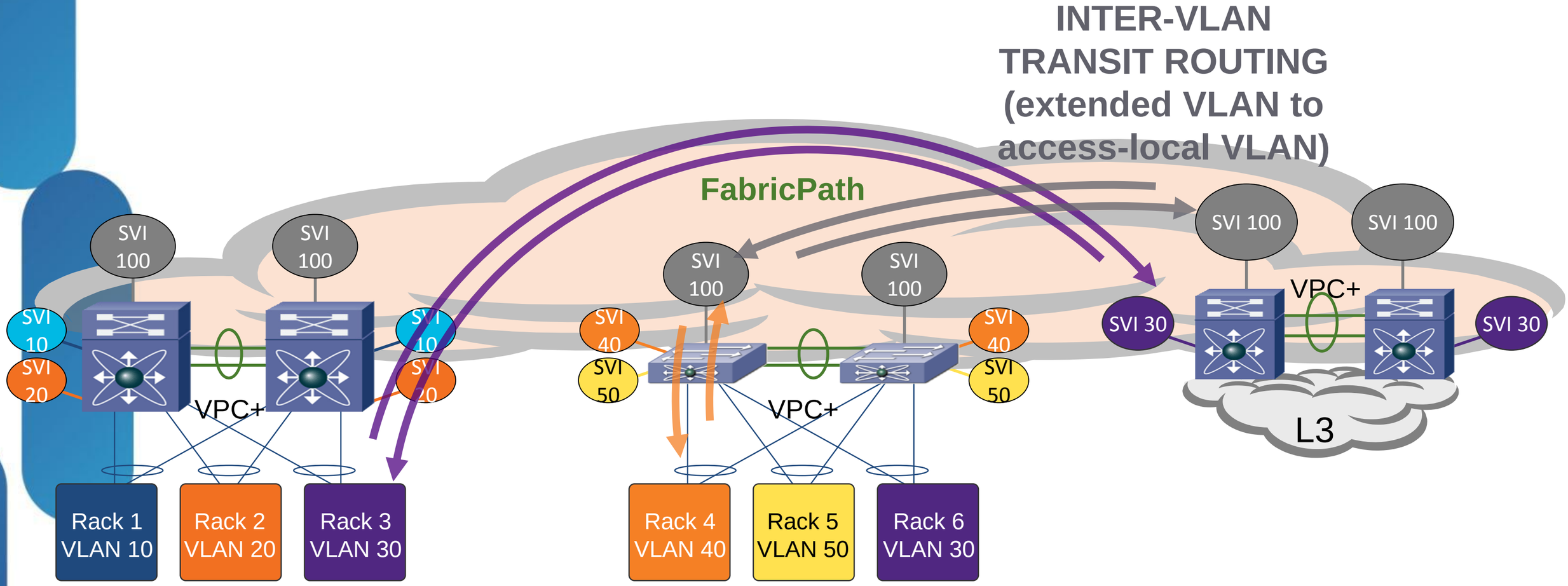
Selective VLAN Extension



© 2013 Cisco and/or its affiliates. All rights reserved.

