DCUFT

Troubleshooting Cisco Data Center Unified Fabric

Volume 1

Version 5.0

Student Guide

Text Part Number: 97-3206-01

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.**

Students, this letter describes important course evaluation access information!

Welcome to Cisco Systems Learning. Through the Cisco Learning Partner Program, Cisco Systems is committed to bringing you the highest-quality training in the industry. Cisco learning products are designed to advance your professional goals and give you the expertise you need to build and maintain strategic networks.

Cisco relies on customer feedback to guide business decisions; therefore, your valuable input will help shape future Cisco course curricula, products, and training offerings. We would appreciate a few minutes of your time to complete a brief Cisco online course evaluation of your instructor and the course materials in this student kit. On the final day of class, your instructor will provide you with a URL directing you to a short post-course evaluation. If there is no Internet access in the classroom, please complete the evaluation within the next 48 hours or as soon as you can access the web.

On behalf of Cisco, thank you for choosing Cisco Learning Partners for your Internet technology training.

Sincerely,

Cisco Systems Learning

Table of Contents

Volume 1

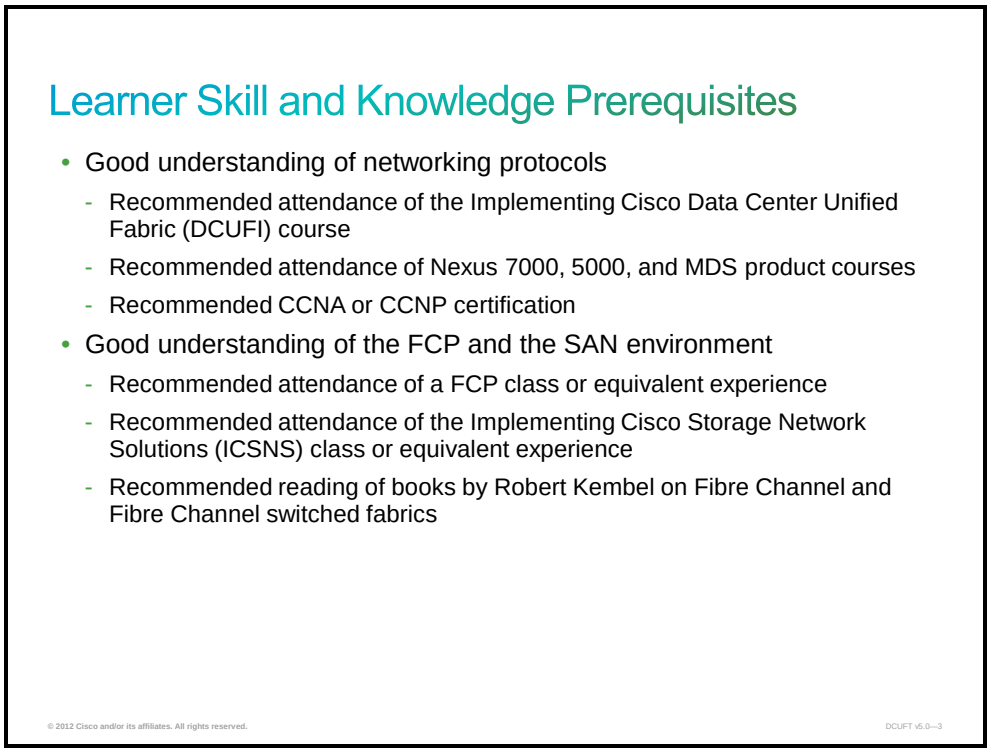
Course Introduction

Overview

Troubleshooting Cisco Data Center Unified Fabric (DCUFT) v5.0 is a three-day instructor-led course. It is designed for systems and field engineers, consulting systems engineers, and Cisco integrators and partners who install, implement, maintain, and troubleshoot the Cisco Nexus 7000 and 5000 Switches, the Cisco Nexus 2000 Fabric Extenders, and Cisco MDS Multilayer Fabric Switches. The course covers the key components and procedures needed to troubleshoot and resolve common issues with the Cisco Nexus 7000, 5000, MDS Switches, and Nexus 2000 Fabric Extenders in the network and SAN environment.

Learner Skills and Knowledge

This subtopic lists the skills and knowledge that learners must possess to benefit fully from the course. The subtopic also includes recommended Cisco learning offerings that learners should first complete to benefit fully from this course.



Before attending this course learners should be familiar with networking protocols and technologies, the SAN environment, and the Fibre Channel Protocol (FCP).

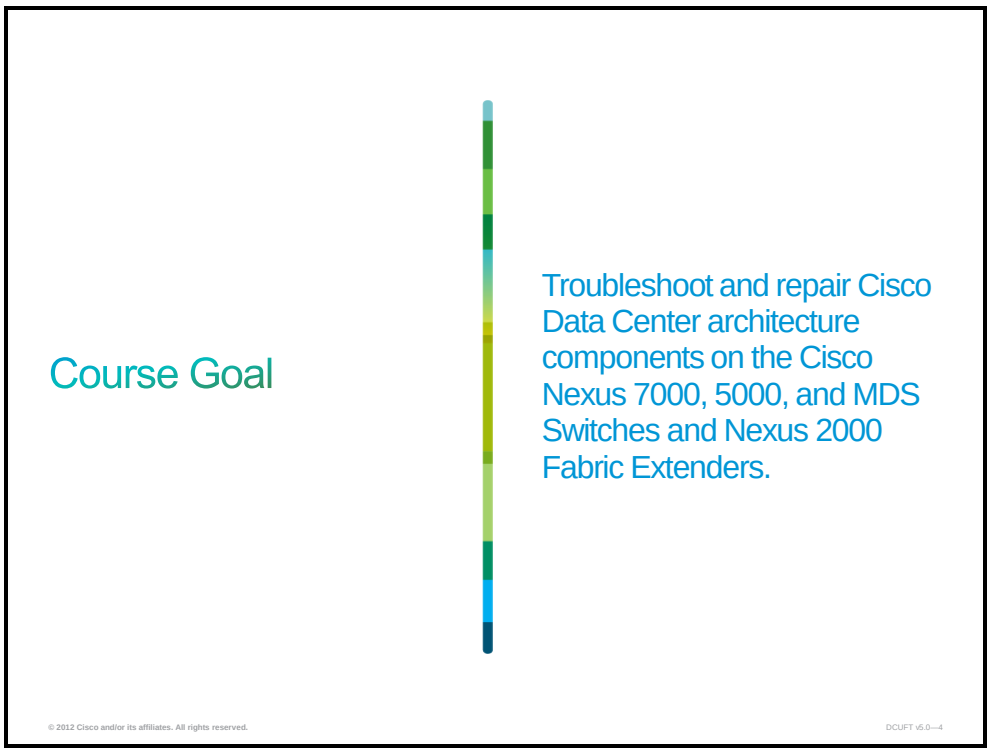
The learner should have attended the *Implementing Cisco Data Center Unified Fabric* (DCUFI) course and Nexus 7000, 5000, and MDS product courses. Cisco CCNA® or CCNP® level of knowledge is also recommended for students attending the DCUFT course.

Note	The recommended courses for CCNA are the *Interconnecting Cisco Network Devices Part 1* (ICND1) and *Interconnecting Cisco Network Devices Part 2* (ICND2) courses.

In order to attain the appropriate level of knowledge of the FCP and SAN environment, the learner should have attended an FCP course such as the *Implementing Cisco Storage Network Solutions* (ICSNS) course. The recommended reading includes the books by Robert Kembel on Fibre Channel and Fibre Channel switched fabrics.

Course Goal and Objectives

This topic describes the course goal and objectives.

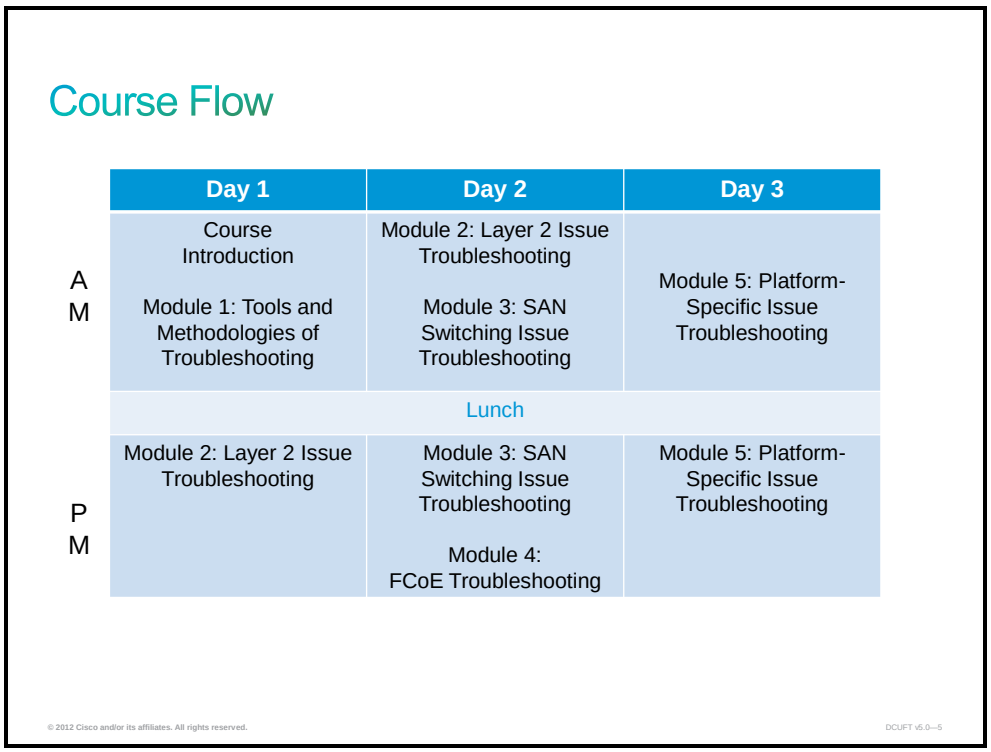


Upon completing this course, you will be able to meet these objectives:

- Describe the troubleshooting tools and methodologies that are available from the CLI and in Cisco DCNM that are used to identify and resolve issues in a Cisco Data Center network architecture
- Identify and resolve issues that are related to VLANs and PVLANs
- Identify and resolve issues that are related to port channels and vPCs
- Identify and resolve issues that are related to Cisco FabricPath
- Identify and resolve issues that are related to OTV
- Identify and resolve issues that are related to Fibre Channel interface operation
- Identify and resolve issues that are related to Fibre Channel switching when a Cisco Nexus switch is used in switched mode or in NPV mode
- Identify and resolve issues that are related to FCoE in the Cisco Data Center architecture
- Identify and resolve issues that are related to Cisco Nexus 7000 Series Switches
- Identify and resolve issues that are specific to Cisco Nexus 5000 Series Switches
- Identify and resolve issues that are specific to Cisco Nexus 2000 Series Fabric Extenders
- Identify and resolve issues that are specific to Cisco MDS Switches

Course Flow

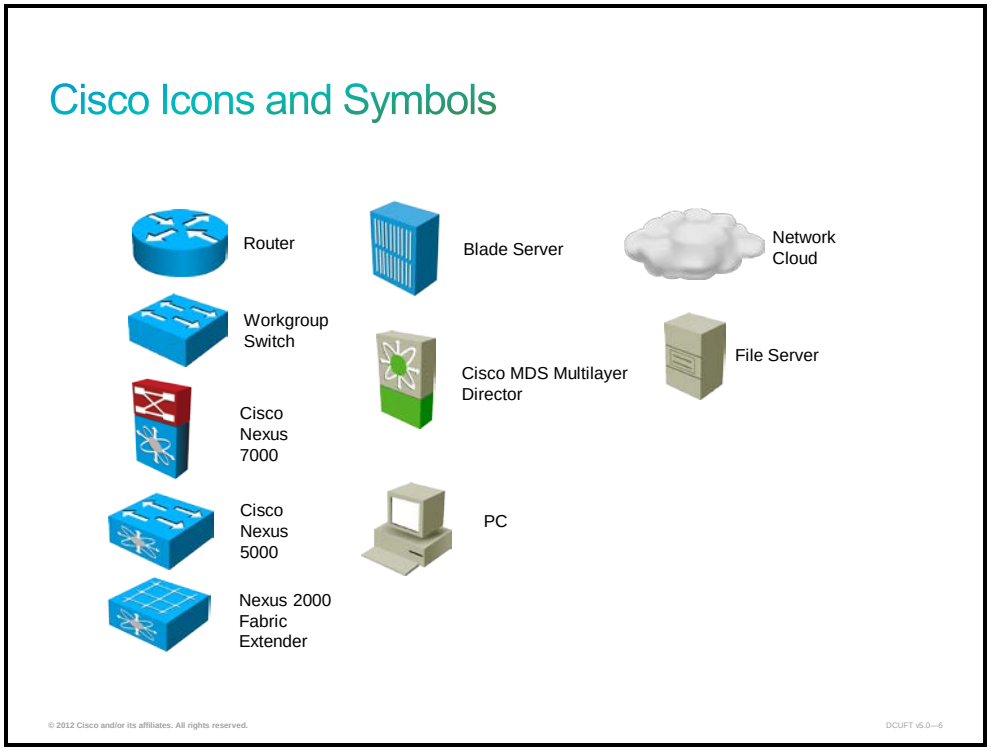
This topic presents the suggested flow of the course materials.



The schedule reflects the recommended structure for this course. This structure allows enough time for the instructor to present the course information and for you to work through the lab activities. The exact timing of the subject materials and labs depends on the pace of your specific class.

Additional References

This topic presents the Cisco icons and symbols that are used in this course, as well as information on where to find additional technical references.

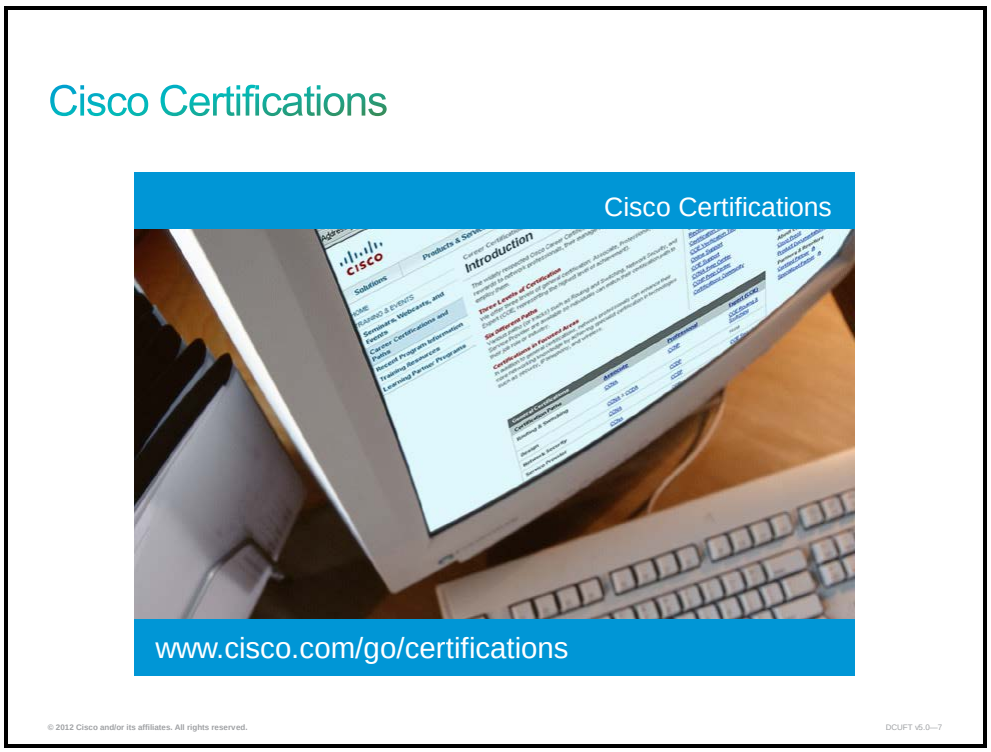


Cisco Glossary of Terms

For additional information on Cisco terminology, refer to the *Cisco Internetworking Terms and Acronyms* glossary of terms at http://docwiki.cisco.com/wiki/Internetworking_Terms_and_Acronyms_%28ITA%29_Guide.

Your Training Curriculum

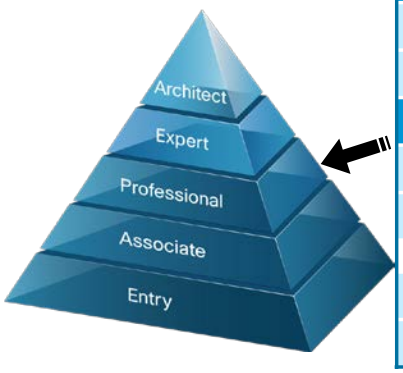
This topic presents the training curriculum for this course.



You are encouraged to join the Cisco Certification Community, a discussion forum open to anyone holding a valid Cisco Career Certification (such as Cisco CCIE®, CCNA®, CCDA®, CCNP®, CCDP®, CCIP®, CCVP®, or CCSP®). It provides a gathering place for Cisco certified professionals to share questions, suggestions, and information about Cisco Career Certification programs and other certification-related topics. For more information, visit www.cisco.com/go/certifications.

