White Paper

Cisco FabricPath Best Practices

Updated February 2016

Contents

What You Will Learn

Cisco® FabricPath has been deployed and running in the field since the end of 2010, and thousands of customers have now acquired Cisco FabricPath licenses for their networks. This document introduces some best practices that are mainly the results of their experiences.

A best practice is a recommendation based on Cisco internal testing and the experience accumulated by customers in their production networks. A best practice is not a rule. A configuration that is beneficial for the majority of the use cases can be described as a best practice but may not be applicable to a particular network design. To help customers determine whether a best practice is applicable, this document details the rationale for each best practice. Cisco support is not bound to strict compliance with the best practices listed in this document.

This document is not an introduction to Cisco FabricPath. A basic knowledge of Cisco FabricPath is assumed. Refer to the Cisco FabricPath white paper and Cisco Nexus 7000 Series NX-OS FabricPath Configuration Guide for more information.

This document is not a network design guide either. It does not attempt to describe each and every possible use case of the technology, and the network topology presented here simply illustrates the use of some common recommended configurations.

Following a high-level overview of the network topology that will be used as an example, the best practices are introduced. More detailed explanations of the best practices are provided when needed.

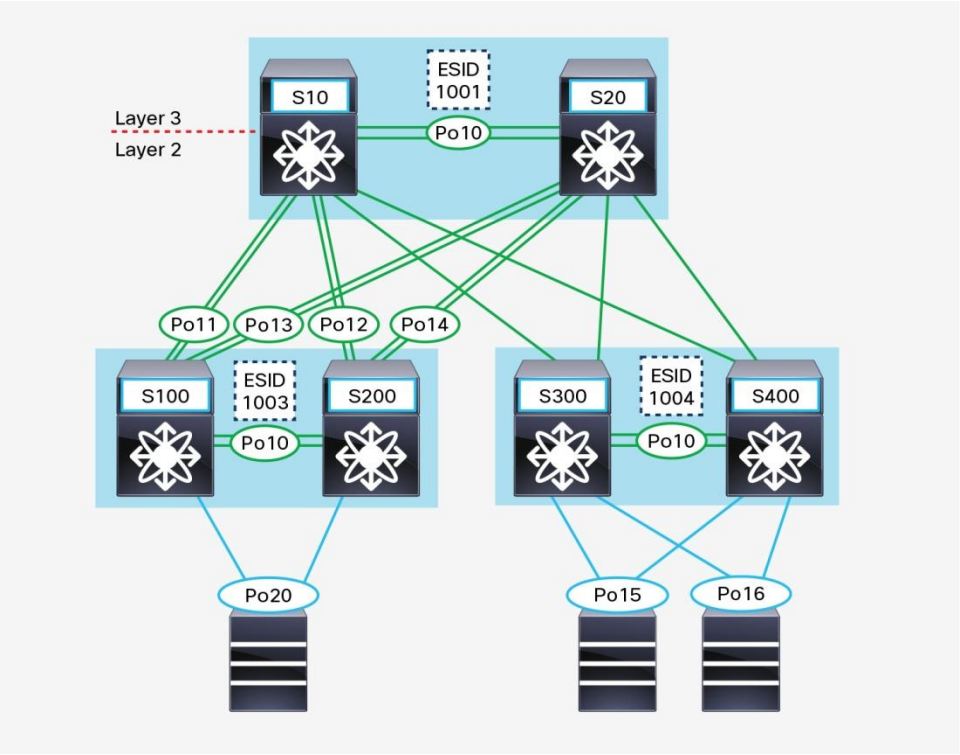
In this revised version of this document, you will find an additional set of best practices pertaining to the following topics:

- Multitopology feature
- Overload bit
- Anycast Hot Standby Router Protocol (HSRP)
- Protection against broadcast storms
- MAC address table scaling at the default gateway level
- Bidirectional forwarding detection (BFD) and Cisco FabricPath

Network Overview

The best practices are introduced using the network represented in Figure 1. This network topology shows the essential Cisco FabricPath features described in this document; however, this network design may not represent the preferred topology for scalability or for helping ensure that the solution continues to meet a particular organization's needs both now and in the future.

Figure 1. Network Topology



This network provides Layer 2 connectivity to all the leaf switches for various VLANs. It includes a set of common fabricwide VLANs and pod-specific VLANs. This separation is achieved with the help of the Cisco FabricPath multitopology feature, discussed later with the best practices. The leaf switches operate only as Layer 2 bridges.

The two Cisco Nexus® 7000 Series Switches at the core act as a Layer 2 and Layer 3 boundary; they allow routing between VLANs or to the outside of the network.

Virtual port channel (vPC+) technology is used both between the Cisco Nexus 7000 Series Switches at the core and between the pair of leaf switches.

Summary of Best Practices

The following best practices are discussed in this document:

- Manual switch ID assignment
- Multidestination trees root configuration and optional multitopology feature configuration
- Consistent VLAN configuration
- Use of port channels between leaf and spine nodes
- Fast convergence options and the overload bit
- Configuration of active-active default gateways with vPC+
- Attachment of devices other than Cisco FabricPath devices to the network fabric
- Cisco FabricPath and Spanning Tree Protocol connectivity options

- Anycast HSRP configuration
- Options for scaling the Layer 2 default gateway
- BFD and Cisco FabricPath

The document also provides additional notes. It concludes with full Cisco FabricPath network configurations.

Manual Switch ID Assignment

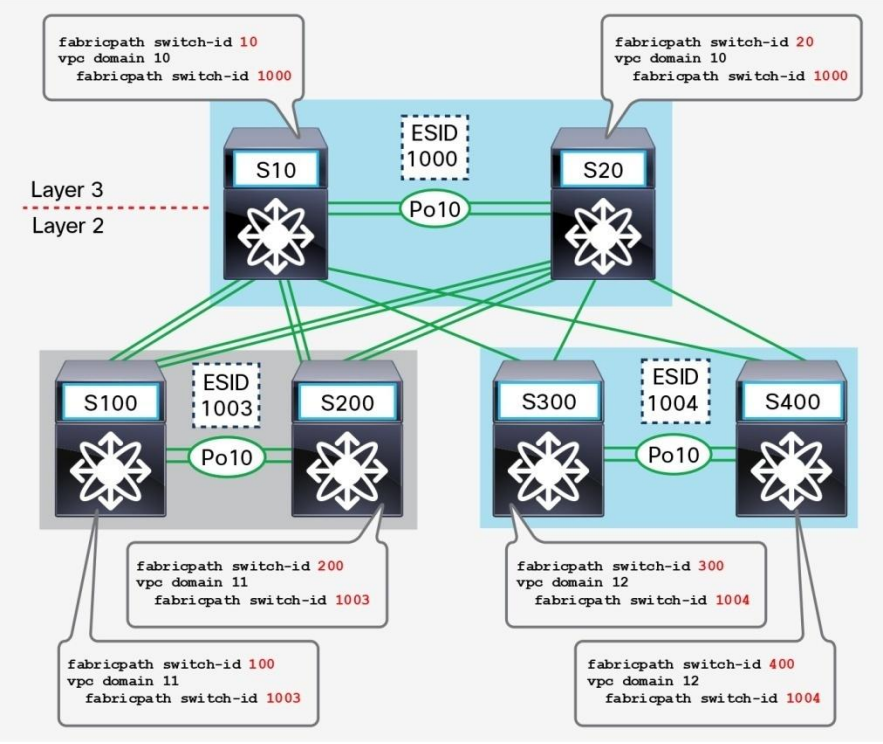
Cisco FabricPath can automatically assign switch IDs to all Cisco FabricPath nodes in the network; however, a meaningful numbering scheme is more convenient. During troubleshooting, a distinct numbering scheme allows faster and easier identification of switch roles.

Here is an example of such a scheme:

- The devices in the core (spine nodes) can be assigned two-digit IDs.
- The devices in the access layer (leaf nodes) can be assigned three-digit IDs.
- The virtual switches (emulated switches and anycast HSRP switches) can be assigned four-digit IDs.

The sample network discussed here used four-digit emulated switch IDs. However, in many real-world deployments, customers choose to match the numerical values of vPC+ domain IDs and their corresponding emulated switch IDs. Other customers prefer to align the vPC+ peer nodes with their associated emulated switch IDs to easily see the correspondence (for example, they use SIDs 101 and 102 for the peer nodes, and ESID 10 for the emulated switch). Figure 2 shows the assigned switch IDs.

Figure 2. Switch IDs



Here is a sample Cisco FabricPath switch ID configuration for a single vPC+ pair of Cisco Nexus switches. The remaining Cisco FabricPath switches are configured in a similar manner. The full configuration is provided in the appendix.

```
S10# show run fabricpath

feature-set fabricpath

fabricpath switch-id 10

vpc domain 10
  fabricpath switch-id 1000

-- output omitted --
```

```
S20# show run fabricpath

feature-set fabricpath

fabricpath switch-id 20

vpc domain 10
  fabricpath switch-id 1000

-- output omitted --
```

Multidestination Trees Root Configuration and Optional Multitopology Feature Configuration

The multitopology feature allows you to configure several Cisco FabricPath Intermediate System–to–Intermediate System (IS-IS) protocol topologies. Cisco NX-OS Software supports a variety of multidestination trees (MDTs) and topologies, depending on the platform used. For the latest information about the scalability parameters supported, refer to the Cisco Nexus 7000, 5000, and 6000 Series Switches verified scalability guides.

By default, only the default Topology 0 is configured on Cisco FabricPath nodes. Cisco FabricPath nodes cannot opt out of membership in Topology 0.

The Cisco FabricPath VLAN can belong to only a single topology; however Cisco FabricPath interfaces can be configured to carry multiple topologies (apart from the default Topology 0).

In the sample network, configured two additional topologies are configured:

- A pair of switches, S100 and S200, which constitute a first pod apart from the default Topology 0, is configured with Topology 1.

- A pair of switches, S300 and S400, which constitute a second pod apart from the default Topology 0, is configured with Topology 2.

A subset of VLANs assigned to Topology 0 is shared across these pods, whereas pod-local VLANs are assigned to their respective topologies. This logical separation helps contain multidestination traffic within the boundaries of that pod, or in other words, within that particular topology.

By default, Cisco FabricPath creates two MDTs in the default Topology 0, and multidestination traffic is mapped to either of those trees for load-balancing purposes. Cisco FabricPath IS-IS protocol elects the switch with the highest configured root priority as the root for MDT 1. The second-highest root priority becomes the root for MDT2. If the root priorities are the same, the switch with the highest system ID and then with the highest switch ID becomes the root.

As a best practice, you should explicitly set the root of all MDTs within each topology, so they provide optimal forwarding behavior. It is highly recommended that you configure a third Cisco FabricPath node with the third-highest priority to help ensure deterministic behavior. If the first two MDT root switches fail, the predetermined node becomes the new MDT root.

When configuring root priority, it is recommended that you not choose the highest configurable value, so that you have more flexibility in the event that the MDT root role needs to move to another Cisco FabricPath node.