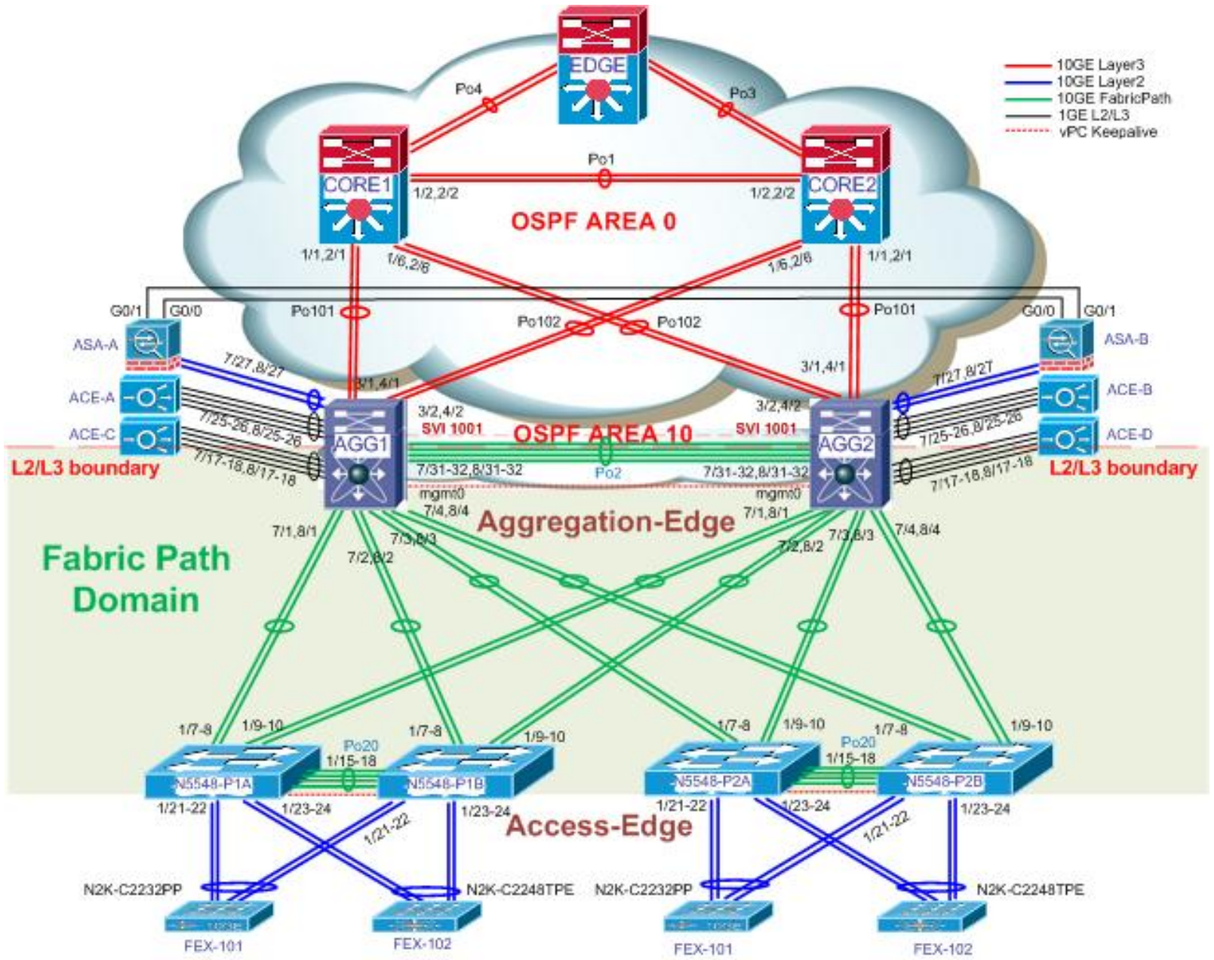
Distributed Routing

Selective VLAN Extension