Cisco Career Certifications: Cisco CCNP Data Center

Expand Your Professional Options and Advance Your Career



Cisco CCNP Data Center
Implementing Cisco Data Center Unified Fabric (DCUFI)
Implementing Cisco Data Center Unified Computing (DCUCI)
Available Exams (pick a group of 2)
Designing Cisco Data Center Unified Computing (DCUCD)
Designing Cisco Data Center Unified Fabric (DCUFD)
or
Troubleshooting Cisco Data Center Unified Fabric (DCUFT)
Troubleshooting Cisco Data Center Unified Computing (DCUCT)

www.cisco.com/go/certifications

Module 1

Tools and Methodologies of Troubleshooting

Overview

This module is designed to show the student some of the common tools and methodologies that are used in troubleshooting the Cisco Data Center architecture.

Module Objectives

Upon completing this module, you will be able to describe the troubleshooting tools and methodologies that are used to identify and resolve issues in the Cisco Data Center network architecture. This ability includes being able to meet these objectives:

- Describe the troubleshooting tools and methodologies that are available from the CLI that are used to identify and resolve issues in a Cisco Data Center network architecture
- Describe the troubleshooting tools and methodologies that are available in Cisco DCNM that are used to identify and resolve issues in a Cisco Data Center network architecture

Lesson 1

Understanding CLI Troubleshooting Tools

Overview

Describe the troubleshooting tools and methodologies that are available from the CLI that are used to identify and resolve issues in the Cisco Data Center network architecture.

Objectives

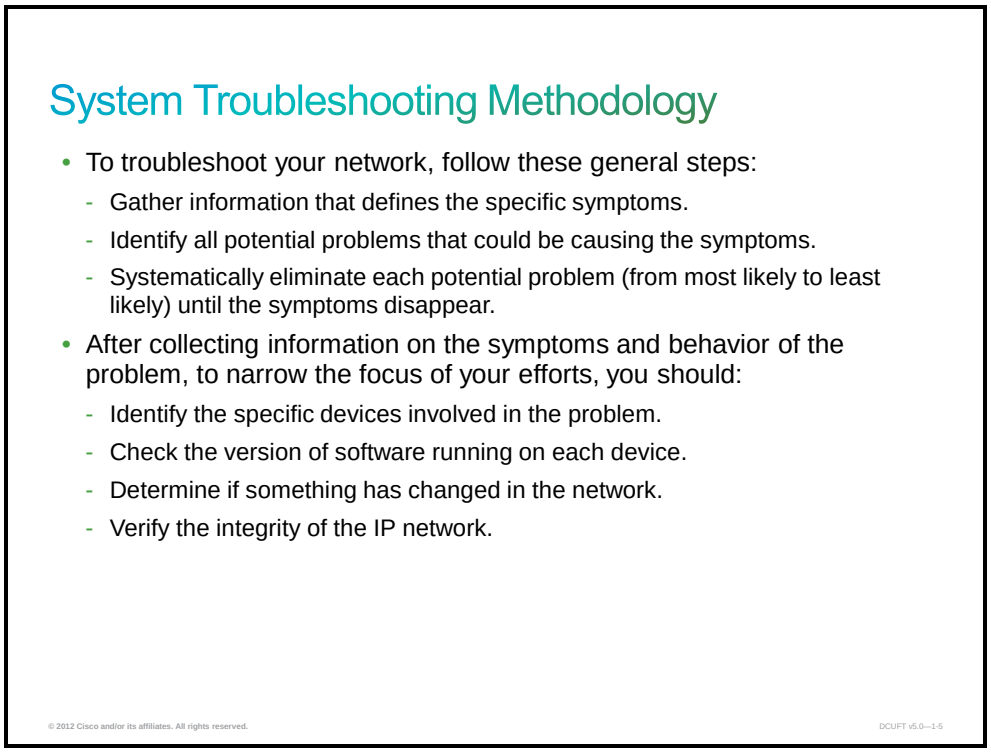
Upon completing this lesson, you will be able to describe the troubleshooting tools and methodologies that are available from the CLI that are used to identify and resolve issues in a Cisco Data Center network architecture. You will be able to meet these objectives:

- Provide the basic concepts, methodology, and general troubleshooting guidelines for problems that may occur when configuring and using Cisco NX-OS
- Explain how to use ping and traceroute to troubleshoot problems with connectivity, and how to use path choices of the pong feature to measure the delay of the network between two points
- Explain how to use the CLI to monitor the system, processes, and the CPU
- Explain how to use the CLI SPAN utility to perform detailed troubleshooting from a particular application host for proactive monitoring and analysis
- Explain how to use Ethanalyzer to troubleshoot your network and analyze control-plane traffic
- Explain how to use the logging feature to log information for monitoring and troubleshooting of the Cisco Nexus or MDS switch
- Explain how to use GOLD to collect diagnostic results and detailed statistics on the Cisco Nexus or MDS switch
- Explain how to use the blue beacon to aid in the troubleshooting or replacement of components in the Cisco Nexus switch

- Explain how to use SNMP and RMON to monitor and troubleshoot a Cisco Nexus or MDS switch
- Explain how to use the CLI debug feature to show real-time information while actively troubleshooting a network

Troubleshooting Methodology

This topic introduces the basic concepts, methodology, and general troubleshooting guidelines for problems that may occur when configuring and using the Cisco Nexus Operating System (NX-OS).

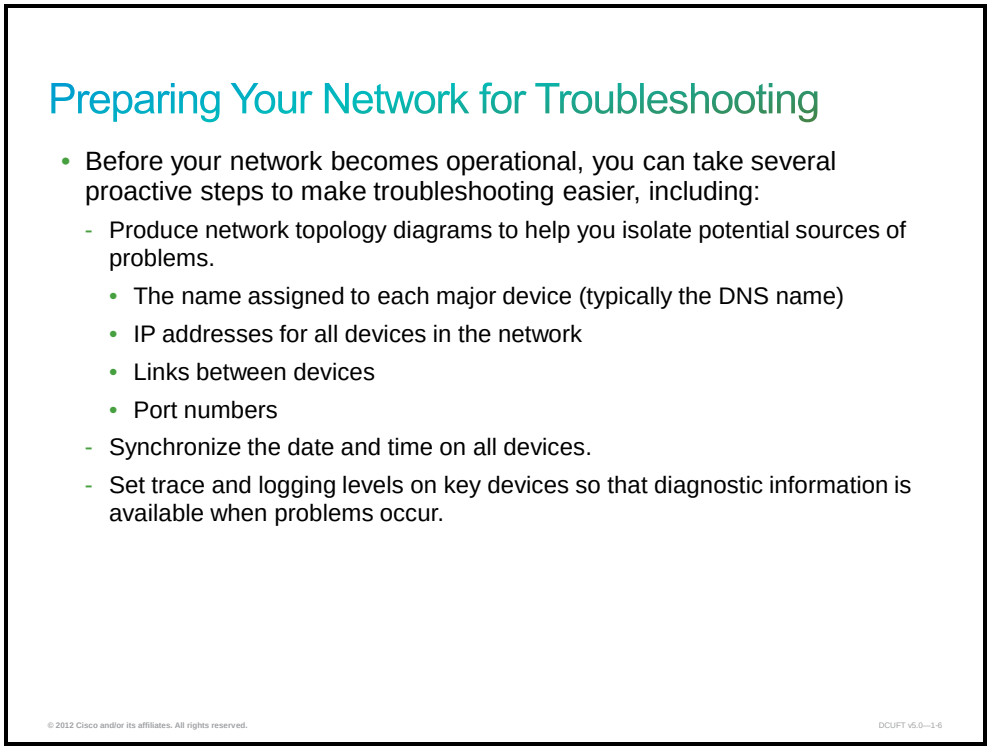


The implementation phase of your network deployment is an excellent time to develop a methodology for troubleshooting the network as a whole. When a problem occurs, the list of potential suspects can be long. You must collect detailed information and systematically narrow the list of potential causes to determine the root problem. To troubleshoot your network, follow these general steps:

1. Gather information that defines the specific symptoms.
2. Identify all potential problems that could be causing the symptoms.
3. Systematically eliminate each potential problem (from most likely to least likely) until the symptoms disappear.

After collecting information on the symptoms and behavior of the problem, to narrow the focus of your efforts and to isolate point(s) of failure, you should do the following:

- Identify the specific devices involved in the problem.
- Check the version of software running on each device.
- Determine if something has changed in the network.
- Verify the integrity of the IP network.



Before your network becomes operational, you can take several proactive steps to make troubleshooting easier, including the following:

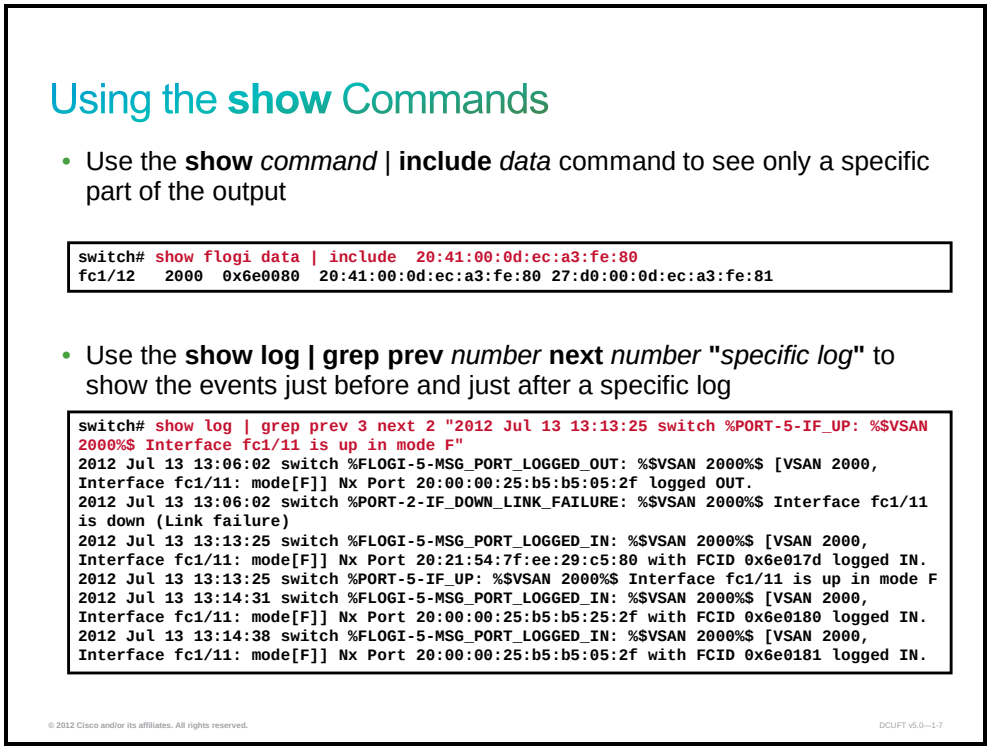
- Produce network topology diagrams to help you isolate potential sources of problems.
- Synchronize the date and time on all servers.
- Set trace and logging levels on key devices so that diagnostic information is available when problems occur.

One of the first lines of defense is possessing current topology information. One of the most important pieces of topology information is a detailed network diagram. At a minimum, your network topology diagrams should include the following information:

- The name assigned to each major device (typically the Domain Name System [DNS] name)
- IP addresses for all devices in the network
- Links between devices
- Port numbers

This information is critical for isolating which components are involved in a particular problem. For medium- to large-sized networks, you may want to take a "layered" approach in your diagrams. Create a high-level diagram that illustrates the overall physical layout of your network, including all sites and the links between them. Then, for each site, create additional diagrams that show detailed addressing information and port number configurations. Frequent changes and upgrades to your network can quickly make these diagrams out-of-date. Inaccurate diagrams slow down the troubleshooting process and may lead to misdiagnosing the problem. Remember to keep these diagrams as current as possible.

Tracing can be enabled on multiple devices and the log file output can be compared to isolate problems. In order to correlate messages for the same activity in different log files, you must compare the message time stamps and the source device MAC and IP addresses. You should synchronize every device to the same date and time source so that the time stamps match. To accomplish this synchronization, set each device to obtain its date and time from the same Network Time Protocol (NTP) source.



In addition to using the general **show** command, you can narrow down the output to a specific part by using the **include** command.

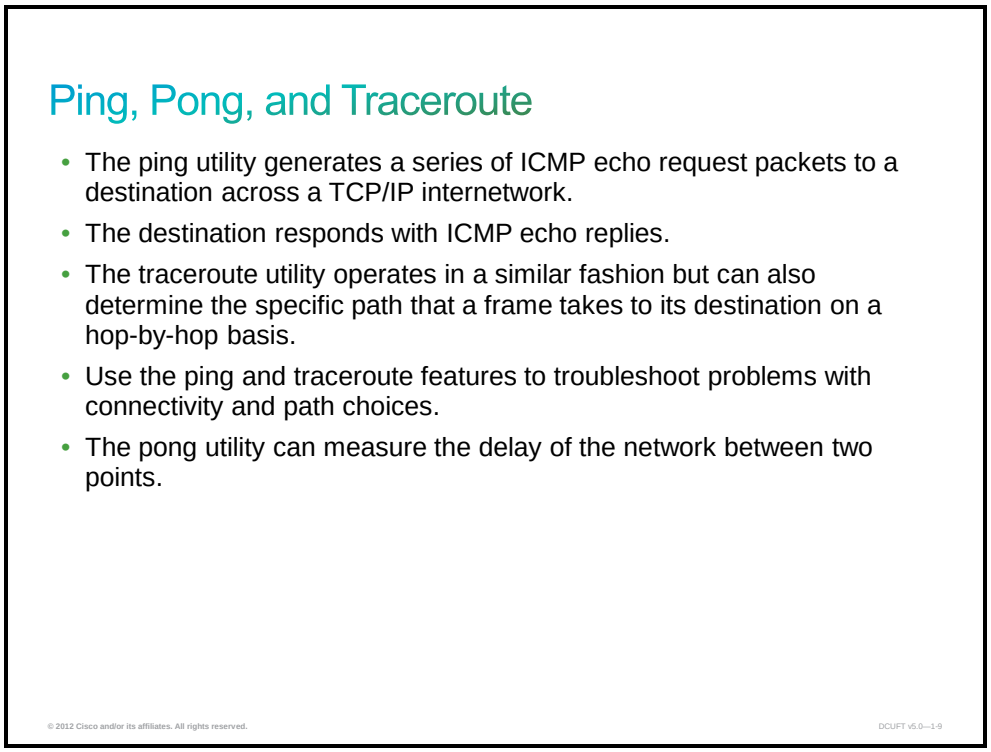
When displaying logging output you can also use the **grep** command to show the events just before and just after a specific log. You can also see any changes made around a certain time by using the **show accounting log | grep** *timestamp* command.

If you would like to see only the last few lines of the output, you can use the **show** *command* **last** *number* command:

```
switch# show log last 5
2012 Jul 18 21:27:18 switch %FLOGI-5-MSG_PORT_LOGGED_OUT: %$VSAN
2000%$ [VSAN 2000, Interface fc1/18: mode[F]] Nx Port
20:00:00:25:b5:b0:05:01 logged OUT.
2012 Jul 18 21:27:19 switch %FLOGI-5-MSG_PORT_LOGGED_IN: %$VSAN 2000%$
[VSAN 2000, Interface fc1/18: mode[F]] Nx Port 20:00:00:25:b5:b0:05:01
with FCID 0x6e0158 logged IN.
2012 Jul 19 03:48:53 switch %AUTHPRIV-3-SYSTEM_MSG:
pam_aaa:Authentication failed for user chlee3 from 10.154.36.171 -
sshd[19584]
2012 Jul 19 03:48:53 switch %DAEMON-3-SYSTEM_MSG: error: PAM:
Authentication failure for chlee3 from 10.154.36.171 - sshd[19583]
2012 Jul 19 03:48:54 switch %DAEMON-3-SYSTEM_MSG: error: ssh_msg_send:
write - sshd[19585]
```

Ping, Pong, and Traceroute

This topic explains how to use ping and traceroute to troubleshoot problems with connectivity, and how to use path choices of the pong feature to measure the delay of the network between two points.



The CLI allows a user to configure and monitor Cisco NX-OS using a local console or remotely by using a Telnet or Secure Shell (SSH) session. The CLI provides a command structure similar to Cisco IOS Software, with context-sensitive help, **show** commands, multiuser support, and role-based access control (RBAC).

The **ping** and **traceroute** commands can be used to troubleshoot problems with connectivity and path choices. These features are not used to identify or resolve network performance issues, but the **pong** command can be used for that.

The **ping** and **traceroute** commands are two of the most useful tools for troubleshooting TCP/IP networking problems. The ping utility generates a series of Internet Control Message Protocol (ICMP) echo request packets to a destination across a TCP/IP internetwork. The destination responds with ICMP echo replies. If pings fail, the ping at the source does not tell you if the problem is with the requests getting to the destination, or the replies coming back.

The **traceroute** command operates in a similar fashion but can also determine the specific path that a frame takes to its destination on a hop-by-hop basis.

The **pong** command can measure the delay of the network between two points.