Here is a sample configuration of Cisco FabricPath nodes S10, S100, and S300 showing the relevant commands for enabling the multitopology feature. Configuration for the rest of the Cisco FabricPath nodes can be found in the appendix. Note that no special configuration is needed for the default Topology 0.

```
S10# show run fabricpath

fabricpath topology 1
  member vlan 110-199
fabricpath topology 2
  member vlan 210-299
vlan 10-99,110-199,210-299
  mode fabricpath

!vPC+ Peer-link
interface port-channel10
  switchport mode fabricpath
  vpc peer-link
  fabricpath topology-member 1
  fabricpath topology-member 2

!Link to S100
interface port-channel11
  switchport mode fabricpath
  fabricpath topology-member 1
```

```
!Link to S200
interface port-channel12
  switchport mode fabricpath
  fabricpath topology-member 1


!Link to S400
interface Ethernet6/3
  fabricpath topology-member 2
  switchport mode fabricpath

!Link to S300
interface Ethernet6/4
  fabricpath topology-member 2
  switchport mode fabricpath

-- output omitted --
```

```
S100# show run fabricpath

fabricpath topology 1
  member vlan 110-199
vlan 10-99,110-199
  mode fabricpath

!vPC+ Peer-link
interface port-channel10
  switchport mode fabricpath
  vpc peer-link
  fabricpath topology-member 1

!Link to S10
interface port-channel11
  switchport mode fabricpath
  fabricpath topology-member 1

!Link to S20
interface port-channel13
```

```
  switchport mode fabricpath
  fabricpath topology-member 1

fabricpath domain default
    topology 1
    root-priority 245

-- output omitted --
```

```
S300# show run fabricpath

fabricpath topology 2
  member vlan 210-299
vlan 10-99,210-299
  mode fabricpath

!vPC+ Peer-link
interface port-channel10
  switchport mode fabricpath
  vpc peer-link
  fabricpath topology-member 2

!Link to S10
interface Ethernet5/1
  fabricpath topology-member 2
  switchport mode fabricpath

!Link to S20
interface Ethernet5/2
  fabricpath topology-member 2
  switchport mode fabricpath

fabricpath domain default
  topology 2
    root-priority 245

-- output omitted --
```

The following figures illustrate the resulting MDTs: (please note that these are examples to illustrate how MDT tree looks like. The MDT tree root selection may not be optimal.)

- Figure 3 shows a topology membership within a Cisco FabricPath network. The presence of a specific color label indicates whether or not that topology is defined on the switch.
- Figures 4 and 5 show MDTs for the default Topology 0.
- Figures 6 and 7 show MDTs for Topology 1.
- Figures 8 and 9 show MDTs for Topology 2.

Figure 3. Red, Green, and Blue Labels Signify Configured Topology on That Cisco FabricPath Node

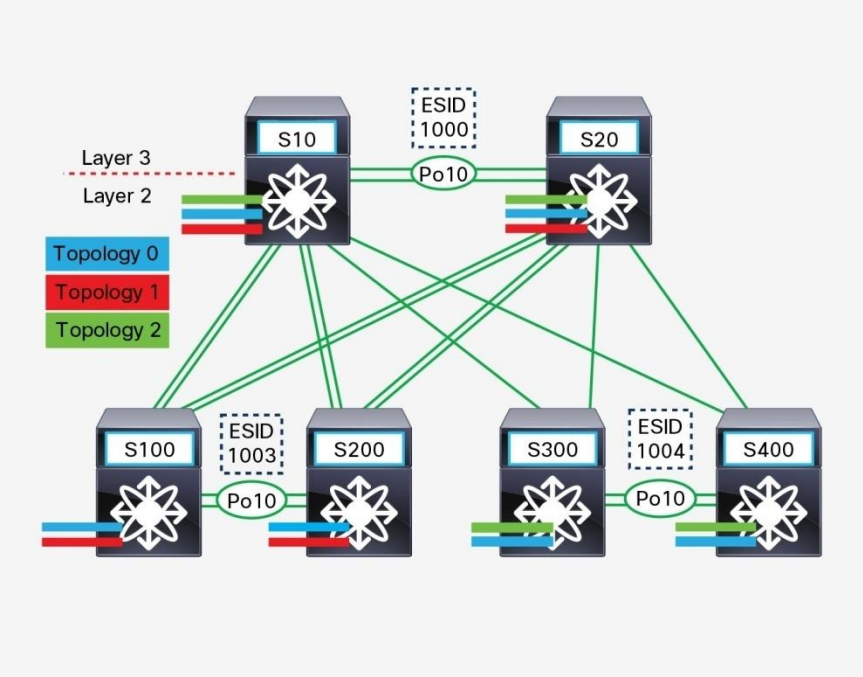


Figure 4. MDT 1 for **Topology 0** Is Rooted at S10

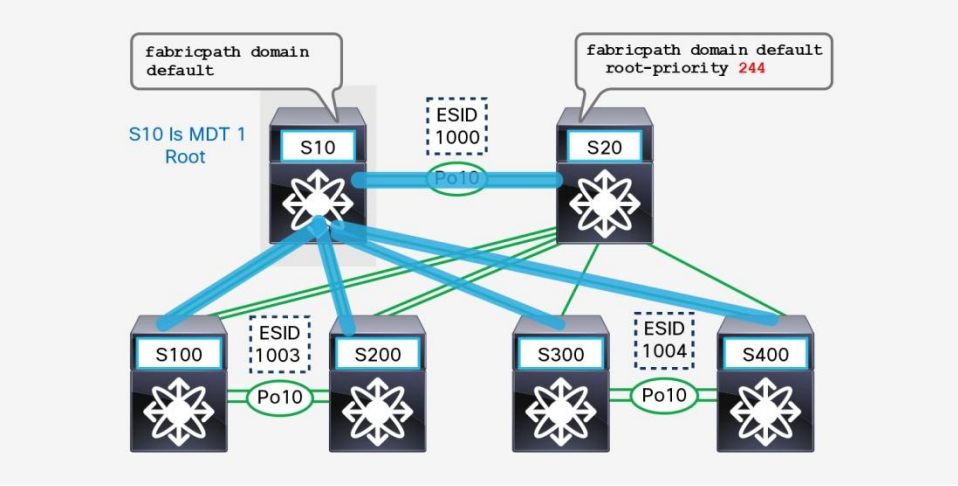


Figure 5. MDT 2 for **Topology 0** Is Rooted at S20

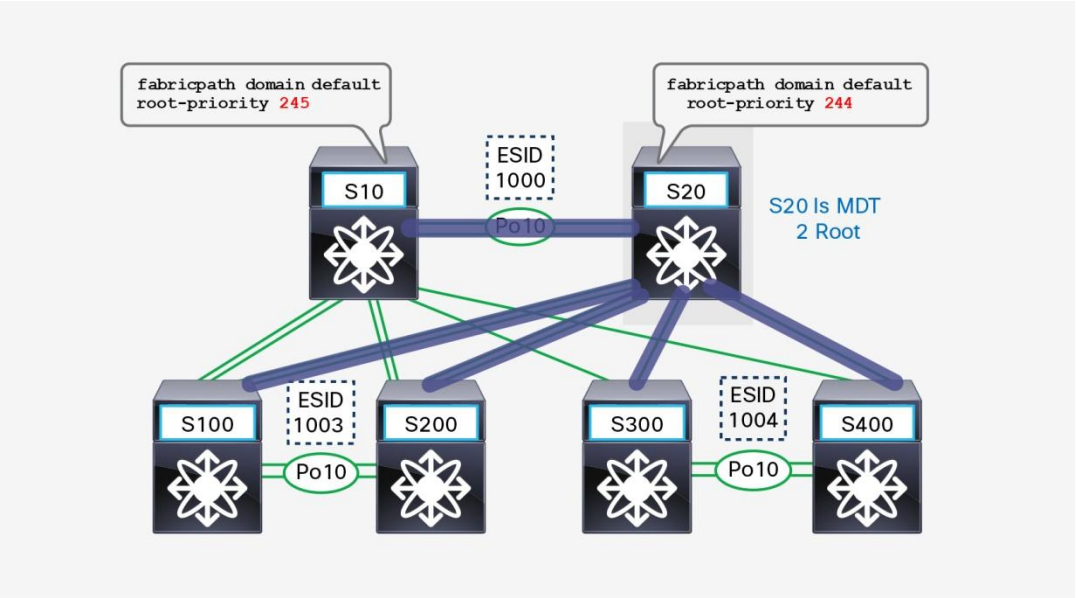


Figure 6. Only S10, S20, S100, and S200 Are Part of **Topology 1**—MDT 1 Root for **Topology 1** Is S100

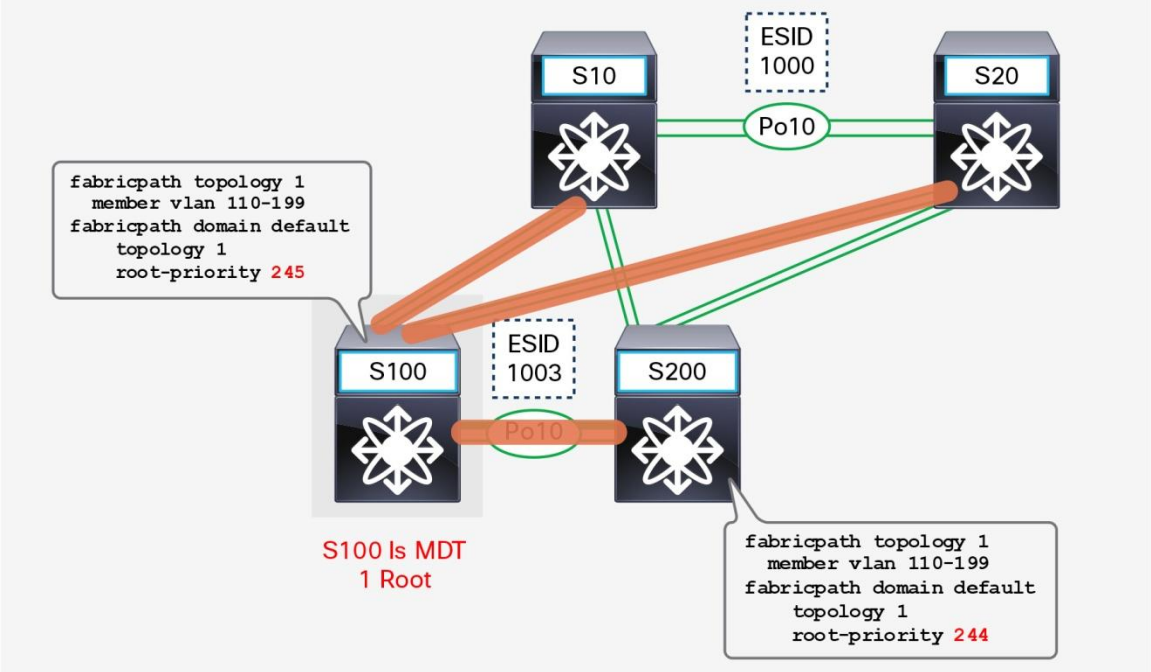


Figure 7. Only S10, S20, S100, and S200 Are Part of **Topology 1**—MDT 2 Root for **Topology 1** Is S200