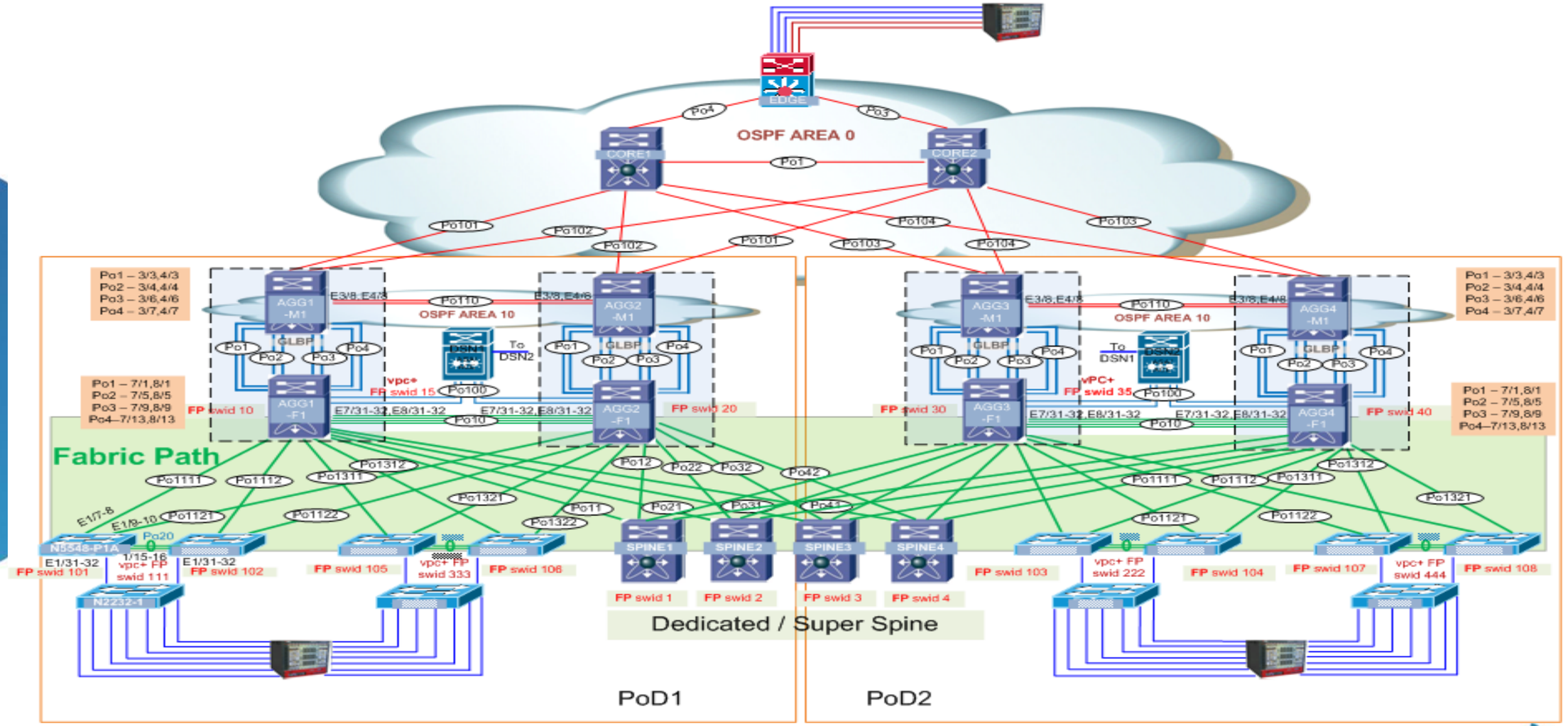
© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC 3.0 Typical DC Topology