Using the **ping** Command

- Use the **ping** command to verify connectivity and latency to a particular destination across an IPv4 routed network.

```
Switch#ping 172.28.230.1 vrf management

PING 172.28.230.1 (172.28.230.1): 56 data bytes
64 bytes from 172.28.230.1: icmp_seq=0 ttl=254 time=1.095 ms
64 bytes from 172.28.230.1: icmp_seq=1 ttl=254 time=1.083 ms
64 bytes from 172.28.230.1: icmp_seq=2 ttl=254 time=1.101 ms
64 bytes from 172.28.230.1: icmp_seq=3 ttl=254 time=1.093 ms
64 bytes from 172.28.230.1: icmp_seq=4 ttl=254 time=1.237 ms
```

- Use the **ping6** command to verify connectivity and latency to a particular destination across an IPv6 routed network.

```
Switch#ping6 2001:0DB8::200C:417A vrf management

PING6 2001:0DB8::200C:417A (2001:0DB8::200C:417A): 56 data bytes
64 bytes from 2001:0DB8::200C:417A: icmp_seq=0 time=2.307 ms
64 bytes from 2001:0DB8::200C:417A: icmp_seq=1 time=2.094 ms
64 bytes from 2001:0DB8::200C:417A: icmp_seq=2 time=1.894 ms
64 bytes from 2001:0DB8::200C:417A: icmp_seq=3 time=1.702 ms
64 bytes from 2001:0DB8::200C:417A: icmp_seq=4 time=1.678 ms
```


The **ping** command is used to verify connectivity and latency to a particular destination across an IPv4 routed network.

The **ping6** command is used to verify connectivity and latency to a particular destination across an IPv6 routed network.

The **ping** command allows a user to send an ICMP echo request to a port or end device. By specifying the IPv4 or IPv6 address, a user can send a series of ICMP echo request packets to a target destination. Once these packets reach the target, the ICMP echo replies are sent back to the source and a time stamp is taken.

There are several **ping** CLI options available; some of the most useful include **count** and **packet-size**:

- **ping *dest-address* count *number***: This command specifies the number of transmissions to send.
- **ping *dest-address* packet-size *bytes***: This command specifies the packet size in bytes to transmit. The range is from 1 to 65468. The default is 56 bytes.

You can use the same commands replacing **ping** with **ping6**.

Using the **pong** Command

- Use the **pong** command to measure port-to-port delays.
- The pong utility can be enabled only on F and M series module ports.
- The example shows the pong service between FabricPath Switch IDs:

```
switch# configure terminal
switch(config)# pong destination-swid 3506 destination-mac 001b.54c2.9a43 vlan 1 count 2

Packet No. 1
--- ----------------- --------------------------
Hop Switch-id Switching time (sec, nsec)
--- ----------------- --------------------------
1 0-1b-54-c2-9a-41 0 4752
2 0-1b-54-c2-9a-43 0 544258088
3 0-1b-54-c2-9a-41 0 4792
Round trip time: 0sec 15416 nsec

Packet No. 2
--- ----------------- --------------------------
Hop Switch-id Switching time (sec, nsec)
--- ----------------- --------------------------
1 0-1b-54-c2-9a-41 0 4752
2 0-1b-54-c2-9a-43 0 522744240
3 0-1b-54-c2-9a-41 0 4736
Round trip time: 0sec 15368 nsec
```


The **pong** command is used to measure port-to-port delays. It uses replies on the IEEE 1588-2008 (v2) Precision Time Protocol (PTP), a time synchronization protocol for nodes distributed across a network. The **pong** utility is similar to the network-monitoring utility ping, but provides for a greater depth of network diagnostics.

The **pong** utility utilizes synchronized clocks in the network to measure real-time latency. Latency is the delay of the network between any two points that are as seen by a frame traveling between the two points.

The **pong** utility can be enabled on Cisco Nexus 7000 F1 and F2 series module ports. The M2 series module supports it as well; however, you need to use the **feature pong** command.

There are several options for use of the **pong** command:

- This example shows the **pong** service between FabricPath Switch IDs:

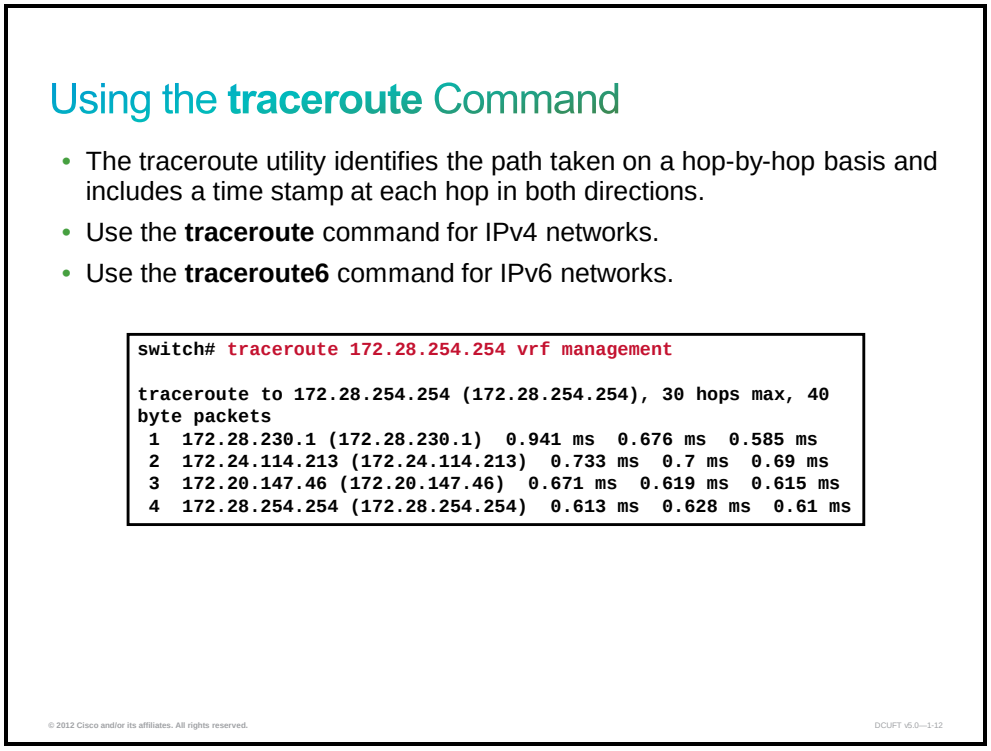
  `switch(config)#` **`pong destination-swid 3506 destination-mac 001b.54c2.9a43 vlan 1 count 3`**

- This example shows the **pong** service using virtual device context (VDC) MACs:

  `switch(config)#` **`pong source 001b.54c2.9a43 destination 001b.54c2.9a42 vlan 1 count 3`**

- This example shows the **pong** service using static MAC for injection:

  `switch(config)#` **`pong source 1.2.3 destination 001b.54c2.9a43 vlan 1 count 1 interface ethernet 4/29 inject`**



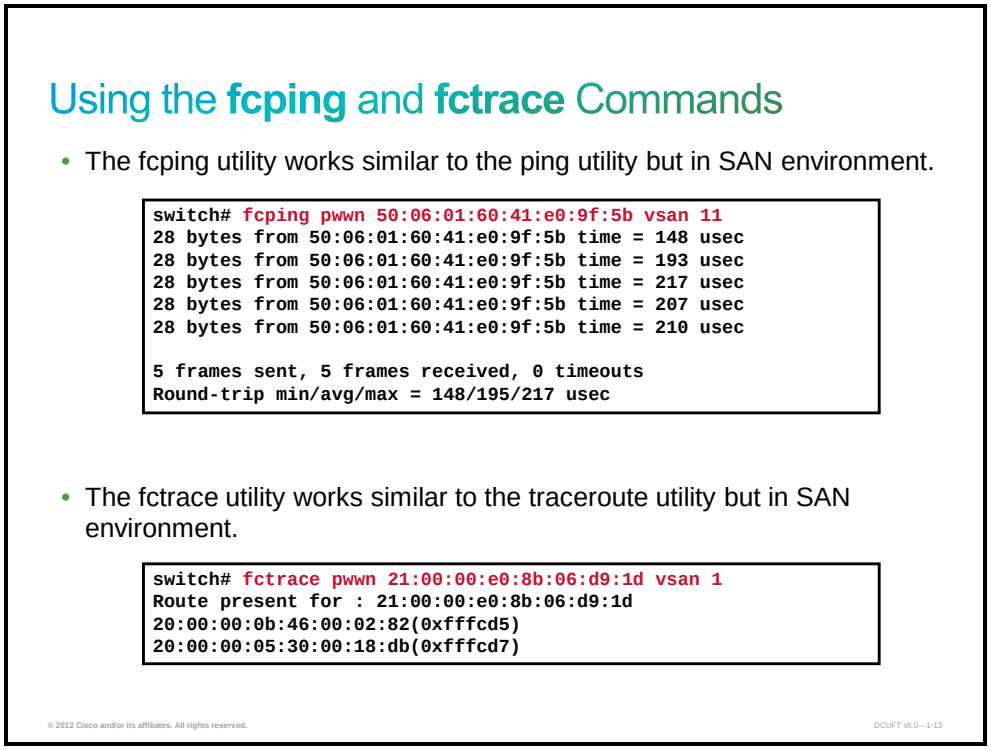
The **traceroute** command is used to do the following:

- Trace the route that is followed by the data traffic
- Compute the interswitch (hop-to-hop) latency

The **traceroute** command identifies the path that is taken on a hop-by-hop basis and includes a time stamp at each hop in both directions. The **traceroute** command can be used to test the connectivity of ports along the path between the generating device and the device closest to the destination.

For IPv4 networks, the **traceroute** command is used. For IPv6 networks, the **traceroute6** command is used. If the destination cannot be reached, the path discovery traces the path up to the point of failure.

To terminate a running **traceroute** command, press **Ctrl-C**.



The **fcping** command verifies the reachability of a node by checking its end-to-end connectivity. You can invoke the **fcping** feature by providing the Fibre Channel ID (FCID), the destination port world wide name (pWWN), or the device alias information.

The *device-alias* option specifies the device alias name. The *fcid* option specifies the FCID of the destination N port, with the format 0xhhhhhh. The *domain-controller-id* option verifies connection to the destination switch. The *pwwn* option specifies the pWWN of the destination N port, with the format hh:hh:hh:hh:hh:hh:hh:hh. The *vsan* option specifies a VSAN_ID.

The **fctrace** command provides the following capabilities:

- Trace the route followed by data traffic
- Compute interswitch (hop-to-hop) latency

You can invoke **fctrace** by providing the FCID, the N pWWN, or the device alias of the destination.

The trace frame is routed normally through the network until it reaches the far edge of the fabric. When the frame reaches the edge of the fabric (the F port connected to the end node with the given pWWN or FCID), the frame is looped back (swapping the source ID and the destination ID) to the originator.

If the destination cannot be reached, the path discovery starts, which traces the path up to the point of failure.

Monitor System, Processes, and CPU

This topic explains how to use the CLI to monitor the system, processes, and the CPU.

Monitor System, Processes, and CPU

- Use the **show processes cpu history** command to display information about CPU utilization by the system processes in the last 60 seconds, 60 minutes, and 72 hours in a graphical format.

```
switch# show processes cpu history

<…>
    11151422111111121111211212112212252212222122231312212223212 2
    86918030999399406989095088970671522791120531128042392533591 4
100
 90
 80
 70
 60
 50    *                         *
 40    * *                       *
 30    * *                 *   *  ** *        * *     * **
 20 ********** ** ********************************* **********
 10 ***#*****************************#************************
    0....5....1....1....2....2....3....3....4....4....5....5....
              0    5    0    5    0    5    0    5    0    5

             CPU% per minute (last 60 minutes)
            * = maximum CPU%   # = average CPU%
<…>
```


There are three options when using the **show processes** command:

```
switch# show processes ?
cpu      Show processes CPU Info
log      Show information about process logs
memory   Show processes Memory Info
```

The **show processes cpu history** command is used to display information about CPU utilization by the system processes in the last 60 seconds, 60 minutes, and 72 hours in a graphical format. The command output includes the following:

```
switch(config)# show processes cpu history
                                2                       1       21
     1186 1    1 1  6 11 543451 1  1   12   1111115 1  114244 1
100
 90
 80
 70
 60
 50
 40
 30
 20                            #                       #       #
 10    ##           #    #  ##                        #       ##
   0....5....1....1....2....2....3....3....4....4....5....5....
             0    5    0    5    0    5    0    5    0    5

                CPU% per second (last 60 seconds)
                      # = average CPU%
```

```
    1115142211111112111121121211221225221222212223131221222321 22
    8691803099939940698909508897067152279112053112804239253359 14
100
 90
 80
 70
 60
 50    *                              *
 40    * *                            *
 30    * *                    *   *  ** *          * *     * **
 20 ********** ** ********************************* **********
 10 ***#*****************************#************************
    0....5....1....1....2....2....3....3....4....4....5....5....
              0    5    0    5    0    5    0    5    0    5

               CPU% per minute (last 60 minutes)
              * = maximum CPU%   # = average CPU%
<rest of the output omitted>
```

Note	You can specify another VDC only from the default VDC, when working with a Cisco Nexus 7000.

Using the **show processes cpu sort** Command

- Use the **show processes cpu** command to display information about CPU processes sorted by CPU utilization.

```
switch# show processes cpu sort

PID    Runtime(ms)  Invoked   uSecs  1Sec    Process
-----  -----------  --------  -----  ------  -----------
 3622         2335      6843    341    50.0% pfstat
    1         2550      4169    611    0.0%  init
    2           13      2676      4    0.0%  migration/0
    3         2091    883525      2    0.0%  ksoftirqd/0
    4           48      6300      7    0.0%  desched/0
    5           10      2816      3    0.0%  migration/1
    6           21   4450597      0    0.0%  ksoftirqd/1
    7           42      6416      6    0.0%  desched/1
    8         1785      8581    208    0.0%  events/0
    9         1560      7426    210    0.0%  events/1
   10           58      2731     21    0.0%  khelper
   15            0        30     25    0.0%  kthread
   24            0         2      5    0.0%  kacpid
  104           12       201     62    0.0%  kblockd/0
  105            4       138     33    0.0%  kblockd/1
  118            0         5     17    0.0%  khubd
  185            0         4      3    0.0%  pdflush
  186          139      3010     46    0.0%  pdflush
  187            0         1      5    0.0%  kswapd0
<…>
```


The **show processes cpu sort** command is used to display information about CPU processes sorted by CPU utilization.

Some high CPU utilization (for instance, 80 percent) is not always a concern. High CPU utilization is a concern only when it occurs for specific processes that are aligned with a particular problem or complaint.

Using the **show system resource** Command

- Use the **show system resources** command to display system-related CPU and memory statistics.

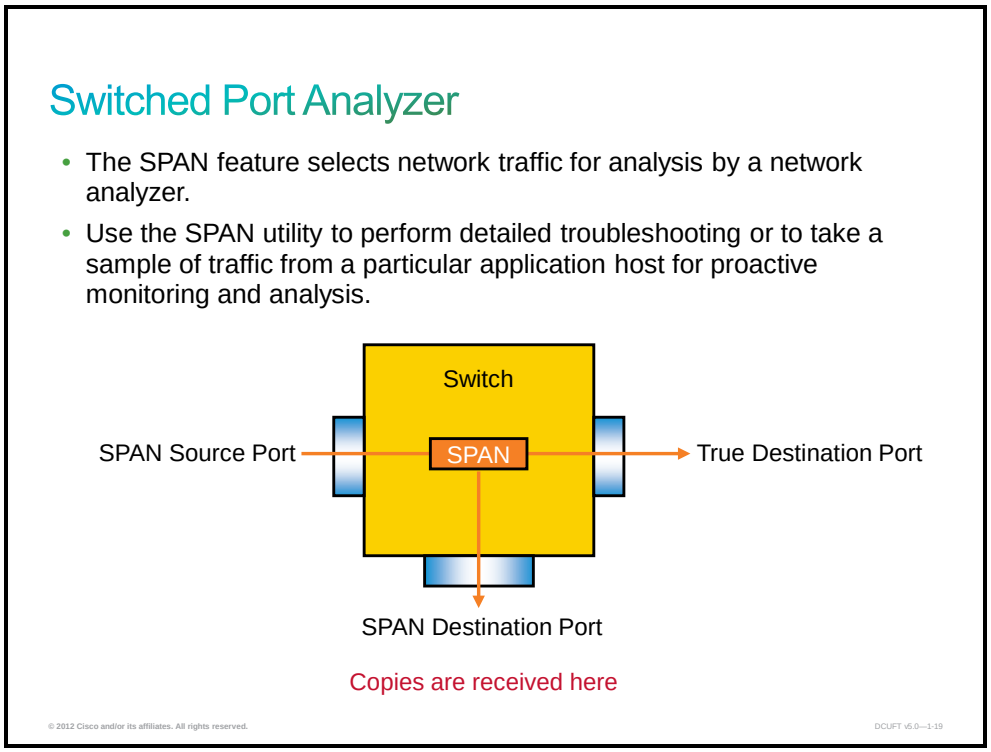
```
switch# show system resources

Load average: 1 minute: 0.30 5 minutes: 0.34 15 minutes: 0.28
Processes : 606 total, 2 running
CPU states : 0.0% user, 0.0% kernel, 100.0% idle
Memory usage: 2063268K total, 1725944K used, 337324K free
2420K buffers, 857644K cache
```


The **show system resources** command is used to display system-related CPU and memory statistics. The output includes the following:

- The load field is defined as the number of running processes. The average reflects the system load over the past 1, 5, and 15 minutes.
- The processes field displays the number of processes in the system, and how many are actually running when the command is issued.
- The CPU states field shows the CPU usage percentage in user mode, kernel mode, and idle time in the last 1 second.
- The memory usage field provides the total memory, used memory, free memory, memory that is used for buffers, and memory that is used for cache in kilobytes. Buffers and the cache are also included in the used memory statistics.

Switched Port Analyzer

This topic describes how to use the CLI Switched Port Analyzer (SPAN) utility to perform detailed troubleshooting from a particular application host for proactive monitoring and analysis.