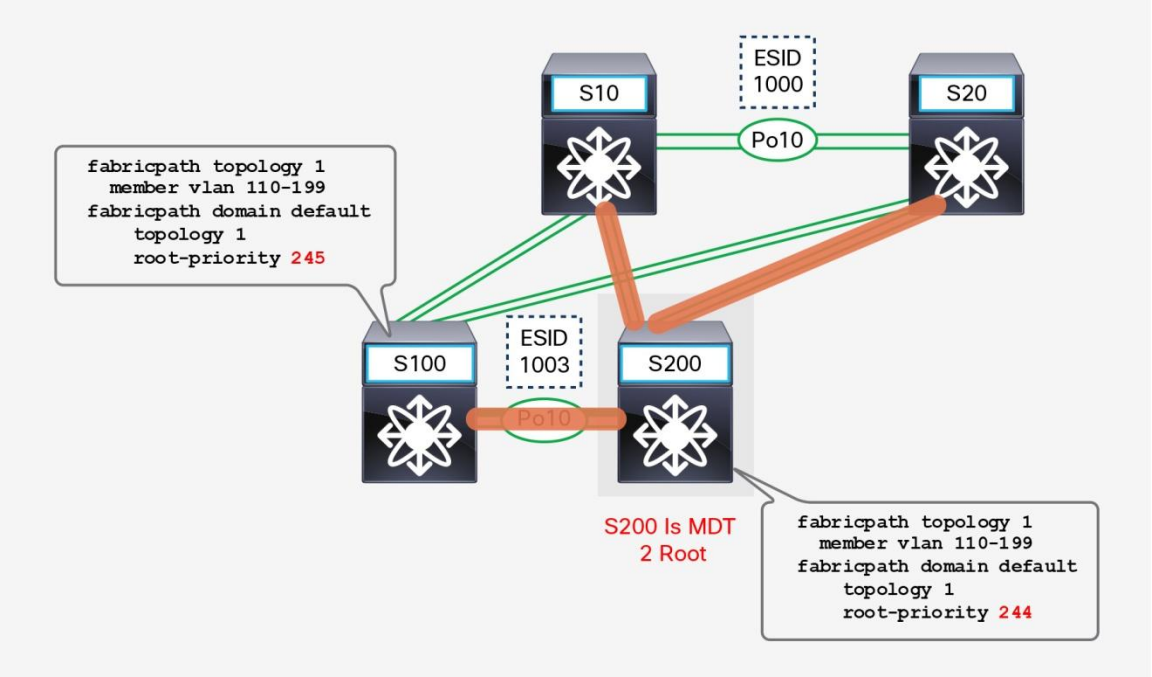


Figure 8. Only S10, S20, S300, and S400 Are Part of **Topology 2**—MDT 1 Root for **Topology 2** Is S300

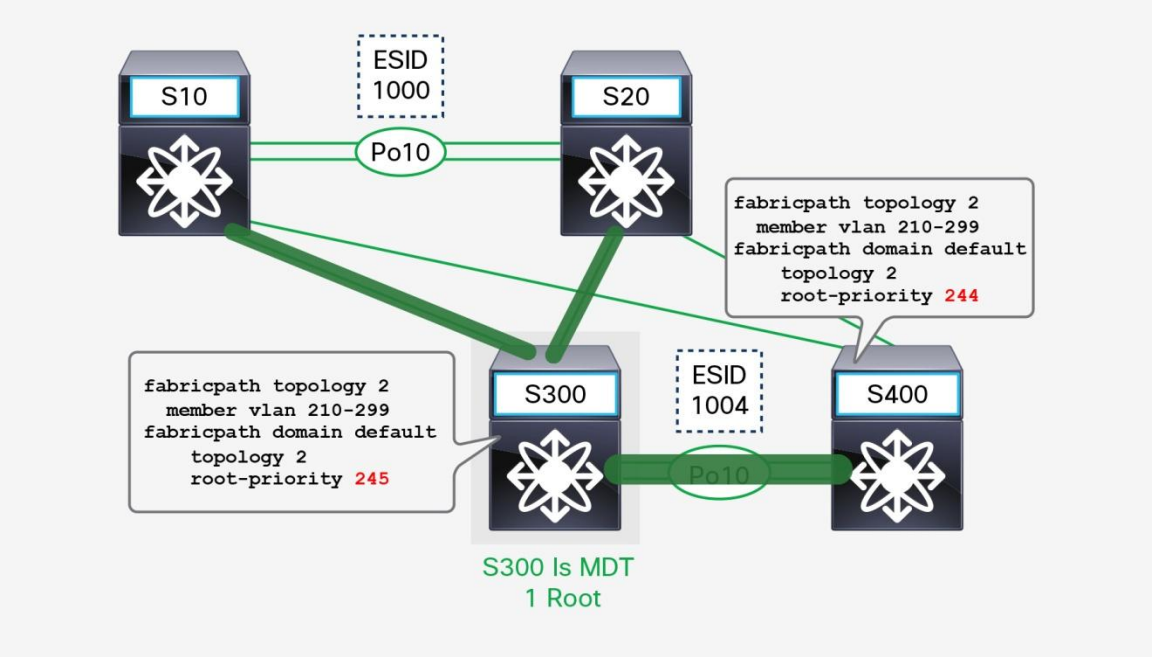
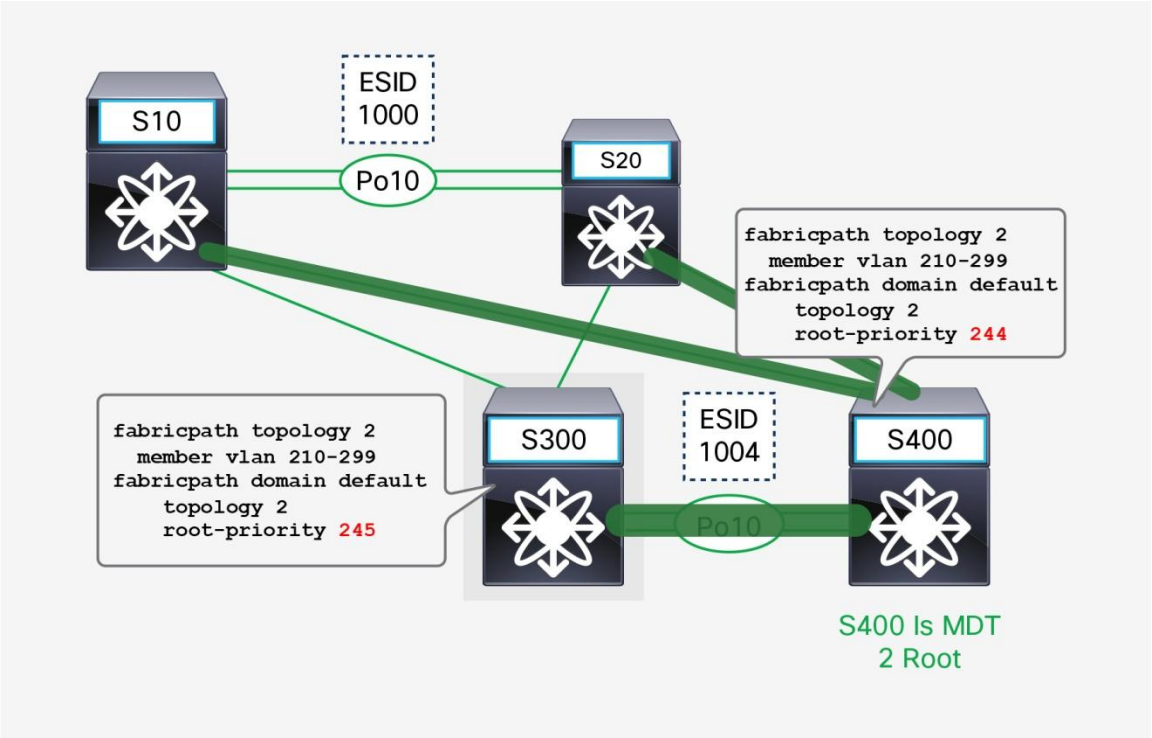


Figure 9. Only S10, S20, S300, and S400 Are Part of **Topology 2**—MDT 2 Root for **Topology 2** Is S400



Consistent VLAN Configuration

You should configure the Cisco FabricPath VLANs consistently on all Cisco FabricPath switches in a particular topology. The sample network has three topologies, and all **mode fabricpath** VLANs are assigned according to Table 1.

Table 1. VLAN Assignment

VLAN IDs	Topology ID	Member Switches
10 to 99	0	S10*, S20**, S100, S200, S300, and S400
110 to 199	1	S10, S20, S100*, and S200**
210 to 299	2	S10, S20, S300*, and S400**

* The Cisco FabricPath switch is configured as the root for MDT 1 in the respective topology.
** The Cisco FabricPath switch is configured as the root for MDT 2 in the respective topology.

All topology-specific VLANs must be configured on all Cisco FabricPath nodes in their respective topologies.

Note that when VLAN 1 is used to forward data, it should also be configured as **mode fabricpath** and placed in the default Topology 0. In the sample network, VLAN 1 is not used to forward data.

For example, VLAN 110 must be configured on switches S10, S20, S100, and S200. However, this same VLAN doesn't need to be configured on switches S300 and S400 because they participate in Topology 0 and Topology 2 only.

This recommendation is applicable only to **mode fabricpath** VLANs, not to **mode CE** VLANs.

Note that when a new **mode fabricpath** VLAN is configured, it is automatically assigned to the default Topology 0. If you need to a newly created VLAN to belong to specific topology from the start, you should assign this VLAN to the desired topology prior to configuring it with the **mode fabricpath** command.

Here are sample configurations for all six Cisco FabricPath nodes:

```
S10# show run vlan

vlan 1,10-99,110-199,210-299
vlan 10-99,110-199,210-299
  mode fabricpath
```

```
S20# show run vlan

vlan 1,10-99,110-199,210-299
vlan 10-99,110-199,210-299
  mode fabricpath
```

```
S100# show run vlan

vlan 1,10-99,110-199
vlan 10-99,110-199
  mode fabricpath
```

```
S200# show run vlan

vlan 1,10-99,110-199
vlan 10-99,110-199
  mode fabricpath
```

```
S300# show run vlan

vlan 1,10-99,210-299
vlan 10-99,210-299
  mode fabricpath
```

```
S400# show run vlan

vlan 1,10-99,210-299
vlan 10-99,210-299
  mode fabricpath
```

Use of Port Channels between Leaf and Spine Nodes

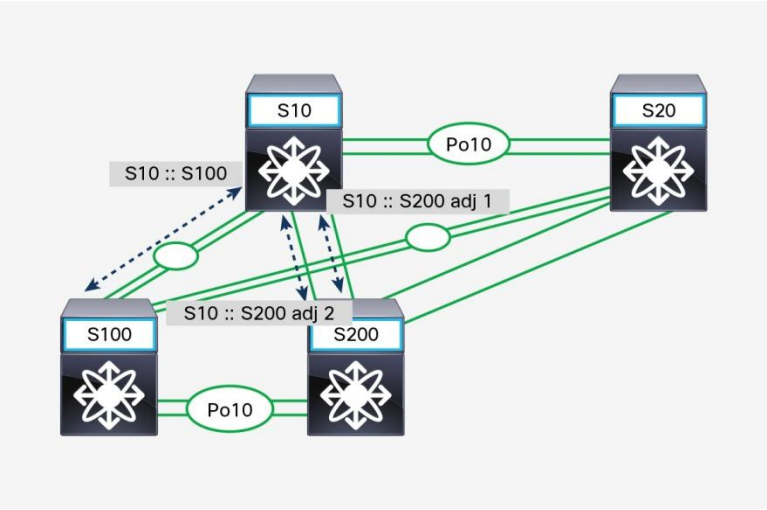
You should group the links between adjacent switches in a port channel. There are two main benefits of bundling links between two adjacent Cisco FabricPath nodes:

- Port channels provide more bandwidth for the MDTs because a port channel is treated as a single logical link in Cisco FabricPath IS-IS protocol.
- With port channels configured, the Cisco FabricPath IS-IS protocol has to handle fewer adjacencies than with individual Cisco FabricPath links and therefore scales better.

Two options are available to set the Cisco FabricPath IS-IS metrics and costs on port channels. By default, Cisco NX-OS adjusts the metrics and costs based on the number of active port-channel members and uses only the lowest metrics for each path. A change in the number of port-channel members always causes a Cisco FabricPath IS-IS recalculation. Statically hard-coded metrics on port channels provide reliable paths independent of the number of available active members of port channels and prevents Cisco FabricPath IS-IS recalculation; however, the remaining individual member links might be oversubscribed in the event of port-channel member link failure. In other words, port channels that are hard coded with specific metrics may be burdened with more traffic than the number of active member links can handle.