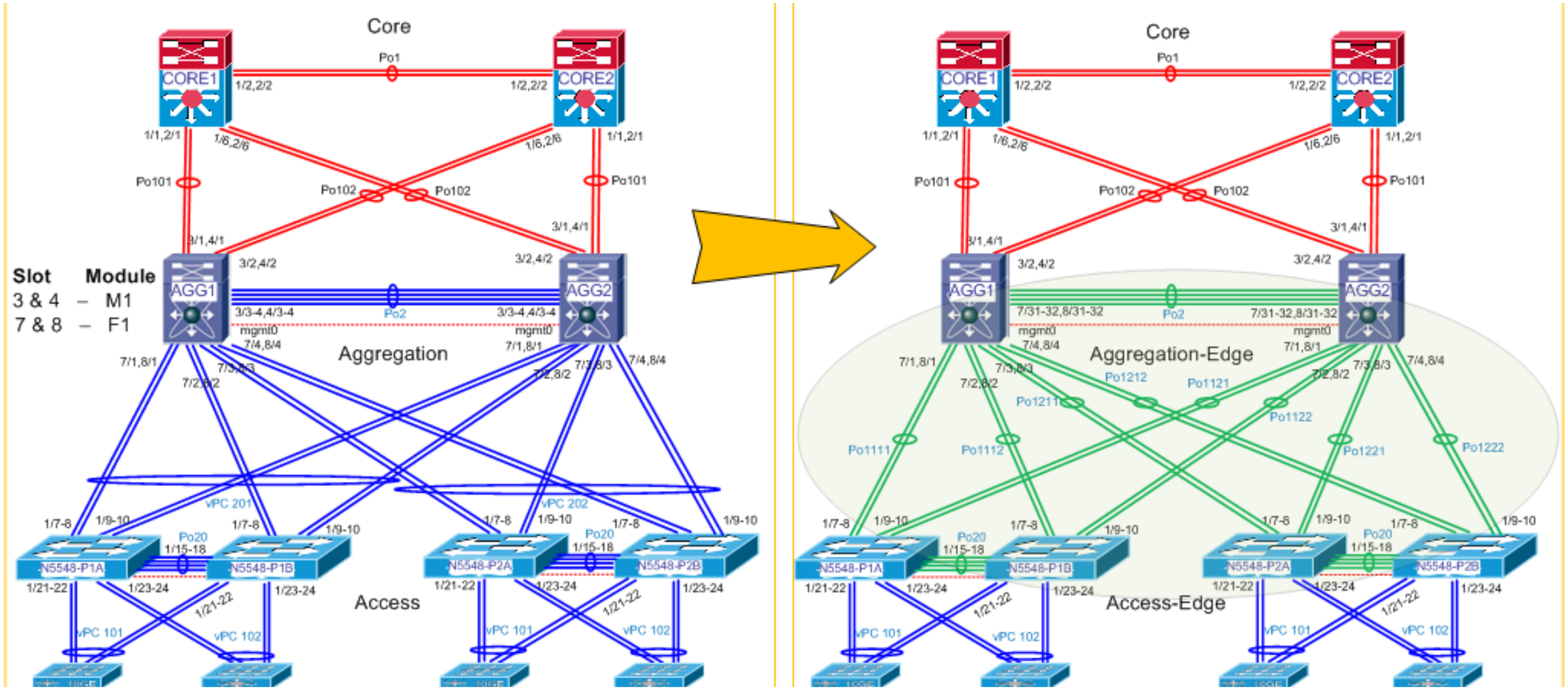
© 2013 Cisco and/or its affiliates. All rights reserved.

VMDC 3.0 – Extended DC Topology



© 2013 Cisco and/or its affiliates. All rights reserved.

vPC to FabricPath Migration in a Single POD



Whitepaper :

http://www.cisco.com/en/US/prod/collateral/switches/ps9441/ps9402/white_paper_c11-709336.html

© 2013 Cisco and/or its affiliates. All rights reserved.

Conclusion

- FabricPath is simple, keeps the attractive aspects of Layer 2
  - Transparent to L3 protocols
  - No addressing, simple configuration and deployment
- FabricPath is efficient
  - High bi-sectional bandwidth (ECMP)
  - Optimal path between any two nodes
- FabricPath is scalable
  - Can extend a bridged domain without extending the risks generally associated to Layer 2 (frame routing, TTL, RPFC)

© 2013 Cisco and/or its affiliates. All rights reserved.

Q & A

Complete Your Online Session Evaluation

Give us your feedback and receive a Cisco Live 2013 Polo Shirt!

Complete your Overall Event Survey and 5 Session Evaluations.

- Directly from your mobile device on the Cisco Live Mobile App
- By visiting the Cisco Live Mobile Site www.ciscoliveaustralia.com/mobile
- Visit any Cisco Live Internet Station located throughout the venue

Polo Shirts can be collected in the World of Solutions on Friday 8 March 12:00pm-2:00pm

Cisco live! 365

Don’t forget to activate your Cisco Live 365 account for access to all session material, communities, and on-demand and live activities throughout the year. Log into your Cisco Live portal and click the "Enter Cisco Live 365" button.
www.ciscoliveaustralia.com/portal/login.ww

Cisco live!

© 2013 Cisco and/or its affiliates. All rights reserved.

© 2013 Cisco and/or its affiliates. All rights reserved.