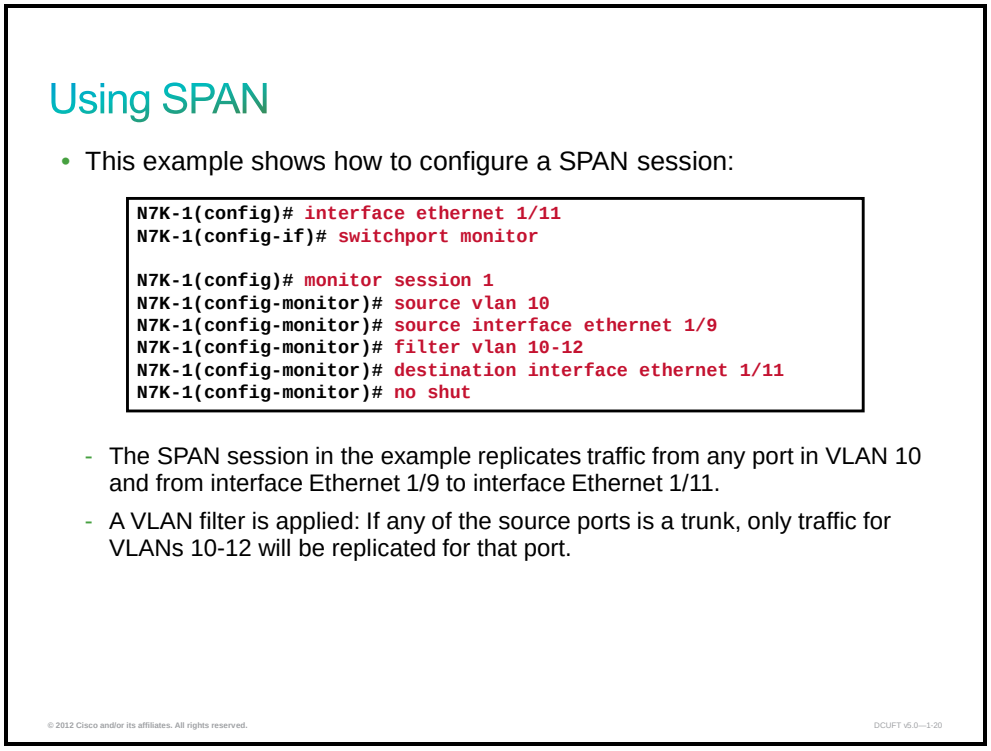


The SPAN feature—sometimes called port mirroring or port monitoring—selects network traffic for analysis by a network analyzer. The network analyzer can be a Cisco SwitchProbe, a Fibre Channel Analyzer, or other Remote Monitoring (RMON) probes.

The SPAN feature that is available on the Cisco Nexus switches allows you to replicate traffic from one or more switch ports or VLANs to another port on the same switch. This feature allows you to connect a system that has packet-sniffing software that is installed to a port on the switch. That port will then receive a copy of the traffic from the selected ports or VLANs, allowing it to be captured and analyzed.

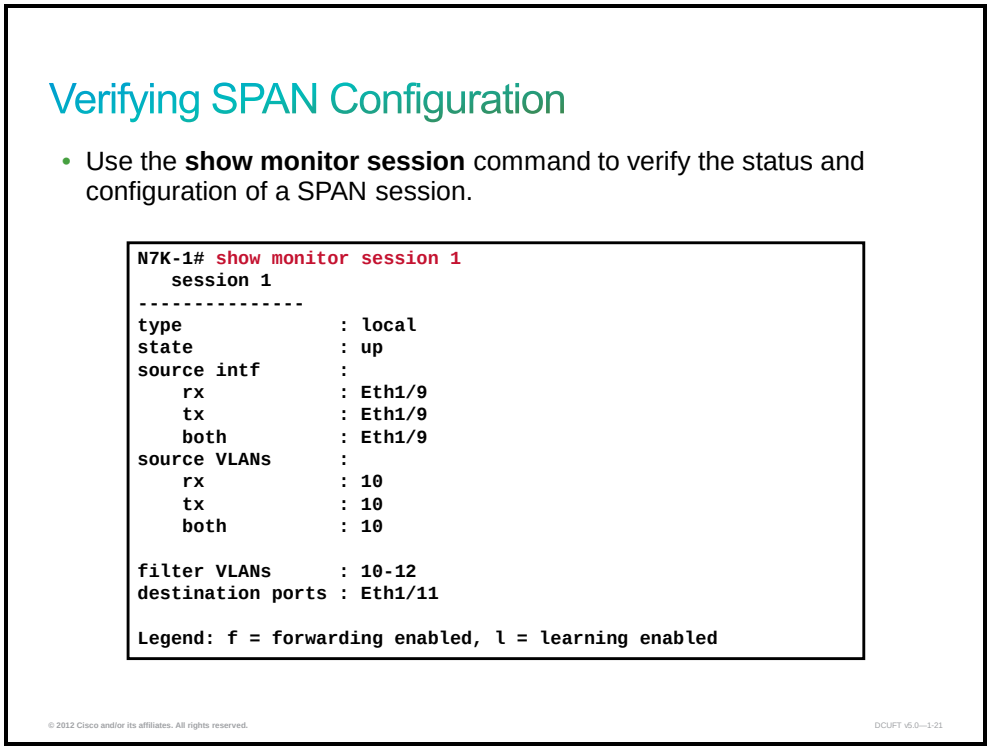
The SPAN sessions are nondisruptive to the devices that are monitored, and the traffic is replicated in hardware to prevent it from affecting the supervisor CPU.

In order to extend SPAN, Remote SPAN (RSPAN) enables remote monitoring of multiple switches across your network. RSPAN is supported on Cisco MDS switches. The Cisco Nexus 7000 and 5000 Series Switches are limited in what they can do with RSPAN. An RSPAN VLAN cannot be used as a SPAN destination. Therefore, the Nexus 7000 switch can only use RSPAN as a transit VLAN or source VLAN and pull data from the RSPAN VLAN. You cannot place anything into the RSPAN VLAN from the Nexus 7000 switch.



The example in the figure shows how to configure a SPAN session. First, the port that is going to be used as the destination port is specifically marked as a SPAN destination port by using the **switchport monitor** command. Next, the monitor session is configured. In this example, both a VLAN and a specific port are configured as SPAN sources. All traffic on any port in VLAN 10 and all traffic on interface ethernet 1/9 will be monitored and replicated to the destination interface ethernet 1/11. A VLAN filter is also applied. A VLAN filter is primarily useful when you are monitoring a trunk. By default, this would replicate traffic from any VLAN on the trunk to the destination port. The VLAN filter limits this to a subset of the VLANs. In this example, only traffic in VLANs 10-12 will be replicated.

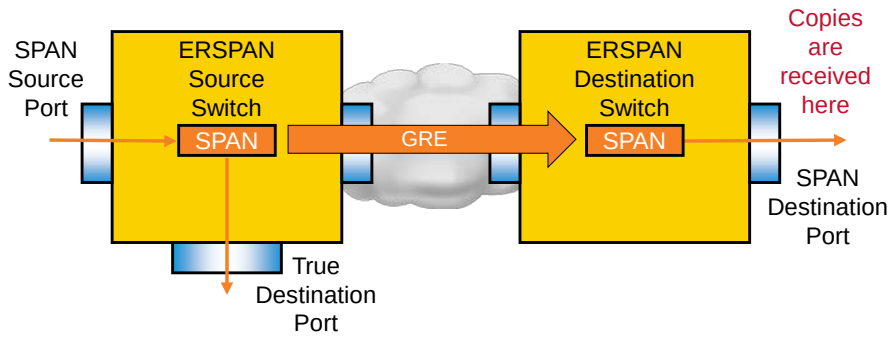
Finally, a **no shut** command is applied to activate the SPAN session.



The **show monitor session** command provides an overview of all the relevant parameters for a SPAN session and its operational state. If the session state does not show as "up," this **show** command displays an indication of the reason that the session is down.

Encapsulated Remote SPAN

- The ERSPAN feature allows traffic from one or more source ports or source VLANs to be encapsulated in GRE packets and sent to another switch to be forwarded to one or more destination ports for capture and analysis.



A major limitation in the use of the SPAN feature as a troubleshooting tool is that the source and destination port must be in the same switch and VDC if using a Cisco Nexus 7000 switch. Encapsulated Remote SPAN (ERSPAN) allows SPAN traffic to be transported across an IP network. The traffic is encapsulated at the source Cisco Nexus 7000 or 5000 switch and is transferred across the IP network using Generic Routing Encapsulation (GRE) packets. The packet is de-encapsulated at the destination switch and then sent to the destination interface.

It is also possible to use an end station as the destination for an ERSPAN session, if the traffic analysis software on the end station supports the ERSPAN format.

Using ERSPAN

- The IP source address to be used for the ERSPAN GRE packets must be configured in the default VDC.

```
N7K-1(config)# monitor erspan origin ip-address 192.168.0.210 global
```

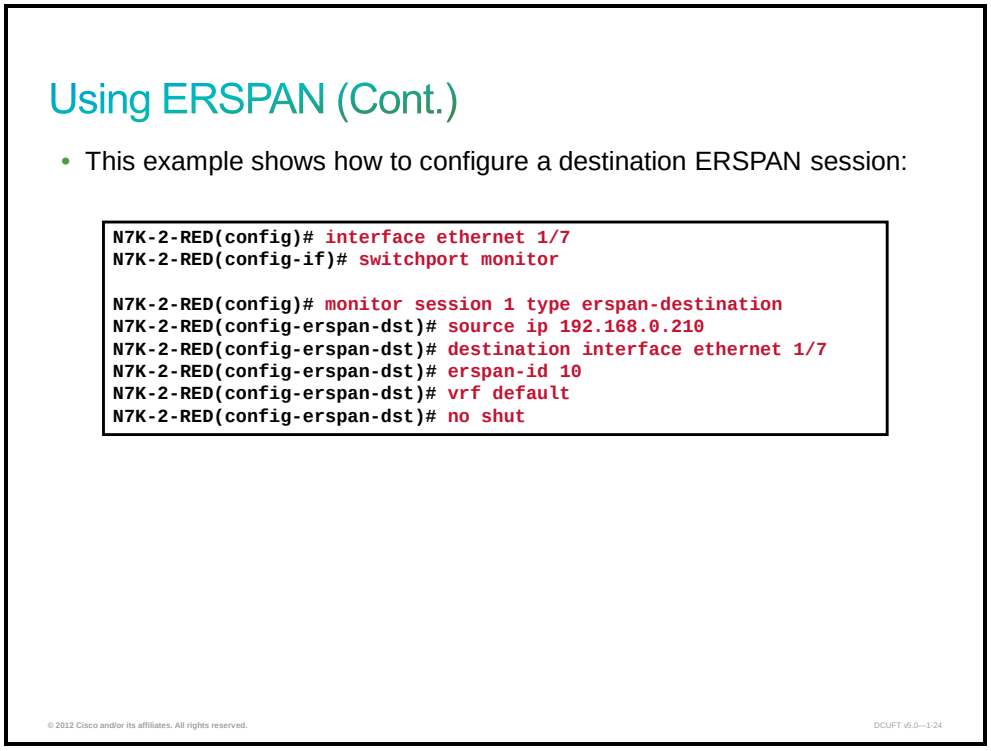
- This example shows how to configure a source ERSPAN session:

```
N7K-1-RED(config)# ip access-list CAPTURE-LIST
N7K-1-RED(config-acl)# permit ip 11.0.0.0 0.255.255.255 any
N7K-1-RED(config-acl)# exit
N7K-1-RED(config)# monitor session 1 type erspan-source
N7K-1-RED(config-erspan-src)# source interface ethernet 1/21
N7K-1-RED(config-erspan-src)# source vlan 10
N7K-1-RED(config-erspan-src)# destination ip 172.16.10.72
N7K-1-RED(config-erspan-src)# erspan-id 10
N7K-1-RED(config-erspan-src)# vrf default
N7K-1-RED(config-erspan-src)# no shutdown
N7K-1-RED(config-erspan-src)# filter access-group CAPTURE-LIST
```


The example in the figure shows how to configure an ERSPAN source session. Before any ERSPAN source session can become active on the switch, it is necessary to configure the source IP address to be used for the ERSPAN traffic in the default VDC by using the **monitor erspan origin ip-address** command.

The example shows how an ERSPAN source monitor session is created and how source interfaces and VLANs are associated with it, using the same syntax that is used for regular SPAN sessions. Then, the destination IP address, virtual routing and forwarding (VRF), and ERSPAN ID are configured. These parameters are all mandatory, and if any of these parameters is not set, the session cannot become active. To enable the session, the **no shut** command is required, similar to regular SPAN sessions.

To limit the information that is captured and transported across the network, an access list can be applied to the ERSPAN session using the **filter access-group** command.



The example in the figure shows how to configure an ERSPAN destination session. Similar to regular SPAN sessions, the intended destination port must first be configured as a SPAN destination port using the **switchport monitor** command.

In the monitor session configuration, the source IP address, ERSPAN ID, and VRF for the ERSPAN session are configured to identify the ERSPAN encapsulated traffic. The de-encapsulated traffic is then forwarded to the ports that are indicated as destination ports. Similar to regular SPAN sessions, it is necessary to issue the **no shut** command to activate the ERSPAN session.

Verifying ERSPAN Configuration

- Use the **show monitor session** command to verify the status and configuration of a source ERSPAN session.

```
N7K-1-pod1# show monitor session 1 brief
   session 1
---------------
type              : erspan-source
state             : up
erspan-id         : 10
vrf-name          : default
acl-name          : CAPTURE-LIST
ip-ttl            : 255
ip-dscp           : 0
destination-ip    : 172.16.10.72
origin-ip         : 192.168.0.210 (global)
source intf       :
    rx            : Eth1/21
    tx            : Eth1/21
    both          : Eth1/21
source VLANs      :
    rx            : 10
    tx            : 10
    both          : 10
```


The **show monitor session** command provides an overview of the relevant parameters for the ERSPAN source and destination sessions. When the session is active, the state of the session will be listed as “up.” If the session is not active, the reason will be listed here to indicate missing parameters or other misconfigurations. The example in the figure shows typical output for an active source ERSPAN session.

Verifying ERSPAN Configuration (Cont.)

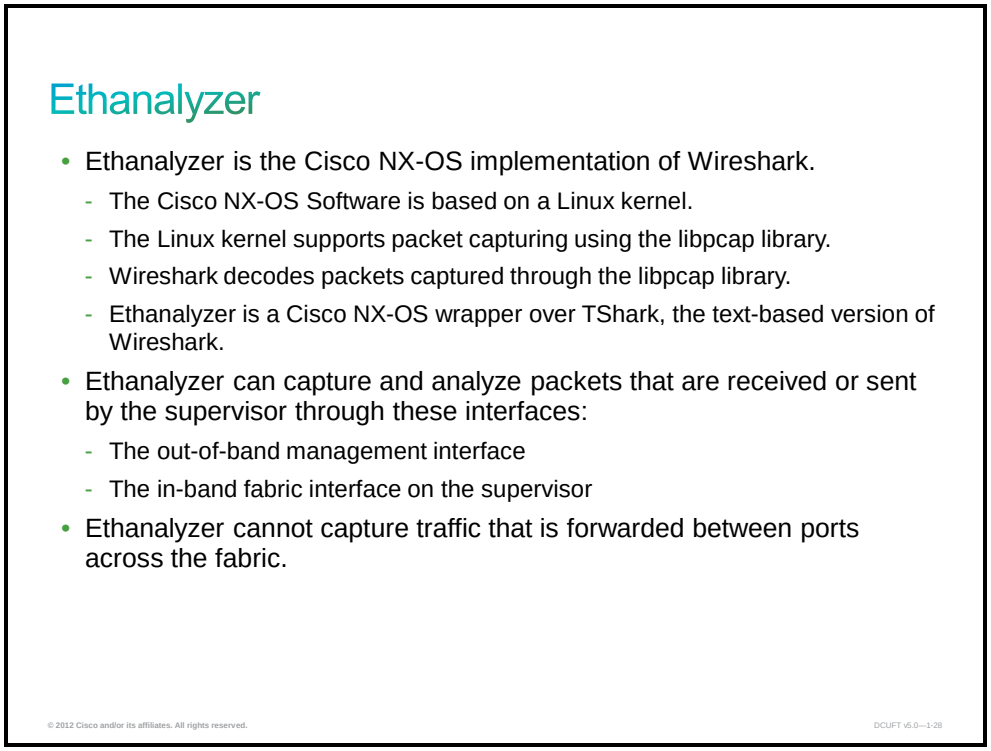
- Use the **show monitor session** command to verify the status and configuration of a destination ERSPAN session.

```
N7K-2-pod2# show monitor session 1
   session 1
---------------
type              : erspan-destination
state             : up
erspan-id         : 10
vrf-name          : default
source-ip         : 192.168.0.210
destination ports : Eth1/7

Legend: f = forwarding enabled, l = learning enabled
```


The example in this figure shows the typical output for the **show monitor session** command for a destination ERSPAN session.

Ethanalyzer

This topic describes how to use Ethanalyzer to troubleshoot your network and analyze control-plane traffic.



The Cisco NX-OS Software is a modern, modular operating system that runs a Linux kernel. The architecture makes it easy to embed common tools that are used by network administrators who are working in Linux-based environments. An example of this integration is support for an integrated packet analyzer for the network traffic that is destined to or generated by the Cisco Nexus 7000 or 5000 Series Switches or MDS supervisor.