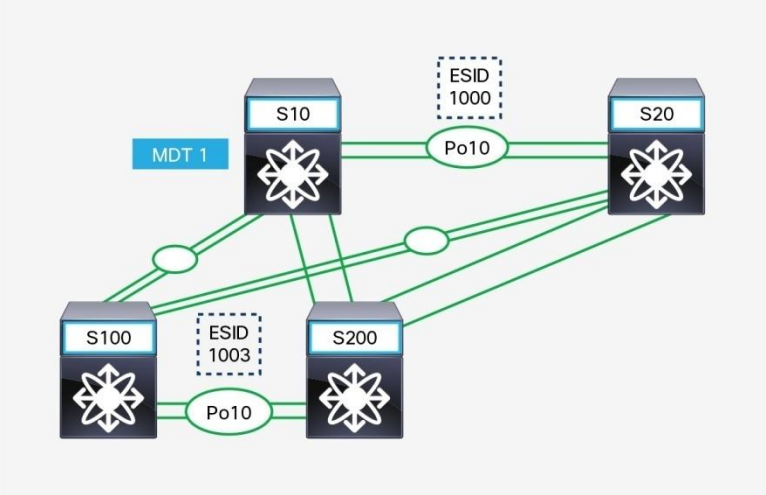
In Figure 10, Cisco FabricPath IS-IS protocol sees only a single logical link and a single adjacency between switches S10 and S100. However, Cisco FabricPath IS-IS protocol has to handle two links and two adjacencies between switches S10 and S200.

Figure 10. Each Individual Logical Link Requires Neighbor Adjacency



In Figure 11, Cisco FabricPath IS-IS protocol can take advantage of two links bundled into a port channel between switches S10 and S100 when building an MDT, gaining more bandwidth for multidestination traffic, such as multicast. However, Cisco FabricPath IS-IS protocol has to use only a single link for MDT traffic between switches S10 and S200.

Figure 11. Multidestination Tree (MDT) Distribution



Fast Convergence Options and the Overload Bit

Cisco FabricPath uses IS-IS protocol in its control-plane infrastructure.

Similar to IS-IS protocol for IPv4 or IPv6, IS-IS protocol for Cisco FabricPath is a link-state protocol, and in many convergence cases it uses event-based logic. For example, in addition to timer-based regular IS-IS Hello (IIH) frames, it sends a link-state protocol-data-unit (LSP) in the event of a link failure. The LSP enables fast convergence without the need to wait for the next IS-IS Hello frame.

As with other types of routing protocols, Cisco NX-OS allows adjustment of various timers for Cisco FabricPath (for example, **lsp-gen-interval** and **spf-interval**). However, the recommended approach is to leave the timers at their default values and maintain the built-in exponential timer behavior in the event of a link or node failure. The default Cisco FabricPath IS-IS exponential LSP timers and shortest-path first (SPF) timers respond within a few milliseconds to any network topology changes. In times of network instability (such as that caused by a flapping link), however, the timers increase to throttle the rate of response to network events. This scheme helps ensure fast convergence when the network is stable, and it helps moderate CPU cycles when the network is unstable.

Proceed with caution when lowering these timers because in some cases ultra-short timer values increase overall CPU utilization.

Several features are available that can help with convergence. One such feature is the Cisco FabricPath linkup delay timer. The linkup delay timer starts when the Cisco FabricPath core link is brought up and the Cisco FabricPath adjacencies is built. During linkup delay, each Cisco FabricPath switch receives FabricPath IS-IS updates from all other Cisco FabricPath switches in the network and detects any switch ID conflicts, but the FabricPath port is in suspended state for data packet. After the timer expires, the Cisco FabricPath switch resolves any switch ID conflicts and brings up the FabricPath port to be data forwarding state. You can configure the timer with a value greater than the default value (the default value is 10 seconds), as shown here, to give the Cisco

FabricPath switch more time to detect switch ID conflicts and to exchange enough neighbor information for the FabricPath IS-IS shortest path calculations, leading to optimized convergence. It is recommended that the link up delay is set consistently in the Cisco FabricPath fabric.

```
S100# show run fabricpath

fabricpath timers linkup-delay 60
```

Another such feature is the overload bit of the IS-IS protocol. The Cisco FabricPath IS-IS overload-bit allows you to exclude the advertising node from use as a transit route for best-route calculations. That is, whenever a Cisco FabricPath node advertises the overload bit in a specific Cisco FabricPath IS-IS type-length-value (TLV) specification, the remaining systems will route traffic around it. The primary goal of the overload bit is to eliminate traffic disruption in Cisco FabricPath node restoration scenarios. After the overload bit is turned off in the IS-IS LSP messages, adjacent Cisco FabricPath nodes will reinstate that node as a possible transit route for user data traffic. The information that follows needs to be taken into consideration when deploying the overload bit.

When a Cisco FabricPath node needs to be taken offline, issue the **set-overload-bit always** command. This command helps ensure that the overload-bit is always advertised until the **no** option for this command is issued.

```
S100# show run fabricpath

fabricpath domain default
  set-overload-bit always
```

To help ensure that Cisco FabricPath node has enough time to complete all relevant Cisco FabricPath IS-IS tasks during a switch restoration procedure, the network administrator can issue the **set-overload-bit on-startup X** command and specify amount of time that a Cisco FabricPath node advertises the overload bit. After timer elapses, the Cisco FabricPath node is smoothly reinserted into the Cisco FabricPath network.

You can check whether the Cisco FabricPath IS-IS protocol is advertising LSPs with the overload bit with the command shown here. In this case, switch S10 is configured with **set-overload-bit always**.

```
S100# sh fabricpath isis database
Fabricpath IS-IS domain: default LSP database
  LSPID                 Seq Number   Checksum  Lifetime   A/P/O/T
  d867.d903.f342.00-00  0x00001F30   0x6B67    1197       0/0/1/1
  d867.d903.f343.00-00  0x00001EED   0xE8D0    1197       0/0/0/1
  d867.d903.f344.00-00* 0x00001F2F   0xAB03    895        0/0/0/1
  d867.d903.f345.00-00  0x00001F02   0x0531    1133       0/0/0/1
  d867.d90a.0b42.00-00  0x00001F36   0x0EA5    1197       0/0/0/1
  d867.d90a.0b43.00-00  0x00001073   0xE983    1197       0/0/0/1
  d867.d90a.0b44.00-00  0x00001F3A   0x3CC6    748        0/0/0/1
  d867.d90a.0b45.00-00  0x00001EFC   0xC82E    1026       0/0/0/1
```

```
S100# sh fabricpath isis hostname
Fabricpath IS-IS domain: default dynamic hostname table
  Level  System ID       Dynamic hostname
  1      d867.d903.f342  S10
  1      d867.d903.f343  S30
  1      d867.d903.f344* S100
  1      d867.d903.f345  S300
  1      d867.d90a.0b42  S20
  1      d867.d90a.0b43  S40
  1      d867.d90a.0b44  S200
  1      d867.d90a.0b45  S400
```

The value of the on-startup timer depends on multiple factors, such as the size of the Cisco FabricPath network and delayed activation of hardware-specific line cards. The ultimate goal is for the overload-bit on-startup timer not to elapse before all relevant Cisco FabricPath IS-IS tasks are completed; otherwise, Cisco FabricPath IS-IS recomputation may occur.

Note: The Cisco FabricPath overload-bit feature is designed for transit nodes only; future overload-bit enhancements will apply to additional deployment options. Also be aware that Cisco NX-OS on every Cisco FabricPath node needs to understand the overload-bit IS-IS TLV specifications. Refer to configuration guide or release notes for more details.

```
S100# show run fabricpath

fabricpath domain default
  set-overload-bit on-startup <x>
```

The configured timer value should be evaluated based on a platform in use and the number of adjacent neighbors. The recommended starting value is approximately 300 to 600 seconds.

Refer to the Cisco Nexus 7000 Series FabricPath white paper and configuration guide for more details.

Note: For the overload bit to work correctly, all nodes in the Cisco FabricPath network should use the supported Cisco NX-OS code. Check the release notes for the latest details.

Note: The overload bit doesn't work with a vPC+ switch. When the overload bit is set on a vPC+ switch that is in the transit path, the vPC+ switch continues attracting traffic.

Note: On the Cisco Nexus 7000 Series Switches, the overload-bit function with Anycast HSRP is supported starting with Cisco NX-OS 6.2(8). Check the release notes for the latest details.

Configuration of Active-Active Default Gateways with vPC+

Configure vPC+ in conjunction with HSRP to use both the vPC+ peers as active-active default gateways.

All devices, whether they are connected through vPC+ or native Cisco FabricPath core ports, can now use multiple paths to two active gateways for the routed traffic (south-north and inter-VLAN traffic).

The following is a partial configuration on switches S10 and S20 that enables active-active default gateway capabilities for specified VLANs. Refer to Figure 1 for the network topology.

```
S10# sh run vpc

vpc domain 10
  peer-keepalive destination 1.1.1.2 source 1.1.1.1 vrf vpc
  fabricpath switch-id 1000

interface port-channel10
  vpc peer-link

S10# sh run hsrp

interface Vlan10
  ip address 10.10.0.254/24
  hsrp version 2
  hsrp 10
    ip 10.10.0.1

interface Vlan110
  ip address 10.110.0.254/24
  hsrp version 2
  hsrp 110
    ip 10.110.0.1
```

```
S20# sh run vpc

vpc domain 10
  peer-keepalive destination 1.1.1.1 source 1.1.1.2 vrf vpc
  fabricpath switch-id 1000

interface port-channel10
  vpc peer-link
```

```
S20# sh run hsrp

interface Vlan10
  ip address 10.10.0.253/24
  hsrp version 2
  hsrp 10
    ip 10.10.0.1

interface Vlan110
  ip address 10.110.0.253/24
  hsrp version 2
  hsrp 110
    ip 10.110.0.1
```

Attachment of Devices Other Than Cisco FabricPath Devices to the Network Fabric

Devices other than Cisco FabricPath devices (servers, switches, and services appliances) can be dual-attached to vPC+ switches using IEEE standard port channels without the need to use Spanning Tree Protocol to provide redundancy. VLANs carried on vPC+ member ports must be Cisco FabricPath mode VLANs.

The vPC peer switch was first introduced as a vPC feature to help ensure that both vPC peer devices present themselves as a single unique Spanning Tree Protocol root device using the same bridge ID to send the exact same Bridge Protocol Data Unit (BPDU). The design of Cisco FabricPath requires the entire FabricPath domain to appear as a single Spanning Tree Protocol bridge to any connected Spanning Tree Protocol domains, so all FabricPath bridges share a common bridge ID (C84C.75FA.6000). According to the design, both vPC+ switches should send the exact same BPDU, so a peer-switch command is not needed for a vPC+ environment by default. However, a known issue for Cisco Nexus 6000 and 5000 Series Switches requires a peer-switch command to force both vPC+ switches send the same BPDU.

Here are sample port-channel interface configurations for vPC+ peers. Refer to Figure 1 for the network topology.

```
S100# sh run spanning-tree
spanning-tree domain 10
spanning-tree pseudo-information
  vlan 10-99,110-199 root priority 4096

S100# sh run int e6/36
interface Ethernet6/36
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 10-11
  channel-group 20 mode active
  no shutdown
```

```
S200# sh run spanning-tree
spanning-tree domain 10
spanning-tree pseudo-information
  vlan 10-99,110-199 root priority 4096

S200# sh run int e6/36
interface Ethernet6/36
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 10-11
  channel-group 20 mode active
  no shutdown
```

```
S100# sh run int po 20

interface port-channel20
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 10-11
  vpc 20
```

```
S200# sh run int po 20

interface port-channel20
  switchport
  switchport mode trunk
  switchport trunk allowed vlan 10-11
  vpc 20
```

Cisco FabricPath and Spanning Tree Protocol Connectivity Options

Cisco FabricPath supports not only direct Classic Ethernet host connections, but also connection of traditional spanning-tree switches attached to edge ports. Each Cisco FabricPath edge switch must be configured as the root for all FabricPath VLANs. If you are connecting spanning-tree devices to the Cisco FabricPath fabric, make sure that you configure all edge switches as the spanning-tree root. Note in the following code that the **spanning-tree priority** command would work:

```
S100(config)# spanning-tree vlan x root primary
Or S100(config)# spanning-tree vlan x priority 8192
```

However, this configuration would change the priority for the spanning tree depending on whether the switch is sending regular BPDUs (when Cisco FabricPath is not running) or sending BPDUs when Cisco FabricPath is operational on the switch. In some scenarios, this change can have undesirable side effects. (For a detailed explanation, see Appendix B.)

As a best practice, on all Cisco FabricPath switches that have Classic Ethernet (CE) ports connected to Classic Ethernet switches, configure the same root priority using the **spanning-tree pseudo information** command shown here:

```
S100(config)# spanning-tree pseudo-information
S100(config-pseudo)#  vlan 10-100 root priority 4096
```

Make sure that the configured priority is the best (lowest) in the network, so that the Cisco FabricPath region is the root of the spanning-tree. If the Classic Ethernet edge ports receive a superior spanning-tree Bridge Protocol Data Unit (BPDU), those ports will be blocked from forwarding traffic.

When a traditional spanning-tree domain is connected to a Cisco FabricPath network, the whole Cisco FabricPath network will be perceived as a single spanning-tree switch, which, however, does not pass spanning-tree BPDUs by default.

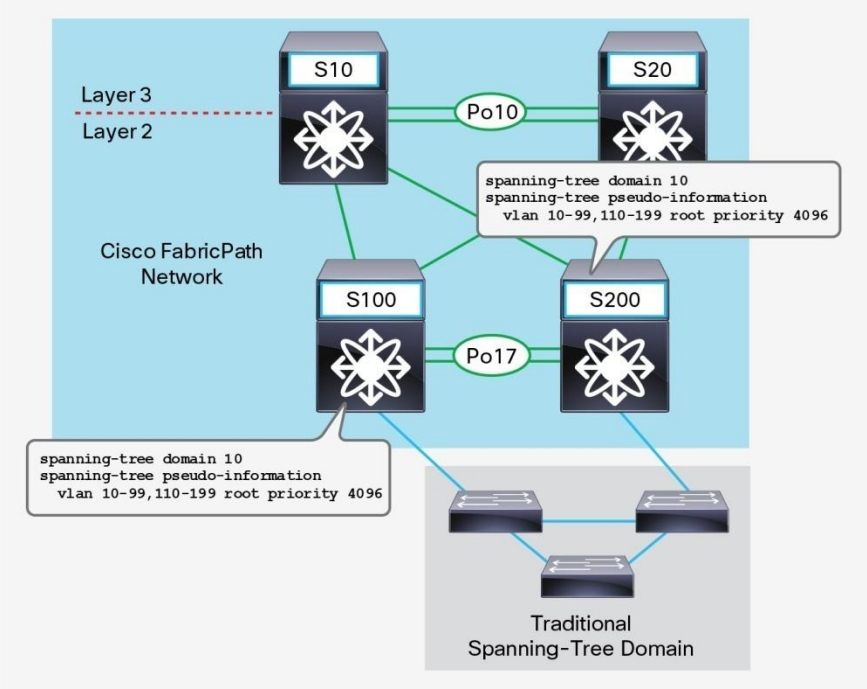
To provide correct and consistent end-to-end MAC address learning, the Cisco FabricPath IS-IS could transport topology change notification (TCN) information to the Cisco FabricPath edge switches and initiate a flush of the MAC address table. To achieve this behavior, additional steps are required:

- Identify all Cisco FabricPath edge ports connected to the same external spanning-tree domain.
- Configure all identified Cisco FabricPath edge ports with the exact same spanning-tree domain ID using the command **spanning-tree domain <x>**.

Note that a given Cisco FabricPath leaf can be configured with only a single spanning tree domain ID.

These steps help ensure a loop-free environment by forwarding relevant BPDUs between all Cisco FabricPath edge ports configured with the identical spanning-tree domain ID (Figure 12).

Figure 12. Spanning-Tree Configuration on Leaf Nodes Connecting to External Spanning-Tree Domain



Anycast HSRP Configuration

Users can configure Anycast HSRP to provide the default gateway function for more than two devices (up to four default gateways on Cisco Nexus 5000, 6000, and 7000 Series Switches.

The Anycast HSRP configuration consists of several components:

- Switch virtual interface (SVI), where the HSRP virtual address is configured
- Anycast HSRP bundle, where the following subcomponents are configured:
  - Anycast bundle ID
  - Anycast switch ID
  - Anycast bundle priority (used for election of active, standby, and listen roles)
  - List of VLANs for which anycast HSRP will be provided

To help ensure that the Anycast HSRP bundle is functioning, the following conditions must be met:

- Anycast switch ID must be configured
- All SVIs of all VLANs defined for the bundle must be configured with HSRP Version 2

If these conditions are not met, either you will not be able to bring up a bundle or a partial traffic drop will occur. The condition in which traffic is partially dropped may occur during runtime if the SVI for the VLAN in the Anycast HSRP bundle goes down as a result of VLAN definition removal, SVI shutdown, or shutdown of all available ports in that VLAN. Cisco Nexus 7000 Series Switches with Cisco NX-OS 6.2(10) and later and Cisco Nexus 5000 and 6000 Series Switches with Cisco NX-OS 7.1(0)N1(1a) and later introduce a new command: **force gateway-down**. This command causes Anycast bundle reconvergence and provides smooth operation if a failure trigger occurs.

The network administrator should help ensure that all Anycast HSRP routers establish Layer 3 routing adjacency with the Interior Gateway Protocol (IGP) of choice. This adjacency is necessary to help ensure consistent routing behavior across all Anycast HSRP routers. As a best practice, configure a transit VLAN and establish Layer 3 routing adjacency among the Anycast HSRP routers using the transit VLAN. Server VLAN SVIs are set as IGP passive interfaces and do not have routing peering adjacency.

The best practices for configuring anycast HSRP are presented here.

1. Create the SVI on all four nodes.

```
SXX# sh run int vlan 10
interface Vlan10
  no shutdown
  ip address 10.10.0.25X/24
  ip ospf passive-interface
  ip router ospf 1 area 0.0.0.0
  hsrp version 2
  hsrp 10
    ip 10.10.0.1
```

2. Verify that the SVI is up on all nodes.

```
S10# show int vlan 10
Vlan4 is up, line protocol is up, autostate enabled
```

3. At this point, VLAN 10 will work as a standard HSRP VLAN and should be in the active-standby listen-listen state. Ignore the priority because this value will be set after the VLANs are mapped to the anycast group.

```
S10# sh hsrp interface vlan 10 brief
*:IPv6 group   #:group belongs to a bundle
                     P indicates configured to preempt.
                     |
 Interface   Grp  Prio P State    Active addr      Standby addr     Group addr
  Vlan10      10   120    Active  local      10.10.0.254            10.10.0.1

S20# sh hsrp interface vlan 10 brief
*:IPv6 group   #:group belongs to a bundle
                     P indicates configured to preempt.
```

```
                     |
 Interface   Grp  Prio P State    Active addr      Standby addr     Group addr
  Vlan10      10   115    Standby   10.10.0.251      local      10.10.0.1

S30# sh hsrp interface vlan 10 brief
*:IPv6 group   #:group belongs to a bundle
                     P indicates configured to preempt.
                     |
 Interface   Grp  Prio P State    Active addr      Standby addr     Group addr
  Vlan10      10   100    Listen   10.10.0.251      10.10.0.254      10.10.0.1

S40# sh hsrp interface vlan 10 brief
*:IPv6 group   #:group belongs to a bundle
                     P indicates configured to preempt.
                     |
 Interface   Grp  Prio P State    Active addr      Standby addr     Group addr
  Vlan10      10   100    Listen   10.10.0.251           10.10.0.254
10.10.0.1
```

4. Add the VLANs in the bundle, configure the Anycast switch ID, and configure the priority on all the switches. Start with the switch with the highest priority because doing so will help ensure a stable Anycast HSRP active role from the start and prevent repeated role reelection.

```
S-XX(config)# hsrp anycast 100 ipv4
S-XX(config-anycast-bundle)# switch-id 1100
S-XX(config-anycast-bundle)# vlan 10-11
S-XX(config-anycast-bundle)# priority X <0-127>
S-XX(config-anycast-bundle)# no shutdown
```

In addition, to remove VLANs from the Anycast bundle, follow these steps:

1. Remove the VLAN from the bundle on all switches starting with the one with the lowest priority to avoid HSRP active-standby state reconvergence.
2. Delete the SVI.

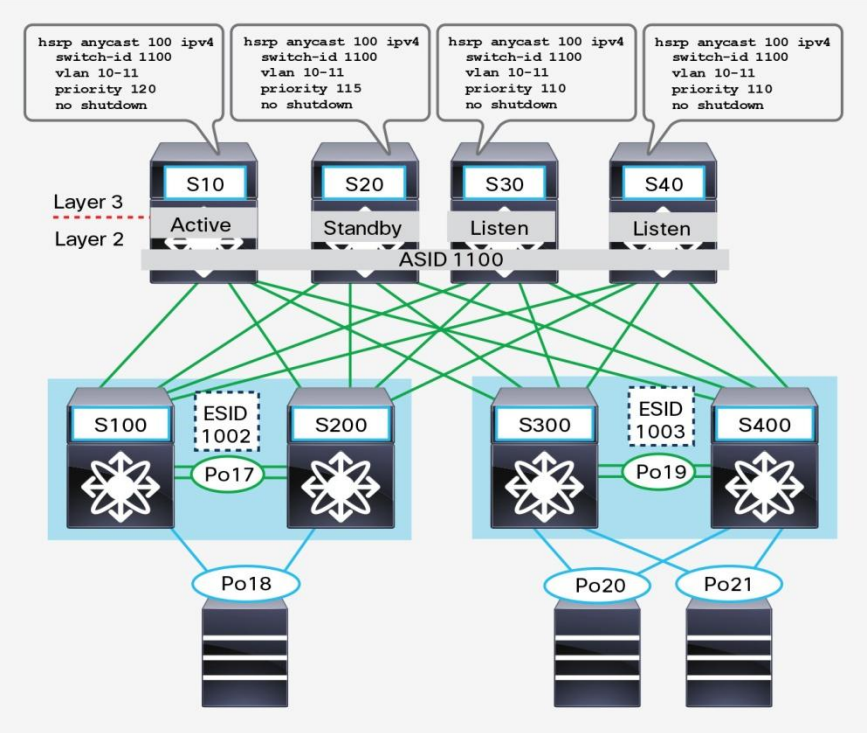
```
S-XX(config)# hsrp anycast 100 ipv4
S-XX(config-anycast-bundle)# no vlan 10
```

Note that all Cisco FabricPath nodes in a network must use the following code software releases or later to support Anycast HSRP:

- Cisco Nexus 7000 Series: Cisco NX-OS 6.2(2) (Anycast HSRP gateway support), minimum recommended release is Cisco NX-OS 6.2(10)
- Cisco Nexus 5500 platform: Cisco NX-OS 6.0(2)N2(1) (Anycast HSRP leaf support only); minimum recommended release is Cisco NX-OS 7.1(0)N1(1a) (Anycast gateway support)
- Cisco Nexus 6000 Series: Cisco NX-OS 6.0(2)N2(1) (Anycast HSRP leaf support only); minimum recommended release is Cisco NX-OS 7.1(0)N1(1a) (Anycast gateway support)

Figure 13 shows a sample Cisco FabricPath network with the Anycast HSRP function and relevant configuration information.