Similar to the debug capability of the Cisco router product line, the Cisco MDS storage switches have a Fibre Channel analyzer to examine packets (via the **fcanalyzer local** command). The Fibre Channel analyzer examines packets to and from the entities that the switch provides. The Fibre Channel analyzer is able to debug frames that the switch is responsible for receiving or sending to a storage device.

Using the command-line version of Wireshark, called TShark, as a basis, Cisco developed the Cisco NX-OS Ethanalyzer. The Linux kernel supports the capturing of packets using the libpcap library and TShark provides additional packet filtering and analysis functionality. Ethanalyzer is a Cisco NX-OS wrapper over TShark that allows packet capturing and analysis from the Cisco NX-OS CLI.

Ethanalyzer can interactively analyze packets that are sent to or generated by the supervisor. More specifically, it can capture traffic received by the supervisor from both the out-of-band management port (mgmt0) and the in-band fabric interface that connects the supervisor to the I/O modules. Ethanalyzer can only capture traffic that is processed by the supervisor CPU.

However, the Cisco Nexus 7000 Series does provide a method for enabling Ethanalyzer to capture data traffic. In this mode of operation, Ethanalyzer gives network administrators a powerful, easy-to-use tool that increases visibility into application behavior and increases their ability to exert control over the network environment.

Using Ethanalyzer

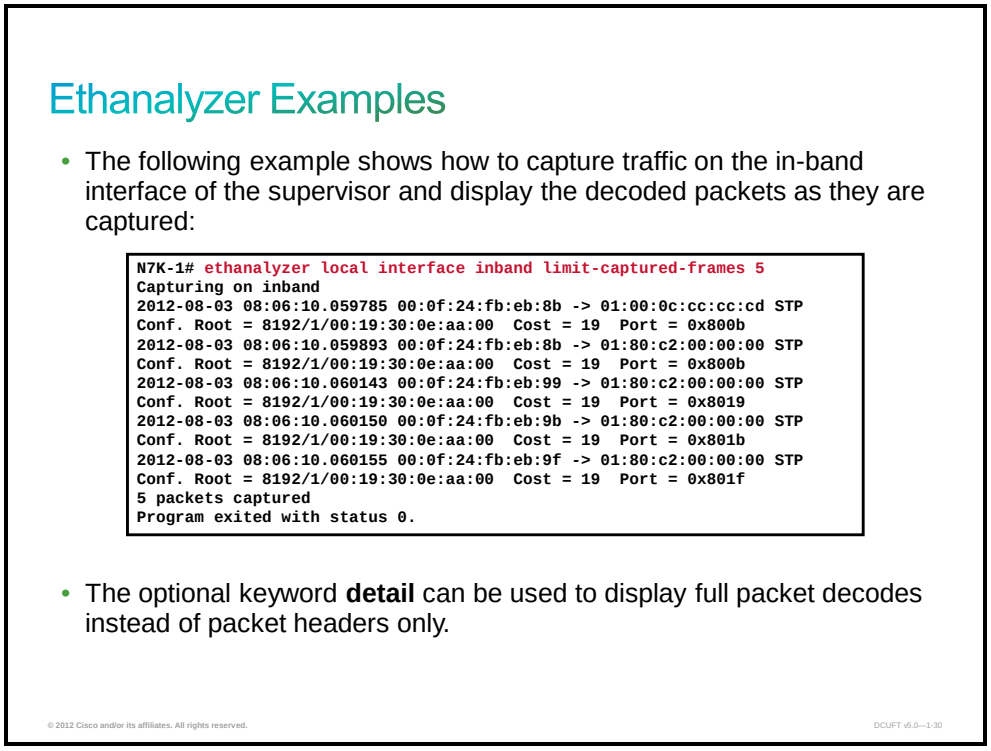
- Use Ethanalyzer to troubleshoot your network and analyze the control-plane traffic.
- Ethanalyzer can capture traffic from a local interface on the supervisor, or read from a file that contains a packet capture in pcap format.
- Ethanalyzer can display decoded packets on the terminal as they are captured or write the captured packets to a file in pcap format.
- On the Cisco Nexus 7000 switch, Ethanalyzer can only be used from the default VDC.
- TCP dump is used for capture filters and Wireshark is used for display filters.

Ethanalyzer captures and processes packets in the libpcap format, which is also used by many other packet capture and analysis tools, such as tcpdump. Ethanalyzer can capture packets live, or read from a file in pcap format that was captured earlier.

When you capture packets using Ethanalyzer, you have the choice to display the decoded packets as they are captured or to write the packets to a file in pcap format to analyze offline.

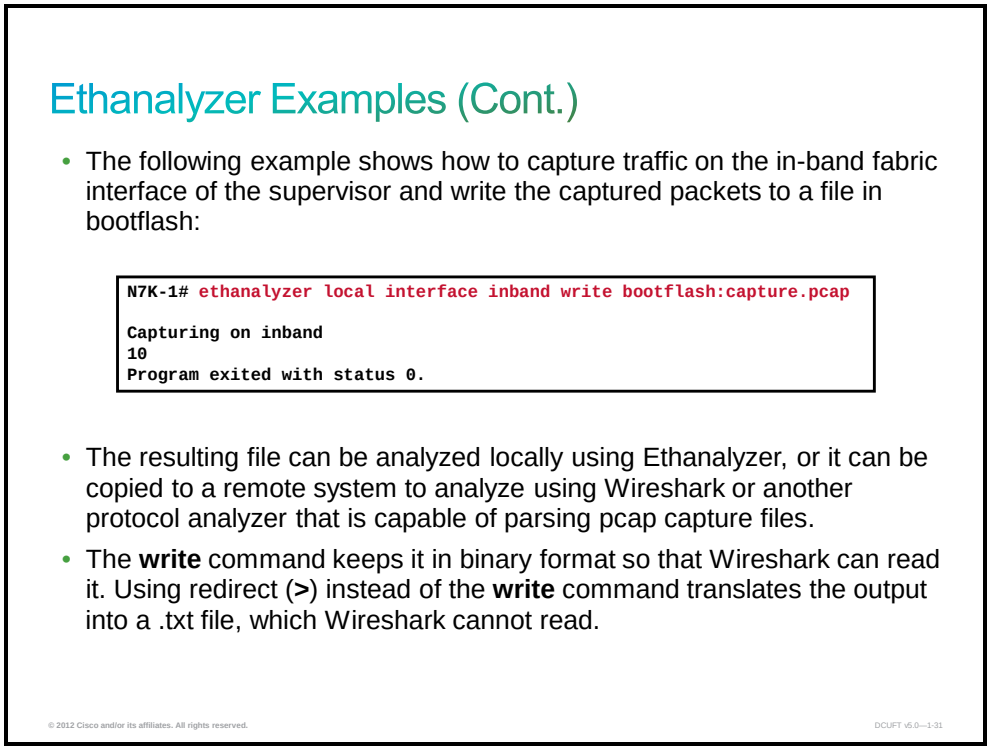
On the Cisco Nexus 7000 switches, Ethanalyzer is only available in the default VDC, as it allows control- and management-plane traffic from all VDCs to be captured. Because Ethanalyzer is based on Wireshark, it supports the same set of powerful capture and display filters that are available in Wireshark.

Note	For more detailed information about the filtering capabilities of Wireshark, TShark, TCP dump and Ethanalyzer, refer to the Wireshark documentation wiki at http://wiki.wireshark.org.



The example in the figure shows how to capture traffic from the in-band fabric interface of the supervisor using Ethanalyzer. By default, the number of captured frames is limited to 100 frames. This value can be changed using the **limit-captured-frames** command option.

By default, the **ethanalyzer** command only displays the packet headers, which can be useful if you only need to confirm that specific types of packets are sent or received by the switch. If you need to see the full packet decode, you should add the **detail** option to the **ethanalyzer** command.



The example in the figure shows how to capture traffic and write the captured packets to a file in pcap format. By default, the packets are not displayed as they are captured when you write the capture to a file. However, if you prefer to see the packets that are written to the file as they are captured, you can add the **display** option to the command.

The captured packets are written to the file in pcap format. This allows the capture file to be transferred to a remote system and analyzed using Wireshark or another network analyzer that supports the pcap format.

Using redirect (“>”) translates the output into a text file, which Wireshark cannot read. Here is an example of redirecting a detailed capture to a file:

```
N7000# ethanalyzer local interface detail > cpu-1.txt
```

Ethanalyzer Examples (Cont.)

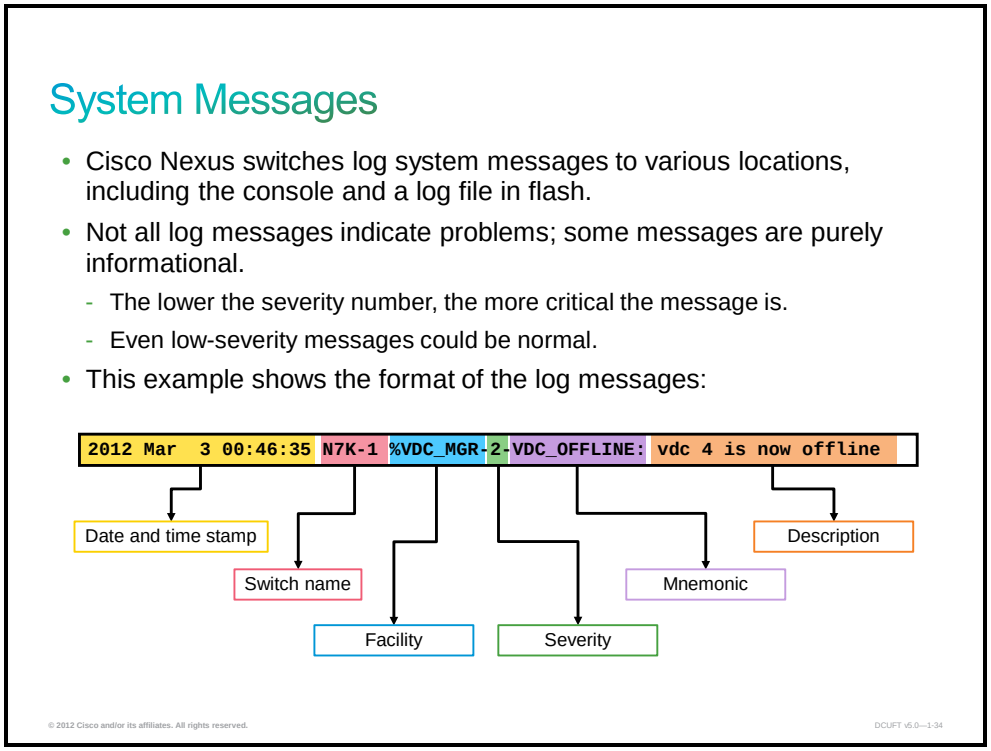
- The following example shows how to display the content of a previously captured pcap file:

```
N7K-1# ethanalyzer local read bootflash:capture.pcap detail
Frame 1 (183 bytes on wire, 151 bytes captured)
    Arrival Time: Mar 30, 2012 13:07:06.616542000
    [Time delta from previous captured frame: 0.000000000 seconds]
    [Time delta from previous displayed frame: 0.000000000 seconds]
    [Time since reference or first frame: 0.000000000 seconds]
    Frame Number: 1
    Frame Length: 183 bytes
    Capture Length: 151 bytes
    [Frame is marked: False]
    [Protocols in frame: eth:llc:stp]
IEEE 802.3 Ethernet
    Destination: 01:80:c2:00:00:00 (01:80:c2:00:00:00)
        Address: 01:80:c2:00:00:00 (01:80:c2:00:00:00)
        .... ...1 .... .... .... .... = IG bit: Group address (multicast/broadcast)
        .... ..0. .... .... .... .... = LG bit: Globally unique address (factory
default)
    Source: c8:4c:75:fc:66:8b (c8:4c:75:fc:66:8b)
        Address: c8:4c:75:fc:66:8b (c8:4c:75:fc:66:8b)
        .... ...0 .... .... .... .... = IG bit: Individual address (unicast)
        .... ..0. .... .... .... .... = LG bit: Globally unique address (factory
default)
    Length: 137
<…further output omitted…>
```


Instead of transferring the file for offline analysis, it is also possible to examine the contents of a pcap file locally using Ethanalyzer. The example in the figure shows how to display the contents of a capture file in bootflash. Similar to the live capturing of packets using Ethanalyzer, the **detail** option is required to view the full packet decode instead of only the packet headers.

Logging

This topic describes how to use the logging feature to log information for monitoring and troubleshooting of the Cisco Nexus or MDS switches.



The system software sends system messages to the console and to a log file in flash. Optionally, these log messages can also be sent to a server using syslog. Not all system messages indicate a problem with your device. Some messages are purely informational, while others might help diagnose problems with links, internal hardware, or the device software.

Each message is logged at a specific severity level. The severity levels that are used in system messages include the following:

- **0 – emergency:** System unusable
- **1 – alert:** Immediate action needed
- **2 – critical:** Critical condition
- **3 – error:** Error condition
- **4 – warning:** Warning condition
- **5 – notification:** Normal but significant conditions
- **6 – informational:** Information message only
- **7 – debugging:** Appears during debugging only

In general, a lower severity number indicates a higher criticality. However, even low-severity messages could indicate normal events, such as a new power supply being detected after insertion.

The log messages follow a specific format and include the following fields:

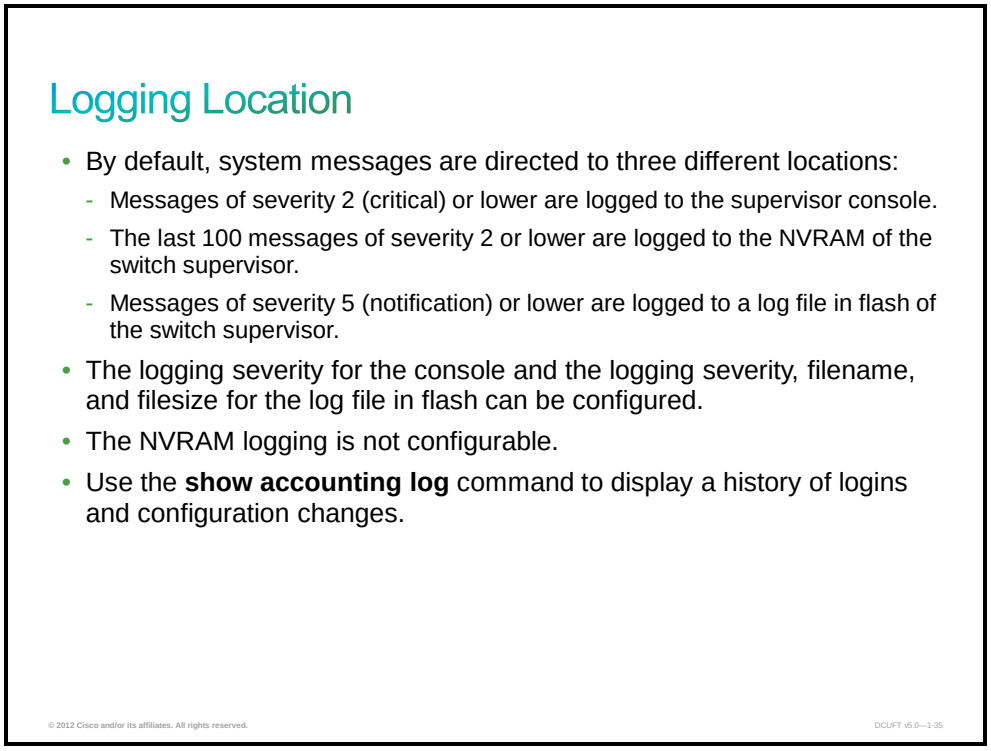
- A month and time stamp that indicates the local switch time when the message was logged
- The hostname of the switch

- The facility, which indicates the Cisco NX-OS Software subsystem or Cisco Nexus hardware component that generated the message

Note	This is not to be confused with the facility value defined in the syslog protocol. The syslog facility is set to local7 by default on Cisco Nexus switches, but can be changed.

- The severity, which indicates the criticality of the message
- A mnemonic, which is a text string that uniquely describes the system message
- A description that describes the event in detail

Note	The mnemonics can be used to find a detailed description of the message in the Cisco NX-OS System Messages Reference at http://www.cisco.com/en/US/docs/switches/datacenter/sw/system_messages/reference/sl_nxos_book.html.



By default, the Cisco NX-OS Software sends log messages to three locations:

- **Console:** Messages of severity 2 (critical) or lower are logged to the console by default. The logging severity level for the console can be changed, but to be able to increase the logging level to a value higher than critical, you must change the console speed to 38,400 b/s.
- **NVRAM:** The device logs the most recent 100 messages of severity 2 or lower to a log in the NVRAM. You cannot configure logging to the NVRAM.
- **Log file:** By default, the system logs all messages of severity 5 (notification) or lower to a logfile in flash. The log filename, maximum file size, and the maximum severity of the messages that are logged are configurable.

Use the **show accounting log** command to display a history of logins and configuration changes.