Figure 13. Cisco FabricPath Network with Anycast HSRP and Configuration Information



You can read more about Cisco FabricPath the Anycast HSRP function in the Cisco FabricPath Interfaces configuration guide and Cisco Nexus 7000 Series FabricPath white paper.

Options for Scaling the Layer 2 Default Gateway

Note that this section of this best-practices document applies only to Cisco Nexus 7000 Series high-density environments in which the total number of hosted virtual machines and hosts exceeds 16,000 MAC addresses. If your projected capacity is lower, then you can skip this section.

Network designers commonly need to know the number devices and hosts, or in other words, the number of MAC addresses, that will be supported on the given Layer 2 network.

Despite Cisco FabricPath's conversational learning, which helps increase the effective MAC address scalability at the leaf-node level, border leaf nodes with the default gateway function have to learn MAC addresses for all hosts and virtual machines that need to leave VLAN boundaries. In other words, the MAC addresses of hosts that send traffic to their default gateways have to be learned and cached on border leaf nodes.

Therefore, in certain scenarios, high-density environments may push scalability requirements beyond the capacity of a single pair of border leaf nodes. In many cases, border leaf nodes are implemented on a pair of Cisco Nexus 7000 Series Switches with F2e-Series enhanced modules, and such nodes can provide up to 16,000 MAC addresses.

The recommended options can help you scale high-density environments beyond the 16,000 limit on MAC addresses.

Option 1

Distribute terminated VLANs across several border leaf node pairs or between several virtual device contexts (VDCs). This approach helps ensure that only MAC addresses from a chosen subset of configured VLANs will have to be learned. This step may involve additional cabling.

With F2-Series and F2e-Series modules, the resulting scale can reach multiples of 16,000 MAC addresses total. With F3-Series modules, the resulting scale can reach multiples of 64,000 MAC addresses total.

In the sample network, several VLANs are configured, and they contain the following numbers of MAC addresses:

- VLANs 10 to 14: 7500 MAC addresses (group A1)
- VLANs 15 to 19: 8000 MAC addresses (group A2)
- VLANs 20 to 24: 4000 MAC addresses (group A3)
- VLANs 25 to 29: 3500 MAC addresses (group A4)
- Cumulative number of MAC addresses: 23,000

This total is more than a single pair of Cisco Nexus 7000 Series Switches with F2e-Series modules can accommodate. However, suppose that you split these VLAN ranges into two groups:

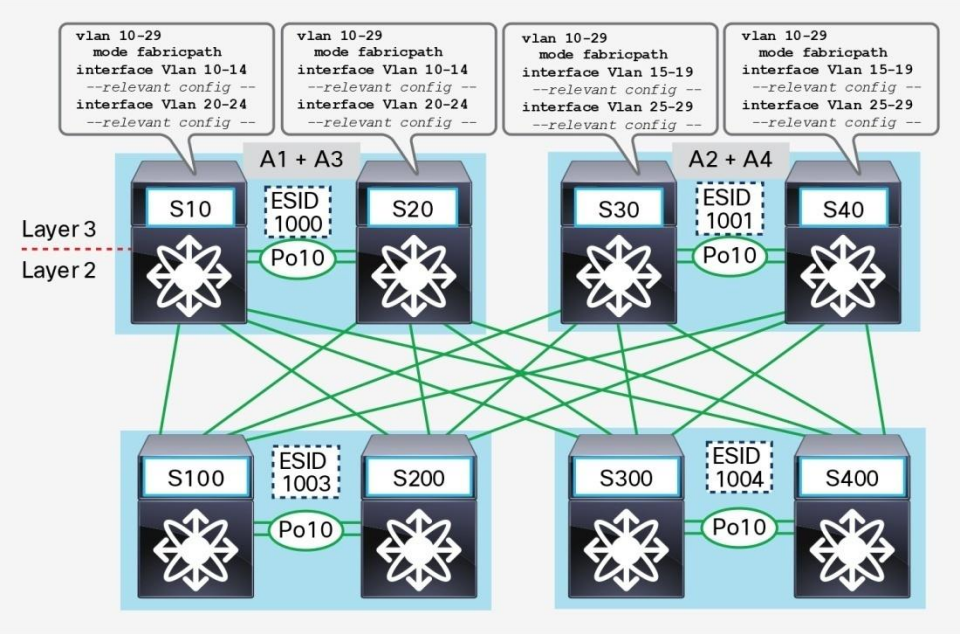
- A1 with A3: 11,500 MAC addresses (okay: fewer than 16,000 MAC addresses)
- A2 with A4: 11,500 MAC addresses (okay: fewer than 16,000 MAC addresses)

One of two groups can be terminated on a single pair of Cisco Nexus 7000 Series with F2e-Series modules.

You can either implement a second pair of Cisco Nexus 7000 Series Switches or split the existing Cisco Nexus 7000 Series Switches into two VDCs. Each of these VDCs will terminate SVIs for one of the two VLAN ranges.

Note that all VLANs in a given Cisco FabricPath topology still have to be configured on all relevant default gateways. This option assumes that only a subset of the SVIs is configured on a given default gateway. Refer to Figure 14 for details.

Figure 14. Segregation of Terminated VLANs Across Two Separate vPC+ Pairs

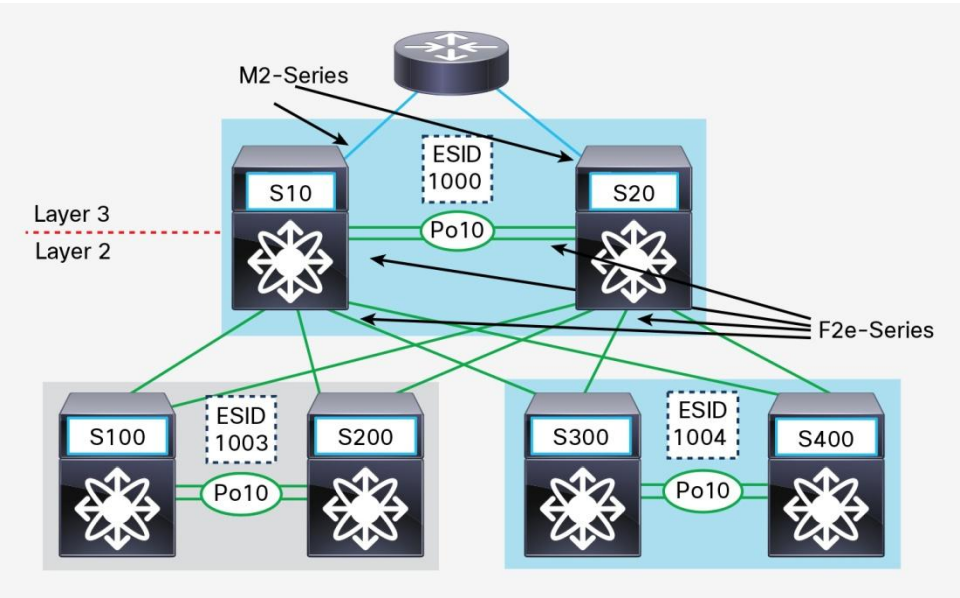


Option 2

Enable proxy Layer 2 learning mode using M2-Series modules. Beginning with Cisco NX-OS Release 6.2(2), F2e-Series modules can be placed with M2-Series modules in the same VDC. F2e-Series modules work in Layer 2 mode only; however, additional functions allow the use of the MAC address tables of the M2-Series modules.

This option applies only to Cisco Nexus 7000 Series Switches with F2e-Series and M2-Series modules. It allows a single pair of border leaf nodes to scale up to 128,000 MAC addresses (Figure 15).

Figure 15. Combine F2e and M2 Modules in the Same VDC for a High Scale MAC Address Table



To implement this option, follow these steps:

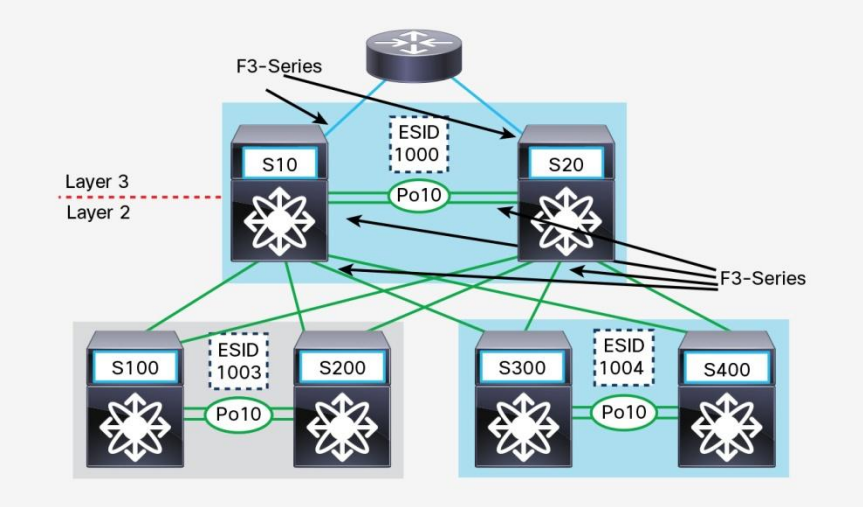
1. Verify that the Cisco FabricPath nodes with the default gateway function (also known as border leaf nodes) have F2e-Series and M2-Series I/O modules installed and initialized.
2. Disable remote MAC address learning on all border leaf nodes:
   `no mac address-table fabricpath remote-learning`
3. Disable Cisco FabricPath core port MAC address learning on all switch-on-a-chip (SOC) devices with core ports connected:
   `no hardware fabricpath mac-learning module <x> [port-group <x>]`
4. If you have leaf nodes that contain F2-Series modules, also disable Cisco FabricPath core port MAC address learning on all F2-Series SOC devices with core ports connected. Note that, in this case, you cannot combine core ports and Classic Ethernet edge ports on a single SOC.
   `no hardware fabricpath mac-learning module <x> [port-group <x>]`
5. Prune the list of allowed VLANs on Classic Ethernet edge ports:
   `switchport trunk allowed vlan <vlans>`

To achieve maximum MAC address scalability, you should avoid connecting Classic Ethernet devices to border leaf nodes. That is, border leaf nodes should have only Cisco FabricPath core ports (connected to an F2e-Series module), and routed Layer 3 interfaces (on an M2-Series modules).

Option 3

If your projected scale requirements do not exceed 64,000 MAC addresses, you can use F3-Series modules on border leaf nodes. Beginning with Cisco NX-OS 6.2(6), border leaf functions can be enabled on both Cisco Nexus 7000 Series Switches and Cisco Nexus 7700 platform switches with an F3-Series module, which has a capacity of 64,000 MAC addresses (Figure 16).

Figure 16. Using F3-Series Modules on Border Leaf nodes



Bidirectional Forwarding Detection and Cisco FabricPath

BFD provides fast failure detection in the network. The Cisco FabricPath network supports the BFD feature for Cisco Nexus 7000 Series Switches with Cisco NX-OS 7.2(0)D1(1) and later and Cisco Nexus 5000 and 6000 Series Switches with Cisco NX-OS 7.0(0)N1(1) and later. The supported BFD features are Cisco FabricPath IS-IS BFD and BFD on a Cisco FabricPath VLAN SVI.

Cisco FabricPath IS-IS BFD is enabled on a FabricPath IS-IS interface. BFD on a Cisco FabricPath VLAN SVI is enabled on a FabricPath VLAN SVI in a FabricPath network.

Cisco FabricPath IS-IS BFD

Cisco FabricPath IS-IS BFD can be enabled globally on all FabricPath IS-IS interfaces, as shown here:

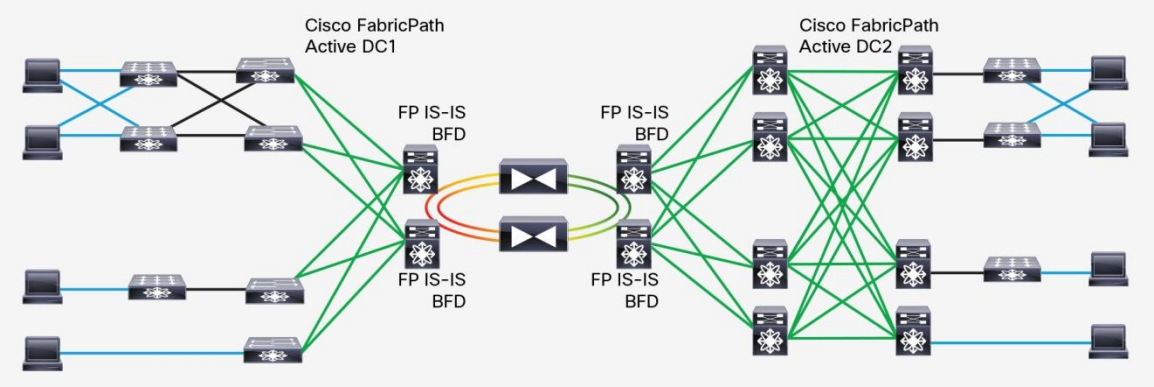
```
//Enable global FabricPath BFD on all FabricPath IS-IS interfaces:
FabricPath domain default
Bfd
```

Or FabricPath IS-IS BFD can be enabled on a specific interface, as shown here:

```
//Enable FabricPath BFD on a specific interface:
interface <interface_id>
fabricpath isis bfd
```

Whether you need to enable BFD on Cisco FabricPath IS-IS interfaces depends on your system. Cisco FabricPath network switches by default peer on physical point-to-point links. With a physical point-to-point link, the Layer 1 physical failure detection mechanism can be used, and detection occurs instantaneously. BFD doesn't provide much additional benefit. But if a Layer 1 device occurs in the path—for example, a dense wavelength division multiplexing (DWDM) transmission component between two Cisco FabricPath switches—then BFD needs to be enabled to help ensure fast failure detection: for instance, if the DWDM component doesn't relay the failure reliably from one side of the path to the other side. Figure 17 shows a Cisco FabricPath network scenario with multiple pods and multiple buildings and with FabricPath IS-IS BFD enabled on the FabricPath switches that connect to the DWDM component.

Figure 17. Cisco FabricPath Multiple-Pod and Multiple-Building Network Scenario



Cisco FabricPath IS-IS BFD is supported for Cisco Nexus 5000 and 6000 Series Switches with Cisco NX-OS 7.0(0)N1(1) and later, and for Cisco Nexus 7000 Series Switches with Cisco NX-OS 7.2(0)D1(1) and later. However, in a mixed environment with Cisco Nexus 5000/6000 and 7000 Series Switches, if you need to establish Cisco FabricPath ISIS BFD between Cisco Nexus 5000/6000 Series Switches and Cisco Nexus 7000 Series Switches, then the Cisco Nexus 5000/6000 Series Switches need to use Cisco NX-OS 7.2(0)N1(1) or later, and you need to enable a new encap-ce command to so that the switches can interoperate with Cisco Nexus 7000 Series Switches with Cisco NX-OS 7.2(0)D1(1) or later.

```
//one side of the link is N5000/6000, use minimum 7.2(0)N1(1) release and enable
encap-ce:
interface <interface_id>
       bfd fabricpath encap-ce
       fabricpath isis bfd

//the other side of the link is N7000, use minimum 7.2(0)D1(1) release:
       interface <interface_id>
       fabricpath isis bfd
```

Note that the BFD feature requires the LAN Base license, which is a Layer 3 license. For Cisco Nexus 5000 and 6000 Series Switches, Cisco In Service Software Upgrade (ISSU) is not supported with a Layer 3 license installed.

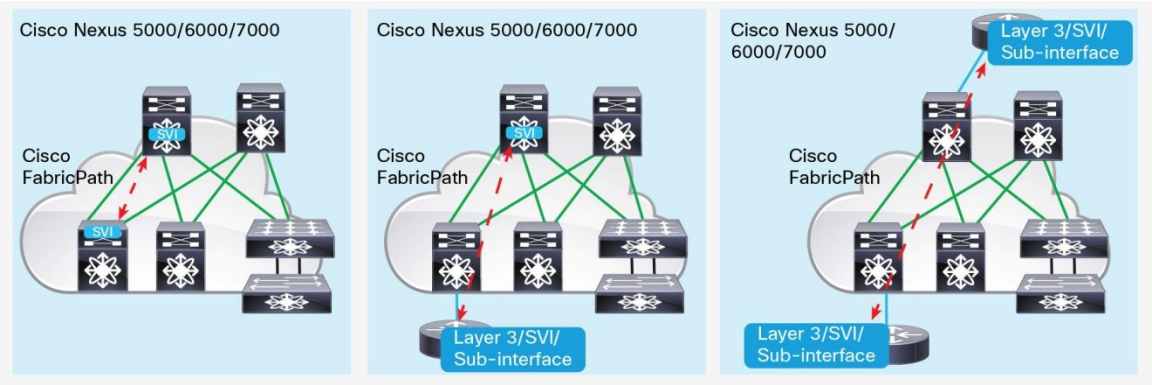
BFD on Cisco FabricPath VLAN SVI

BFD should also be used over a Cisco FabricPath network in other scenarios. For example, if a routing protocol is enabled on top of the Cisco FabricPath network, or if HSRP gateway communication occurs across the Cisco FabricPath network, then BFD needs to be enabled on the FabricPath VLAN SVI. BFD is used to detect failures more quickly so that routing and HSRP converge more quickly after a failure, rather than relying on routing protocol "hello" mechanisms for convergence. The general recommendation for these use cases is to set an appropriate BFD timer, default routing protocol timer, and HSRP hello timer based on customer requirements and failure scenario testing.

Figure 18 shows three examples of routing over BFD:

- SVI enabled on Cisco FabricPath switches, and Layer 3 routing peer with each other
- SVI enabled on a Cisco FabricPath switch, and Layer 3 routing peer with an external Layer 3 router
- Two external Layer 3 routers peer with each other over the Cisco FabricPath network

Figure 18. Routing Peers over Cisco FabricPath Network



When you use BFD on a Cisco FabricPath VLAN SVI, a less aggressive setting for the BFD timer is recommended. The reason is that for a Cisco FabricPath network, BFD supports equal-cost multipath (ECMP) for unicast traffic. In the event of a single link failure, the preferred approach is to give the Cisco FabricPath network time to reconverge and route unicast traffic through alternative routes. With a less aggressive BDF timer setting, BFD doesn't react to a single link failure; instead, BFD reacts to failures caused by multiple ECMP link failures. The BFD timer value set depends on your failure scenario and convergence time requirements.

```
//BFD timer example: with this setting, this local device detects a fault in the
forwarding path when there are 3 missing BFD hello messages from another BFD
device. Which means failure is detected after 0.75 second. The convergence time
is 0.75 seconds.


bfd interval 250 min_rx 250 multiplier 3
```

Additional Notes

Disable Optimized Multicast Flooding

When IPv6 is used in a network, the Internet Group Management Protocol (IGMP) Optimized Multicast Flooding (OMF) feature must be disabled for that VLAN. (Note that this requirement applies only to F2-Series modules, not to F2e- and F3-Series modules.)

```
S100(config)# vlan 10
S100 (config-vlan-config)# no ip igmp snooping optimise-multicast-flood
```

Configure Multicast Load Balancing with Cisco FabricPath When Using Fabric Extenders

When attaching a fabric extender (FEX) to a Cisco FabricPath Cisco Nexus 7000 Series leaf switch, the following configuration is required:

```
vpc domain 11
   fabricpath multicast load-balance
```

Protecting Against Broadcast Storms in a Cisco FabricPath Network

As a technology, Cisco FabricPath does not suppress broadcast storms, so additional configuration on the edge port is recommended. Network administrators need to evaluate the level of suppression required according to the hosted applications used and their density. Refer to the Configuring Traffic Storm Control configuration guide for more details.

Appendix A: Full Cisco FabricPath Network Configurations

This appendix presents configurations for three topologies and their respective VLANs:

- Topology 0: Default topology; VLANs 10 to 99
- Topology 1: S100 and S200 constitute a first pod; VLANs 110 to 199
- Topology 2: S300 and S400 constitute a second pod; VLANs 210 to 299

For simplicity, only a subset of VLANs is terminated with Anycast HSRP: VLANs 10 to 11.

In the following output, IGP configuration is omitted.

Configuration for Switch S10

```
S10# sh run fabricpath
version 6.2(6a)
feature-set fabricpath

fabricpath topology 1
  member vlan 110-199
fabricpath topology 2
  member vlan 210-299
vlan 10-99,110-199,210-299
  mode fabricpath
fabricpath switch-id 10
vpc domain 10
  fabricpath switch-id 1000


interface port-channel10
  switchport mode fabricpath
  fabricpath topology-member 1
  fabricpath topology-member 2

interface port-channel11
  switchport mode fabricpath
  fabricpath topology-member 1
```

```
interface port-channel12
  switchport mode fabricpath
  fabricpath topology-member 1

interface Ethernet6/3
  fabricpath topology-member 2
  switchport mode fabricpath

interface Ethernet6/4
  fabricpath topology-member 2
  switchport mode fabricpath

fabricpath domain default
  root-priority 245

S10# sh run hsrp

version 6.2(6a)
feature hsrp

hsrp anycast 100 ipv4
  switch-id 1100
  vlan 10-11
  priority 121
  no shutdown


interface Vlan10
  hsrp version 2
  hsrp 10
    ip 10.10.0.1

interface Vlan11
  hsrp version 2
  hsrp 11
    ip 10.11.0.1

S10# sh run vpc
```

```
version 6.2(6a)
feature vpc

vpc domain 10
  peer-keepalive destination 1.1.1.2 source 1.1.1.1 vrf vpc
  fabricpath switch-id 1000

interface port-channel10
  vpc peer-link
```

Configuration for Switch S20

```
S20# sh run fabricpath

version 6.2(6a)
feature-set fabricpath

fabricpath topology 1
  member vlan 110-199
fabricpath topology 2
  member vlan 210-299
vlan 10-99,110-199,210-299
  mode fabricpath
fabricpath switch-id 20
vpc domain 10
  fabricpath switch-id 1000


interface port-channel10
  switchport mode fabricpath
  fabricpath topology-member 1
  fabricpath topology-member 2

interface port-channel13
  switchport mode fabricpath
  fabricpath topology-member 1

interface port-channel14
  switchport mode fabricpath
```

```
  fabricpath topology-member 1

interface Ethernet6/3
  fabricpath topology-member 2
  switchport mode fabricpath

interface Ethernet6/4
  fabricpath topology-member 2
  switchport mode fabricpath

fabricpath domain default
  root-priority 244


S20# sh run hsrp

version 6.2(6a)
feature hsrp


hsrp anycast 100 ipv4
  switch-id 1100
  vlan 10-11
  priority 120
  no shutdown


interface Vlan10
  hsrp version 2
  hsrp 10
    ip 10.10.0.1

interface Vlan11
  hsrp version 2
  hsrp 11
    ip 10.11.0.1


S20# sh run vpc
```

```
version 6.2(6a)
feature vpc

vpc domain 10
  peer-keepalive destination 1.1.1.1 source 1.1.1.2 vrf vpc
  fabricpath switch-id 1000

interface port-channel10
  vpc peer-link
```