The following output shows possible configuration of logging (you can see that the configuration of NVRAM is not possible):

```
N7K1-POD1(config)# logging ?
  console           Set console logging
  event             Interface events
  ip                IP configuration
  level             Facility parameter for syslog messages
  logfile           Set File logging
  message           Interface events
  module            Set module(linecard) logging
  monitor           Set terminal line(monitor) logging level
  server            Enable forwarding to Remote Syslog Server
  source-interface  Enable Source-Interface for Remote Syslog Server
  timestamp         Set logging timestamp granularity
```

Viewing the Logs

- To view the log messages in the NVRAM, use the **show logging nvram** command:

```
N7K-1-RED# show logging nvram last 3
2012 Mar 3 23:29:38 N7K-1-RED %$ VDC-4 %$ %SYSMGR-2-SERVICE_CRASHED:
 Service "orib" (PID 1149) hasn't caught signal 11 (core will be saved).
2012 Mar 2 00:47:09 N7K-1-RED %$ VDC-4 %$ %FEX-2-NOHMS_ENV_FEX_ONLINE:
 FEX-105 On-line (Serial Number JAF1420AHPE)
2012 Feb 8 23:24:36 N7K-1-RED %$ VDC-4 %$ %FEATURE-MGR-2-
 FM_AUTOCKPT_SUCCEEDED: AutoCheckpoint created successfully
```

- To view the log messages in the log file in flash, use the **show logging logfile** command:

```
N7K-1# show logging logfile start-time 2012 Mar 2 00:00:00
2012 Mar  2 00:46:35 N7K-1 %VDC_MGR-2-VDC_OFFLINE: vdc 4 is now offline
2012 Mar  2 00:46:58 N7K-1 %IM-5-IM_INTF_STATE: mgmt0 is DOWN in vdc 4
2012 Mar  2 00:46:58 N7K-1 %IM-5-IM_INTF_STATE: mgmt0 is UP in vdc 4
2012 Mar  2 00:47:04 N7K-1 %VDC_MGR-2-VDC_ONLINE: vdc 4 has come online
2012 Mar  3 09:53:37 N7K-1 %AUTHPRIV-3-SYSTEM_MSG: pam_aaa:Authentication
 failed for user admin from 192.168.0.11 - login
```


To view the log messages in the NVRAM of the switch, use the **show logging nvram** [**last** *number-lines*] command. The **last** option can be used to specify that only the last number of lines in the log should be displayed.

To view the log messages in the logfile in flash, use the **show logging logfile** [**start-time** *yyyy mmm dd hh:mm:ss*] [**end-time** *yyyy mmm dd hh:mm:ss*] command. The **start-time** and **end-time** parameters allow you to specify a time range for the log messages that you are interested in.

To configure logging, use the following configuration commands:

Command	Description
switch(config)# **logging logfile** *logfile-name severity-level* [**size** *bytes*]	▪ Configures the name of the log file used to store system messages and the minimum severity level to log. You can optionally specify a maximum file size.
▪ switch(config)# logging module [*severity-level*]	▪ Enables module log messages that have the specified severity level or higher. If the severity level is not specified, the default of 5 is used.
▪ switch(config)# logging timestamp {microseconds | milliseconds | seconds}	▪ Sets the logging time-stamp units. By default, the units are seconds.
▪ switch(config)# logging server *host* [*severity-level* [use-vrf *vrf-name* [facility *facility*]]]	▪ Configures a host to receive syslog messages.

Onboard Failure Logging

- Cisco NX-OS Software includes the capability to store failure and environmental data in persistent storage.
  - The information can be used to analyze problems caused by hardware or software failures.
  - OBFL is enabled by default.
- This example shows how to display OBFL information for analysis or to provide to Cisco TAC:

```
N7K-1# show logging onboard
<…output omitted…>
---------------------------
    Module:  1
---------------------------

Exception Log Record :  Thu Mar 15 20:24:00 2012 (202862 us)

Device Id         : 112
Device Name       : Ashburton
Device Error Code : c7004201(H)
Device Error Type : ERR_TYPE_HW
Device Error Name : NULL
<…further output omitted…>
```


Cisco NX-OS Software allows you to log failure data to persistent storage, which you can retrieve and display for analysis at a later time. The onboard failure logging (OBFL) feature stores failure and environmental information in nonvolatile memory on the module. The information will help analyze failed modules. OBFL is enabled by default for all modules.

The data that are stored by OBFL include the following:

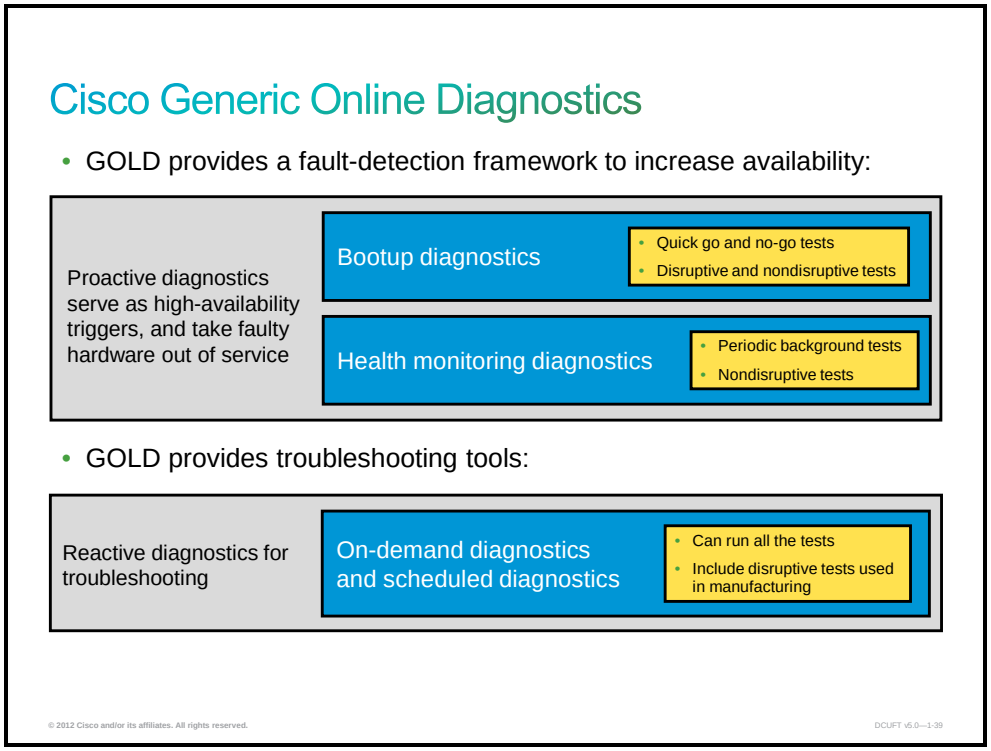
- Time of initial power-on
- Slot number of the module in the chassis
- Initial temperature of the module
- Firmware, BIOS, field-programmable gate array (FPGA), and ASIC versions
- Serial number of the module
- Stack trace for crashes
- CPU hog information
- Memory leak information
- Software error messages
- Hardware exception logs
- Environmental history
- OBFL-specific history information
- ASIC interrupt and error statistics history
- ASIC register dumps

OBFL stores a kernel trace in case Cisco NX-OS Software crashes.

Note	To interpret the data in the OBFL logs, specialist knowledge of the platform is required. The information in the OBFL logs should be presented to the Cisco Technical Assistance Center (TAC) for analysis as part of the troubleshooting process.

Cisco Generic Online Diagnostics

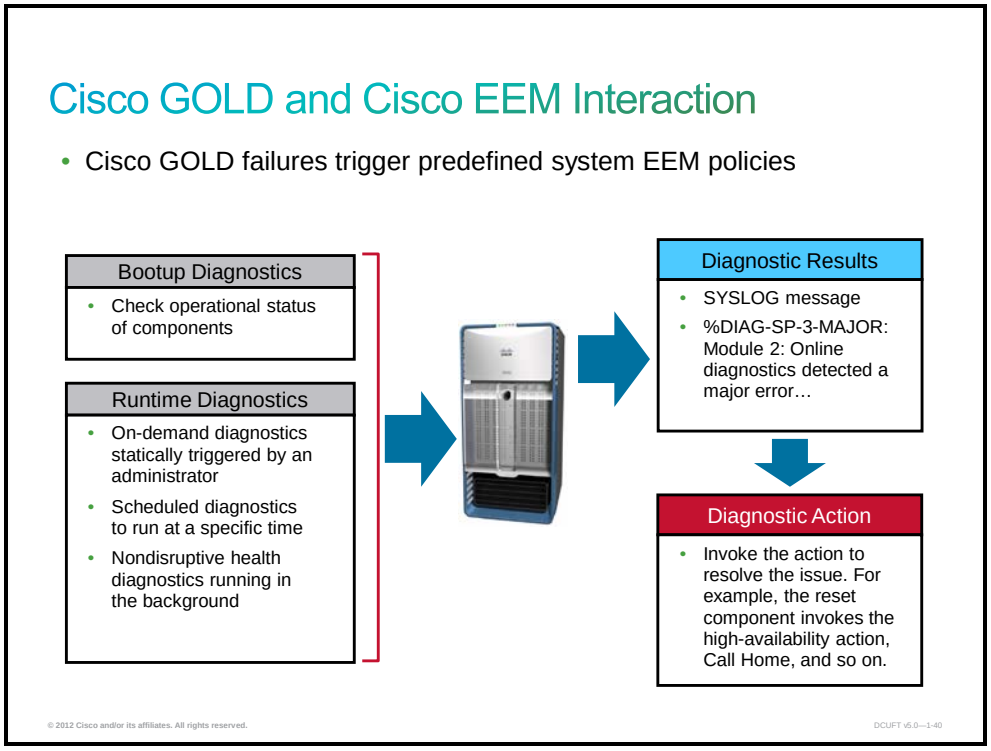
This topic describes how to use Cisco Generic Online Diagnostics (GOLD) to collect diagnostic results and detailed statistics on the Cisco Nexus or MDS switches.



Cisco GOLD is a suite of diagnostics that verifies that the hardware and internal data paths are operating as designed. Boot-time diagnostics, continuous monitoring, on-demand, and scheduled tests are part of the Cisco GOLD feature set. The diagnostics subsystem allows rapid fault isolation and continuous system monitoring.

The diagnostic framework can detect hardware failures while the system is online and operational, and corrective actions are taken through Cisco Embedded Event Manager (EEM) policies.

Note Each module type has a different set of tests.



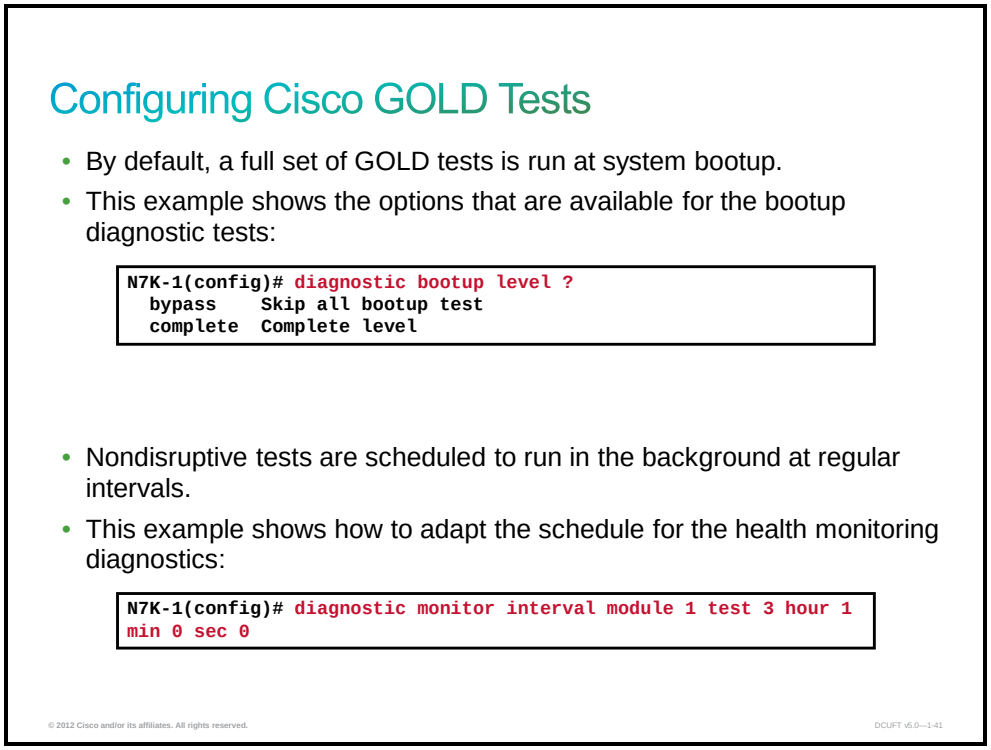
Cisco GOLD uses the Cisco EEM framework to react to failures. The Cisco Nexus 7000 Series Switches ship with a set of system default Cisco EEM policies that include the Cisco GOLD default Cisco EEM. The common default actions for all Cisco GOLD EEM policies include the following:

- Disable the test.
- Trigger the corrective action. For example, place the ports in the error-disabled state or reload the module.

A separate Cisco GOLD failure threshold is used for triggering syslog and Smart Call Home. Both thresholds and actions can be overridden with user-configured policies.

The system default Cisco EEM policies can be examined using the **show event manager system-policy** command. The following example shows the default system policy that is invoked if the RewriteEngineLoopback test fails:

```
N7K-1# show event manager system-policy __RewriteEngineLoopback
       Name : __RewriteEngineLoopback
Description : Do CallHome, log error and disable further HM testing
on affected ports after 10 consecutive failures of GOLD
"RewriteEngineLoopback" test
Overridable : Yes
```

Bootup diagnostics are run during bootup and detect faulty hardware before Cisco NX-OS Software brings a module online. For example, if you insert a faulty module in the device, bootup diagnostics test the module and take it offline before the device uses the module to forward traffic. Bootup diagnostics also check the connectivity between the supervisor and module hardware and the data and control paths for all the ASICs. Bootup diagnostics log failures to the OBFL and syslog and trigger a diagnostic LED indication. You can configure the Cisco Nexus switches to either bypass the bootup diagnostics or run the complete set of bootup diagnostics.

Run-time diagnostics are also called health-monitoring diagnostics. These diagnostics provide information about the health of a live device. They detect run-time hardware errors, memory errors, the degradation of hardware modules over time, software faults, and resource exhaustion. Health monitoring diagnostics are nondisruptive and run in the background to ensure the health of a device that is processing live network traffic. You can enable or disable health-monitoring tests or change their schedule.

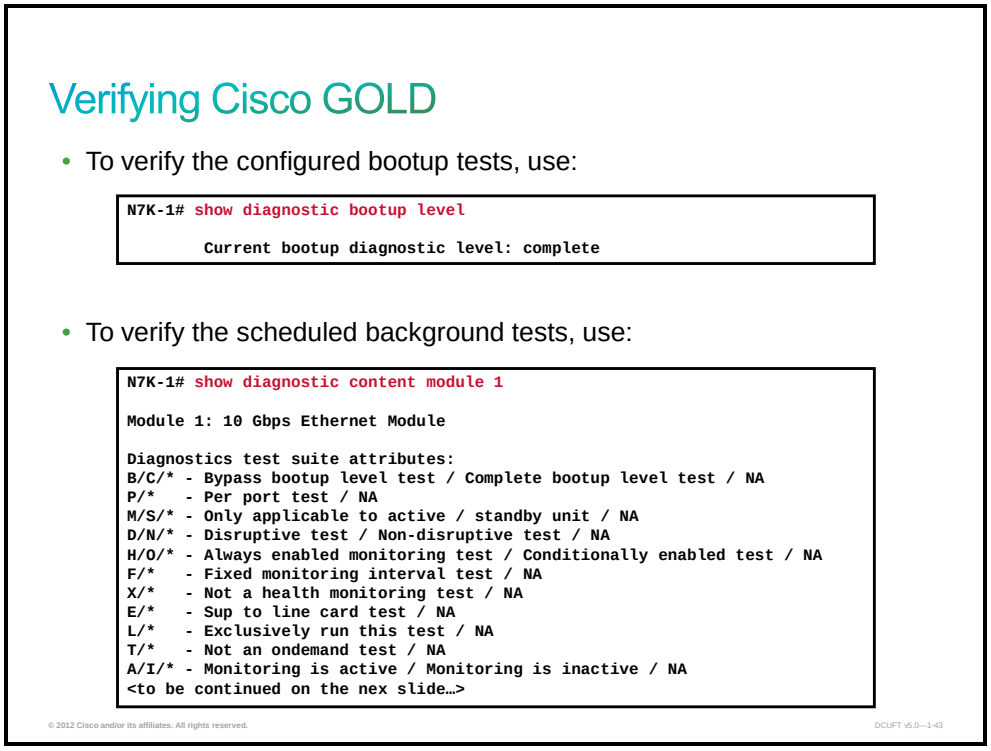
Running On-Demand GOLD Tests

- For troubleshooting purposes, specific GOLD tests can be run on demand.
- Before running the test, you can specify:
  - The number of iterations (default is 1)
  - The action to take when a test fails, which can be to stop the test immediately or continue until a specified number of failures have occurred (default is to stop after one failure has occurred)
- This example shows how to schedule an on-demand test:

```
N7K-1# diagnostic ondemand iteration 20
N7K-1# diagnostic ondemand action-on-failure continue failure-count 5
N7K-1# diagnostic start module 1 test 7 port 9
```


On-demand tests help to test specific hardware components to locate faults. You can schedule on-demand diagnostics to run immediately. Alternatively, you can modify the default interval for the corresponding health-monitoring test.

You can start or stop an on-demand diagnostic test. You can optionally modify the number of iterations to repeat the specified test, and the action to take if the test fails.