Configuration for Switch S30

```
S30# sh run fabricpath

!Command: show running-config fabricpath
!Time: Wed Apr 16 19:32:13 2014

version 6.2(6a)
feature-set fabricpath

fabricpath topology 1
  member vlan 110-199
fabricpath topology 2
  member vlan 210-299
vlan 10-99,110-199,210-299
  mode fabricpath
fabricpath switch-id 30


interface Ethernet6/15
  fabricpath topology-member 1
  switchport mode fabricpath

interface Ethernet6/16
  fabricpath topology-member 1
  switchport mode fabricpath

interface Ethernet6/17
  fabricpath topology-member 2
```

```
  switchport mode fabricpath

interface Ethernet6/18
  fabricpath topology-member 2
  switchport mode fabricpath
fabricpath domain default


S30# sh run hsrp

version 6.2(6a)
feature hsrp

hsrp anycast 100 ipv4
  switch-id 1100
  vlan 10-11
  priority 115
  no shutdown

interface Vlan10
  hsrp version 2
  hsrp 10
    ip 10.10.0.1

interface Vlan11
  hsrp version 2
  hsrp 11
    ip 10.11.0.1
```

Configuration for Switch S40

```
S40# sh run fabricpath

version 6.2(6a)
feature-set fabricpath

fabricpath topology 1
  member vlan 110-199
fabricpath topology 2
  member vlan 210-299
```

```
vlan 10-99,110-199,210-299
  mode fabricpath
fabricpath switch-id 40

interface Ethernet6/15
  fabricpath topology-member 1
  switchport mode fabricpath

interface Ethernet6/16
  fabricpath topology-member 1
  switchport mode fabricpath

interface Ethernet6/17
  fabricpath topology-member 2
  switchport mode fabricpath

interface Ethernet6/18
  fabricpath topology-member 2
  switchport mode fabricpath
fabricpath domain default


S40# sh run hsrp

version 6.2(6a)
feature hsrp

hsrp anycast 100 ipv4
  switch-id 1100
  vlan 10-11
  priority 110
  no shutdown

interface Vlan10
  hsrp version 2
  hsrp 10
    ip 10.10.0.1

interface Vlan11
```

```
  hsrp version 2
  hsrp 11
    ip 10.11.0.1
```

Configuration for Switch S100

```
S100# sh run fabricpath

version 6.2(6a)
feature-set fabricpath

fabricpath topology 1
  member vlan 110-199
vlan 10-99,110-199
  mode fabricpath
fabricpath switch-id 100
vpc domain 11
  fabricpath switch-id 1003
  fabricpath multicast load-balance

interface port-channel10
  switchport mode fabricpath
  fabricpath topology-member 1

interface port-channel11
  switchport mode fabricpath
  fabricpath topology-member 1

interface port-channel13
  switchport mode fabricpath
  fabricpath topology-member 1

interface Ethernet6/29
  fabricpath topology-member 1
  switchport mode fabricpath

interface Ethernet6/30
  fabricpath topology-member 1
  switchport mode fabricpath
```

```
fabricpath domain default
  root-priority 243
  topology 1
    root-priority 245

S100# sh run vpc

version 6.2(6a)
feature vpc

vpc domain 11
  peer-keepalive destination 1.1.1.2 source 1.1.1.1 vrf vpc
  fabricpath switch-id 1003
  fabricpath multicast load-balance

interface port-channel10
  vpc peer-link

interface port-channel20
  vpc 20

interface port-channel30
  vpc 30
```

Configuration for Switch S200

```
S200# sh run fabricpath

version 6.2(6a)
feature-set fabricpath

fabricpath topology 1
  member vlan 110-199
vlan 10-99,110-199
  mode fabricpath
fabricpath switch-id 200
vpc domain 11
  fabricpath switch-id 1003
  fabricpath multicast load-balance


interface port-channel10
  switchport mode fabricpath
  fabricpath topology-member 1

interface port-channel12
  switchport mode fabricpath
  fabricpath topology-member 1

interface port-channel14
  switchport mode fabricpath
  fabricpath topology-member 1

interface Ethernet6/29
  fabricpath topology-member 1
  switchport mode fabricpath

interface Ethernet6/30
  fabricpath topology-member 1
  switchport mode fabricpath

fabricpath domain default
  topology 1
```

```
    root-priority 244

S200# sh run vpc

!Command: show running-config vpc
!Time: Wed Apr 16 19:53:14 2014

version 6.2(6a)
feature vpc

vpc domain 11
  peer-keepalive destination 1.1.1.1 source 1.1.1.2 vrf vpc
  fabricpath switch-id 1003
  fabricpath multicast load-balance

interface port-channel10
  vpc peer-link

interface port-channel20
  vpc 20

interface port-channel30
  vpc 30
```

Appendix B: Why Enable Spanning-Tree Pseudo-Information Commands on Cisco FabricPath Switches

The section "Cisco FabricPath and Spanning Tree Protocol Connectivity Options" noted that if you are connecting spanning-tree devices to the Cisco FabricPath fabric, make sure that you configure all edge switches as spanning-tree root devices by using the **spanning-tree vlan x root primary** command (or manually configure the bridge priority on each switch to force the switch to be root). Note that the following **spanning-tree priority** command would also work;

```
S100(config)# spanning-tree vlan x root primary
Or S100(config)# spanning-tree vlan x priority 8192
```

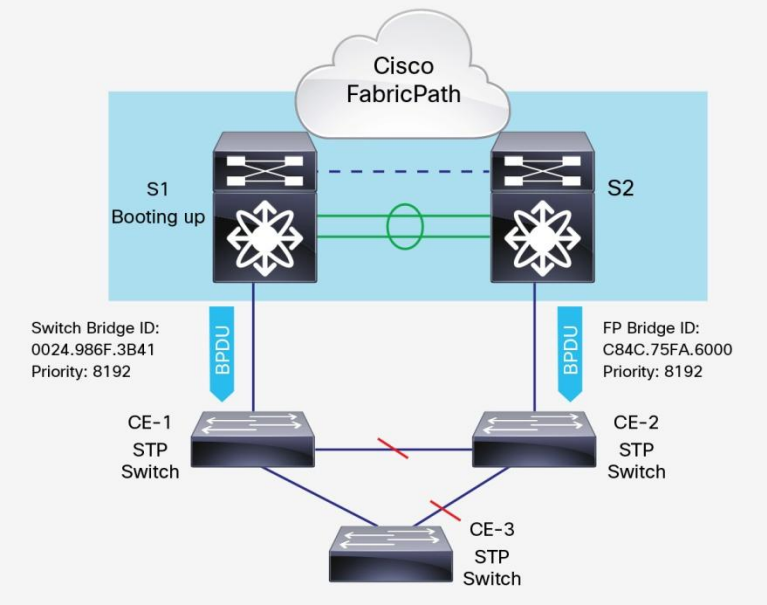
However, this configuration would change the priority for the spanning tree depending on whether the switch is sending regular BPDUs (when Cisco FabricPath is not running) or sending BPDUs when Cisco FabricPath is operational on the switch. In some scenarios, this change can have undesirable side effects. For example, in Figure 19, when a Cisco FabricPath edge switch is reloaded (before it is active in FabricPath), it behaves like a traditional spanning-tree device on its edge ports. It sends a bridge ID with its own system MAC address (0024.986F.3B41) and the configured spanning-tree priority (8192)—not the common Cisco FabricPath bridge ID (C84C.75FA.6000). Therefore, during a reload operation, an edge switch may start to transmit superior BPDUs

(because the local system MAC address may be lower than the common Cisco FabricPath bridge ID) before it becomes active in Cisco FabricPath. Whenever the remaining active Cisco FabricPath edge switch (S2) receives a superior BPDU from the attached spanning-tree switch, it places its customer-edge (CE) port in a Layer 2 gateway Inconsistency state. A syslog message similar to the following is generated:

2015 Jul 30 19:33:03 N7K-SW %STP-2-L2GW_BACKBONE_BLOCK: L2 Gateway Backbone port inconsistency blocking port Ethernet1/1 on VLAN0032.

This condition is cleared after the other Cisco FabricPath edge switch is reconnected to the FabricPath network and starts to send common bridge ID and priority information.

Figure 19. With **Spanning-Tree Priority** Command



To avoid a potential spanning-tree root change during a Cisco FabricPath edge switch reload, configure the same root priority on all FabricPath edge switches using the **spanning-tree pseudo information** command:

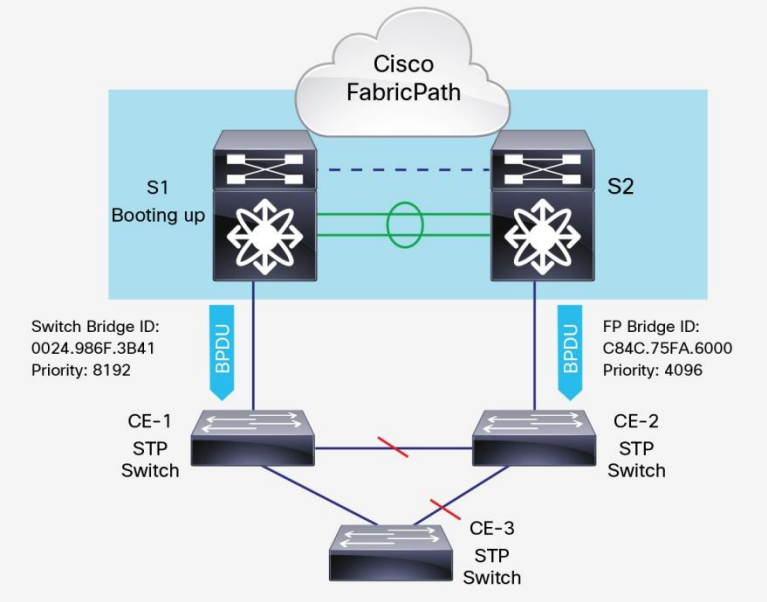
```
S100(config)# spanning-tree pseudo-information
S100(config-pseudo)#  vlan 10-100 root priority 4096
```

The **spanning-tree pseudo-information** knob was originally introduced for vPC environments to enable VLAN load-balancing for spanning-tree attached access devices and to avoid a spanning-tree topology change when a peer device recovers (after a failure or reload). The command works well with vPC+ environments. With the **spanning-tree pseudo-information** command on a Cisco FabricPath edge switch, no matter which FabricPath edge switch is reloading, the remaining FabricPath switch will send a superior BPDU with a common FabricPath bridge ID and the configured spanning-tree pseudo-information priority and remains at spanning-tree root. The reason for this behavior is that before the reloaded Cisco FabricPath edge switch becomes active in FabricPath, it sends a bridge ID with its own system MAC address and the default spanning-tree priority of 32768 (which is an inferior BPDU).

Note that it is fine to use the **spanning-tree priority** command and the **spanning-tree pseudo-information** command together, but the root priority in the pseudo-information needs to be a lower value (higher priority) than the spanning-tree priority to make sure that the remaining working Cisco FabricPath switch is sending a superior BPDU with the common FabricPath bridge ID and the configured spanning-tree pseudo-information priority and remains at the spanning-tree root (Figure 20).

```
S100(config)# spanning-tree vlan x priority 8192
S100(config)# spanning-tree pseudo-information
S100(config-pseudo)#  vlan 10-100 root priority 4096
```

Figure 20. With Both Spanning-Tree Pseudo-Information Command and Spanning-Tree Priority Command



Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at www.cisco.com/go/offices.

Printed in USA C07-728188-02 03/16