The figure shows how to display the currently configured bootup test level and the current schedule for the health-monitoring tests. The flags in the **show diagnostic content** command indicate the characteristics of the tests. For example, the “P” flag in the second line indicates that the test is a per-port test and the “D” or “N” flag in the fourth line indicates whether the test is disruptive or nondisruptive.

Verifying Cisco GOLD (Cont.)

- To verify the scheduled background tests, use:

```
N7K-1# show diagnostic content module 1
<…continuation>
                                                    Testing Interval
ID     Name                                Attributes     (hh:mm:ss)
____   _____________________________       ___________    _____________
 1)    ASICRegisterCheck------------>      ***N******A     00:01:00
 2)    PrimaryBootROM---------------->     ***N******A     00:30:00
 3)    SecondaryBootROM-------------->     ***N******A     00:30:00
 4)    EOBCPortLoopback-------------->     C**N**X**T*     -NA-
 5)    OBFL-------------------------->     C**N**X**T*     -NA-
 6)    PortLoopback------------------>     *P*N***E**A     00:15:00
 7)    RewriteEngineLoopback--------->     *P*N***E**A     00:01:00
 8)    SnakeLoopback----------------->     *P*N***E**A     00:20:00
 9)    FIPS-------------------------->     *P*NO*XE*T*     -NA-
10)    BootupPortLoopback------------>     CP*N**XE*T*     -NA-
```


Note For a description of the tests, refer to the Cisco Nexus 7000 Series NX-OS System Management Configuration Guide at http://www.cisco.com/en/US/docs/switches/datacenter/sw/6_x/nx-os/system_management/configuration/guide/sm_11gold.html#wp1105098.

Verifying Cisco GOLD (Cont.)

- To examine the results of a Cisco GOLD test, use:

```
N7K-1# show diagnostic result module 1 test 7

Current bootup diagnostic level: complete
Module 1: 10 Gbps Ethernet Module

        Test results: (. = Pass, F = Fail, I = Incomplete,
        U = Untested, A = Abort, E = Error disabled)

         7) RewriteEngineLoopback:

          Port   1  2  3  4  5  6  7  8  9 10 11 12 13 14 15 16
          ------------------------------------------------
                 .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .

          Port  17 18 19 20 21 22 23 24 25 26 27 28 29 30 31 32
          ------------------------------------------------
                 .  .  .  .  .  .  .  .  .  .  .  .  .  .  .  .
```

- This command lists results for all types of tests: bootup, scheduled, or on-demand.

This figure shows how to verify the results of a diagnostic test. The **show diagnostic result** command shows the results of all types of tests: bootup tests, scheduled health-monitoring tests, and on-demand tests.

Blink or Beacon

This topic describes how to use the blue beacon to aid in troubleshooting or replacement of components in the Cisco Nexus switch.

Blue Beacon Feature

- On some platforms, you can cause the platform LEDs to blink.
- This feature is a useful way to mark a piece of hardware so that a local administrator can quickly identify the hardware for troubleshooting or replacement.

- Blue beacon LEDs allow for easy identification for servicing.

A single blue LED on all the removable parts (fan trays, power supplies, and modules) can be used to make sure that an operator does not turn off or replace the wrong part on the Nexus system. The network operations center can turn on the LED via the Cisco NX-OS Software and ensure that one of the common causes of errors is avoided.

Using the Blue Beacon Feature

- To flash the LEDs on a hardware entity, use the following commands:

Command	Purpose
locator-led chassis	Flashes the chassis LED
locator-led fan *number*	Flashes one of the fan LEDs
locator-led module *slot*	Flashes the selected module LED
locator-led powersupply *number*	Flashes one of the power supply LEDs
locator-led xbar *number*	Flashes one of the crossbar module LEDs

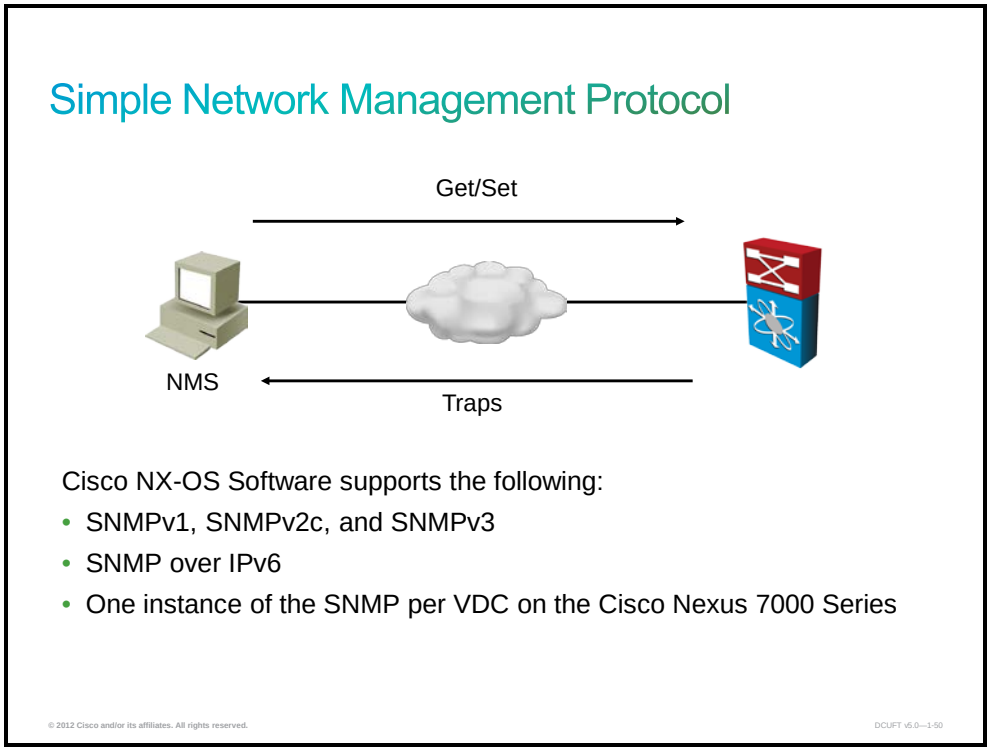
- To flash a single port LED on a module, use the following command in interface configuration mode:

Command	Purpose
locator-led	Flashes the interface LED

To flash the LEDs on a hardware entity or to flash a single port LED on a module, use the commands in the table shown here. The LED can be solid blue (operator has flagged this entity for identification) or off (entity not flagged).

SNMP and RMON

This topic describes how to use Simple Network Management Protocol (SNMP) and Remote Monitoring (RMON) to monitor and troubleshoot a Cisco Nexus or MDS switch.



SNMP provides a standard framework and common language that are used for the monitoring and management of devices in the network.

The SNMP framework consists of three parts:

- **SNMP manager:** Used to control and monitor the activities of network devices using SNMP
- **SNMP agent:** A software component within the managed device that maintains the data for the device and reports the data to managing systems. The Cisco NX-OS Software supports the agent and MIB, enabling the SNMP agent relationship between the manager and the agent, which both must be defined
- **MIB:** Collection of managed objects on the SNMP agent

SNMP notifications are used to indicate improper user authentication, restarts, the closing of the connection, the loss of the connection to a neighbor router, or other significant events. These notifications are generated by the SNMP agent and sent to the SNMP manager.

The Cisco NX-OS Software generates SNMP notifications as either traps or informs. A trap is an asynchronous, unacknowledged message that is sent from the agent to the SNMP managers listed in the host receiver table. Informs are asynchronous messages that are sent from the SNMP agent to the SNMP manager, which the manager must acknowledge. Traps are less reliable than inform messages because the SNMP manager does not send any acknowledgments of receipt of the trap. SNMP managers that receive an inform request acknowledge receipt; if the Cisco NX-OS Software does not receive that receipt, it sends the inform message again.

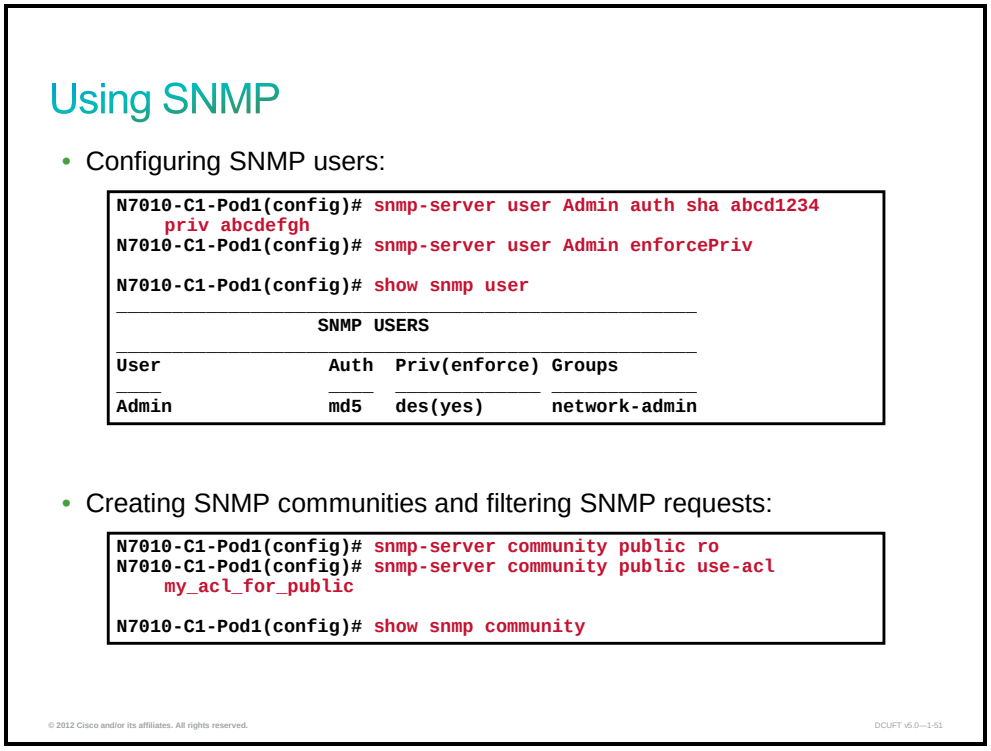
SNMP version 3 (SNMPv3) provides secure access to devices by authenticating and encrypting frames over the network. The security features provided include the following:

- **Message integrity:** Ensures that a message has not been tampered with in-transit
- **Authentication:** Ensures that the message is from a valid source
- **Encryption:** Prevents the message from being seen by unauthorized sources

SNMPv3 provides for security models and security levels. A security model is an authentication strategy that is set up for users, while the security level is the permitted level of security inside a security model. The combination of the security model and security level determines which mechanism is employed when processing SNMP packets.

The Cisco NX-OS Software supports one instance of SNMP per VDC on the Cisco Nexus 7000 Series Switch. SNMP is VRF-aware, and you can configure SNMP to use a particular VRF to reach the SNMP notification host receiver. The SNMP notification filter can configure the SNMP host receiver that is based on the VRF where the notification occurred.

Cisco Nexus 5000 Series Switches only support read-only for MIBs on all versions of SNMP. Nexus 7000 Series Switches support read-write for certain MIBs in version 3 and read-only for all MIBs in all versions.



The figure shows how to configure a user for SNMPv3 and how to create SNMP communities for SNMPv1 or SNMPv2c.

The **snmp-server user** *name* **enforcePriv** command enforces SNMP message encryption for this user.

The **show snmp user** command validates SNMP users.

The **snmp-server community** *community-name* **{ro | rw}** command creates an SNMP community string.

The **snmp-server community** *community-name* **use-acl acl-name** command assigns an access control list (ACL) to an SNMP community to filter SNMP requests.

The **show snmp community** command displays configured SNMP communities.

Using SNMP (Cont.)

- Configuring SNMP notification receivers:

```
N7010-C1-Pod1(config)# snmp-server host 192.0.2.1 informs version 3
    auth NMS
N7010-C1-Pod1(config)# snmp-server host 192.0.2.1 source-interface
    ethernet 2/1
N7010-C1-Pod1(config)# snmp-server user NMS auth sha abcd1234 priv
    abcdefgh engineID 00:00:00:63:00:01:00:10:20:15:10:03
```

- Assigning the SNMP device contact and location information:

```
N7010-C1-Pod1(config)# snmp-server contact Admin
N7010-C1-Pod1(config)# snmp-server location Lab-7

N7010-C1-Pod1(config)# show snmp
sys contact: Admin
sys location: Lab-7
```


The first figure shows the proper steps that are required to configure SNMP notification receivers. The second figure shows how to set SNMP device contact and location information.

The **snmp-server host** *ip-address* **{traps | informs} version 3 {auth | noauth | priv}** *username* **[udp_port number]** command configures a host receiver for SNMPv3 traps or informs. The *ip-address* can be an IPv4 or IPv6 address. The *username* can be any alphanumeric string up to 255 characters. The UDP port number range is from 0 to 65535.

You can configure SNMP to use a specific interface as the source interface for notifications. Use the following command in global configuration mode to configure a host receiver on a source interface:

snmp-server host *ip-address* **source-interface** *if-type if-number* [**udp_port** ***number***]

You must configure a notification target user on the device to send SNMPv3 inform notifications to a notification host receiver. Cisco NX-OS Software uses the credentials of the notification target user to encrypt the SNMPv3 inform notification messages to the configured notification host receiver.

Use the following command in global configuration mode to configure the notification target user:

snmp-server user *name* **[auth {md5 | sha}** *passphrase* **[auto] [priv [aes-128]** *passphrase***] [engineID** *id***]**

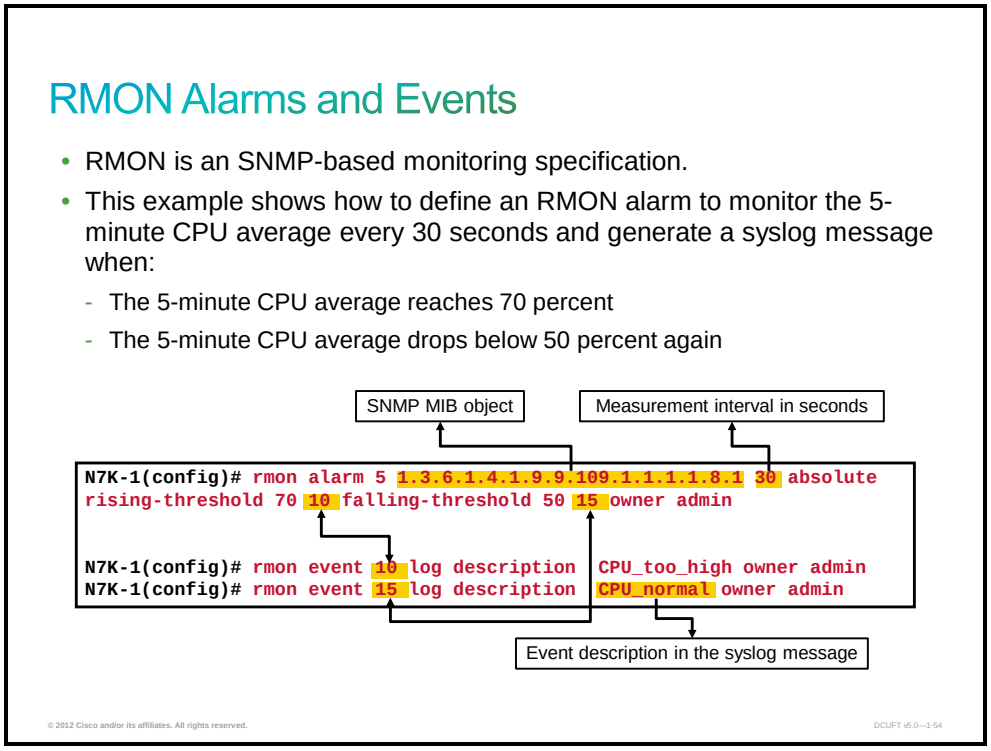
The **show snmp** command displays configured SNMP device contact and location information.

Cisco Embedded Event Manager

- EEM can run shell commands and Tool Command Language (Tcl) commands.
- Cisco NX-OS Software automatically bypasses TACACS authorization (IOS requires an event manager session command for AAA authorization to run the commands).
- Example of EEM script for high CPU:

```
version 5.1(3)
event manager applet Test
  event snmp oid 1.3.6.1.4.1.9.9.109.1.1.1.1.6.1 get-type exact
    entry-op ge
entry-val 10 poll-interval 300
  action 1.0 syslog msg High CPU hit $_event_pub_time
  action 2.0 cli enable
  action 3.0 cli show clock >> bootflash:high-cpu.txt
  action 4.0 cli show policy-map interface control-plane >>
             bootflash:high-cpu.txt
```


Cisco EEM is a powerful tool integrated with Cisco NX-OS Software for system management from within the device itself. Cisco EEM offers the ability to monitor events and take informational, corrective, or any desired action when the monitored events occur or when a threshold is reached. Capturing the state of a router during such situations can be invaluable in taking immediate recovery actions and gathering information to perform root-cause analysis. Network availability is also improved if automatic recovery actions are performed without the need to fully reboot the routing device.



RMON is an SNMP standard-based monitoring specification that allows various network agents and console systems to exchange network monitoring data. Cisco NX-OS Software supports RMON alarms, events, and logs to monitor Cisco NX-OS devices.

An RMON alarm monitors a specific MIB object for a specified interval, triggers an alarm at a specified threshold value, and resets the alarm at another threshold value. You can link RMON alarms to RMON events to generate a log entry or an SNMP notification when the RMON alarm triggers.

RMON is disabled by default and no events or alarms are configured in Cisco NX-OS Software. You can configure your RMON alarms and events by using the CLI or an SNMP-compatible network management station.

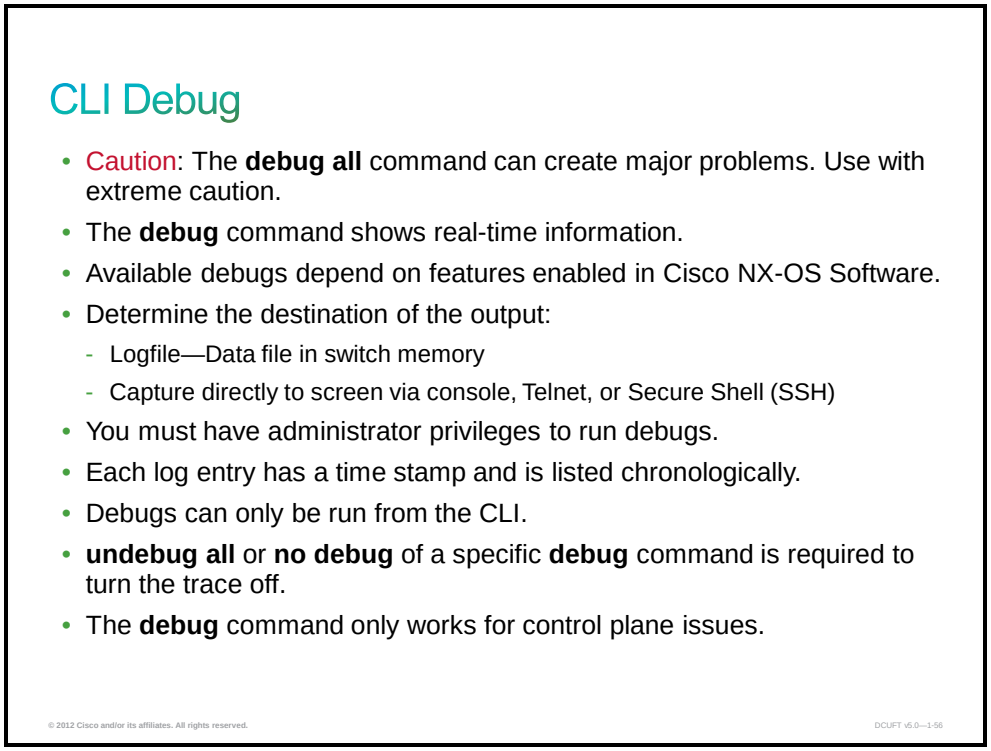
RMON configuration consists of two elements. The first element is the RMON alarm, which defines the MIB variable to monitor, the interval to sample the MIB value with, and a rising and falling threshold that define when the alarm is raised or dropped. The second element is the RMON event, which defines the action that should be taken. The action can either be an SNMP trap or a syslog message. In the RMON alarm definition, you can link separate RMON events to the rising threshold and falling threshold.

The example in the figure creates an RMON alarm that samples the MIB object 1.3.6.1.4.1.9.9.109.1.1.1.1.8.1 cpmCPUTotal5minRev every 30 seconds. When this value exceeds 70 percent, the alarm will be raised and when the value drops below 50 percent the alarm will be dropped. When the alarm is raised, RMON event 10 is triggered, and when the alarm is dropped, RMON event 15 is triggered.

RMON event 10 generates a syslog message with the text “CPU_too_high,” while RMON event 15 generates a syslog message with the text “CPU_normal.” Instead of generating a syslog message, it is also possible to send an SNMP trap, or send both an SNMP trap and a syslog message.

CLI Debug

This topic explains how to use the CLI debug feature to show real-time information while actively troubleshooting a network.

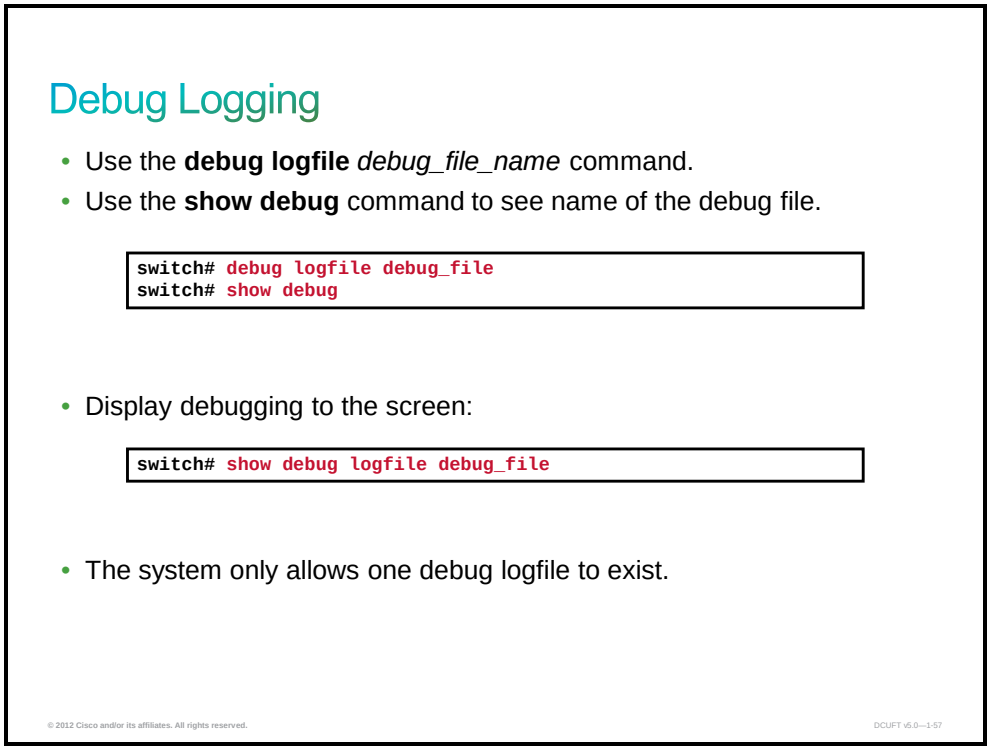


Cisco NX-OS Software supports an extensive debugging feature that is set for actively troubleshooting a network. Using the CLI, a user can enable debugging modes for each feature and view a real-time updated activity log of the control protocol exchanges. Each log entry has a time stamp and is listed chronologically. Access to the debug feature can be limited through the CLI roles mechanism to partition access on a per-role basis. While the **debug** commands show real-time information, the **show** commands can be used to list historical and real-time information.

The **debug** commands should be used very carefully, because some debug commands can affect network performance. By using the **?** option, the options that are available for any feature can be seen.

The debug output shows a time-stamped account of the activity that occurred between the local device and other adjacent devices.

Note	When working with a Cisco Nexus 7000, it is very important to be in the correct VDC, depending on what is being debugged.
Caution	Using **debug** commands can be very dangerous. Use these commands very carefully.



Debug messages can be logged to a special log file, which is more secure and easier to process than sending the debug output to the console.

The log file can, for example, be named “debug_file”, using the **debug logfile** command. Then, the **show debug** command can be used to see name of the debug file.

```
switch# debug logfile debug_file
switch# show debug
```

The following command displays debugging to the screen:

```
switch# show debug logfile debug_file
```

The following example shows the copying of the debug file from a switch to a server with the **copy** command. If no VRF is specified, then the default is used.

```
switch# copy log:debug_file tftp:
Enter vrf: management
Enter hostname for the tftp server: 10.91.42.134
Trying to connect to tftp server......
Connection to Server Established.
|
TFTP put operation was successful
```


To delete the debug logfile, one of the following commands can be used:

```
switch# clear debug-logfile debug_file
switch# undebug all
```

If a user does not use one of these commands, the debug logfile will be cleared and overwritten when the next debug logfile is created.

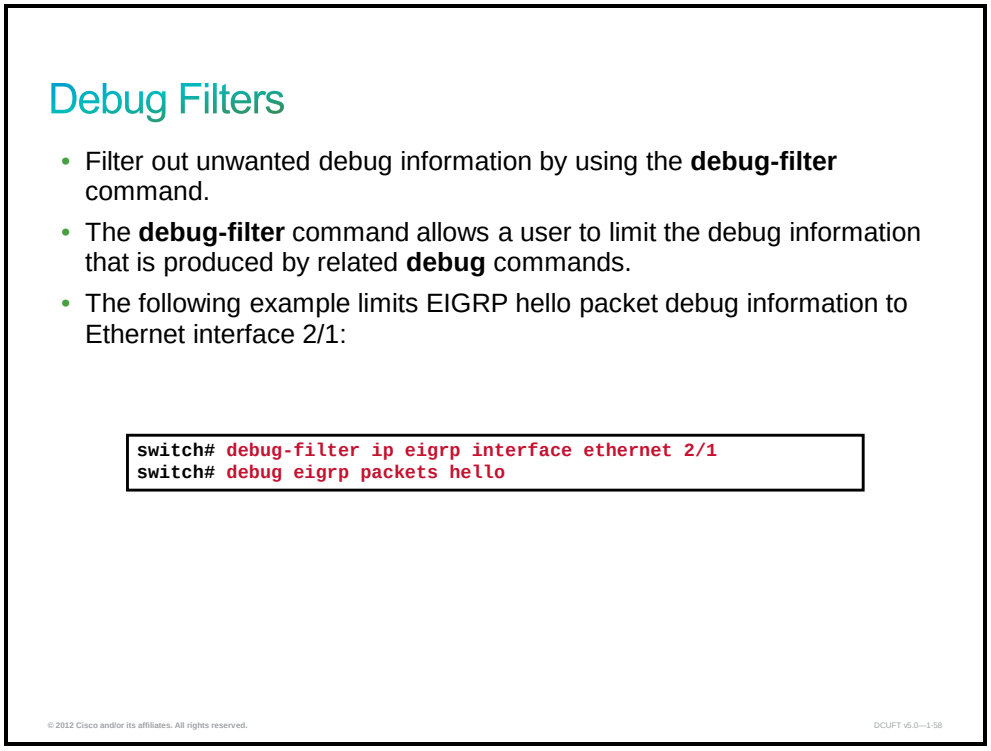
There are 5 files of 4 MB each with specified debug filenames ending with .1, .2 , .3, .4, and .5. This provides up to 20 MB of debugs. These files are stored in logflash, and an example follows:

```
N7K-2#debug logfile zixu-debug1.txt
N7K-2#debug ip packet detail
N7K-2#show debug
Output forwarded to file zixu-debug1.txt (size: 4194304 bytes)
Debug level is set to Minor(1)
default for new sessions logging level: 3
debug ip packet detail
`end`
```

To view the files in the logflash directory, use the following command from the default VDC:

```
dir logflash:///vdc_#/log/
```

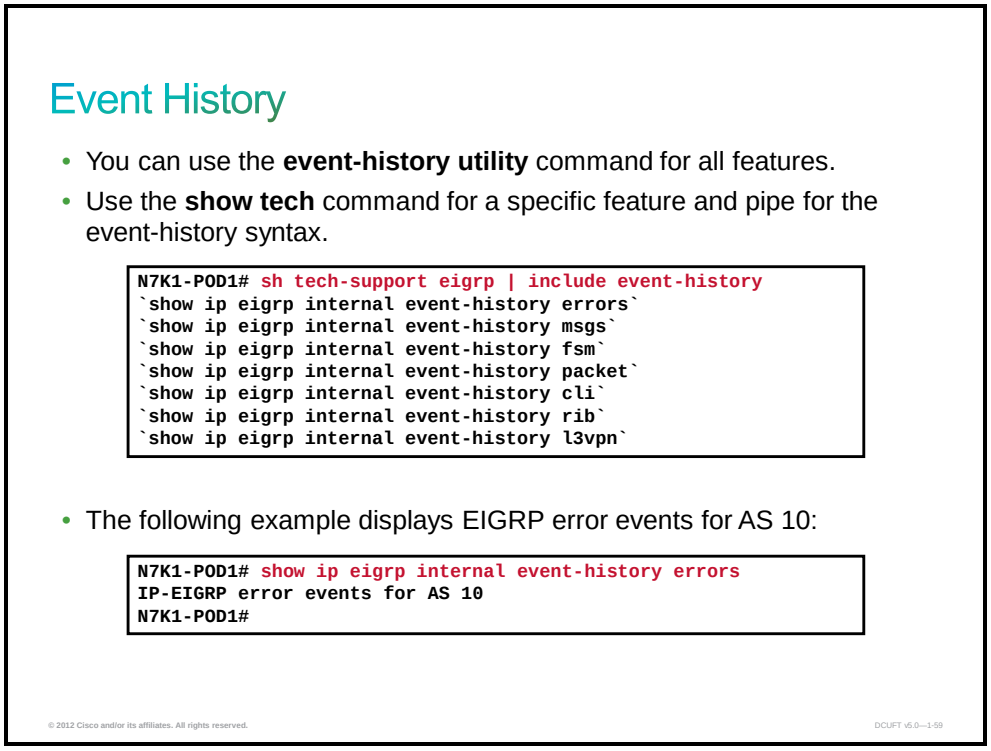
The # should be replaced with the VDC where the debugging was performed.



Unwanted debug information can be filtered out by using the **debug-filter** command. The **debug-filter** command allows you to limit the debug information that is produced by related debug commands.

The following example limits Enhanced Interior Gateway Routing Protocol (EIGRP) hello packet debug information to Ethernet interface 2/1:

```
switch# debug-filter ip eigrp interface ethernet 2/1
switch# debug eigrp packets hello
```

You can use the **event-history** command for all features. This utility acts similar to having a debug running all the time. It can be very useful when looking for the root cause of historical or current problems. You can use the **show tech** command for a specific feature and pipe for the event-history syntax.

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- Basic troubleshooting methodology includes steps to recognize symptoms, gather information, analyze data, determine a root cause and remediate a problem.
- Use the **ping** and **traceroute** commands to troubleshoot problems with connectivity and path choices; the **pong** command can measure the delay of the network between two points.
- The **show processes** command identifies the processes that are running and the status of each process; the **show system resources** command displays system-related CPU and memory statistics.
- The SPAN feature selects network traffic for analysis by a network analyzer.
- Ethanalyzer is a Cisco NX-OS protocol analyzer tool based on a command-line version of Wireshark that captures and decodes packets.
- The system message logging software saves messages in a log file or directs the messages to other devices so they may be later used for monitoring and troubleshooting.

Summary (Cont.)

- Cisco GOLD checks the health of hardware components and verifies proper operation of the system data and control planes.
- The blue beacon feature is a useful way to mark a piece of hardware so that a local administrator can quickly identify the hardware for troubleshooting or replacement.
- RMON alarms and events provide a mechanism for setting thresholds and sending notifications based on changes in network behavior.
- A user can enable debugging modes for each feature and view a real-time updated activity log of the control protocol exchanges.

Lesson 2

Understanding Cisco DCNM Troubleshooting Tools

Overview

This lesson is designed to show the student some of the common tools and methodologies that are used in troubleshooting a Cisco Data Center architecture using the Cisco Data Center Network Manager (DCNM).

Objectives

Upon completing this lesson, you will be able to describe the troubleshooting tools and methodologies that are available in Cisco DCNM that are used to identify and resolve issues in the Cisco Data Center network architecture. You will be able to meet these objectives:

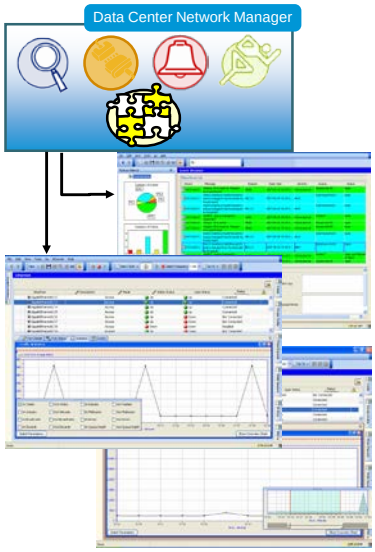
- List the features available in Cisco DCNM for LAN to help troubleshoot issues on the Cisco Unified Fabric
- List the features available in Cisco DCNM for SAN to help troubleshoot issues on the Cisco Unified Fabric

Cisco DCNM for LAN

This topic lists the features that are available in Cisco DCNM for LAN to help troubleshoot issues on the Cisco Unified Fabric.

Cisco DCNM

- GUI management solution
- Operational monitoring of data center infrastructure
  - Proactive monitoring
  - Performance and capacity
  - Topology views
- Data center resource management
  - Automated discovery
  - Configuration and change management
  - Template-based provisioning
- Image management
  - Integration with enterprise systems
  - Web services APIs
  - Event forwarding



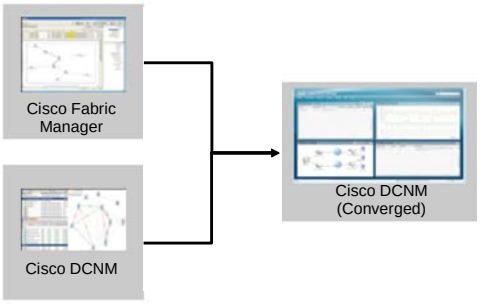
Cisco DCNM is a GUI management solution that maximizes overall data center infrastructure uptime and reliability, which improves business continuity. Focused on the management requirements of the data center network, Cisco DCNM provides a robust framework and rich feature set that fulfills the switching needs of present and future data centers. In particular, Cisco DCNM automates the provisioning process.

Web services application programming interfaces (APIs) provide easy integration with third-party applications and allow accurate flow-through provisioning and data mining.

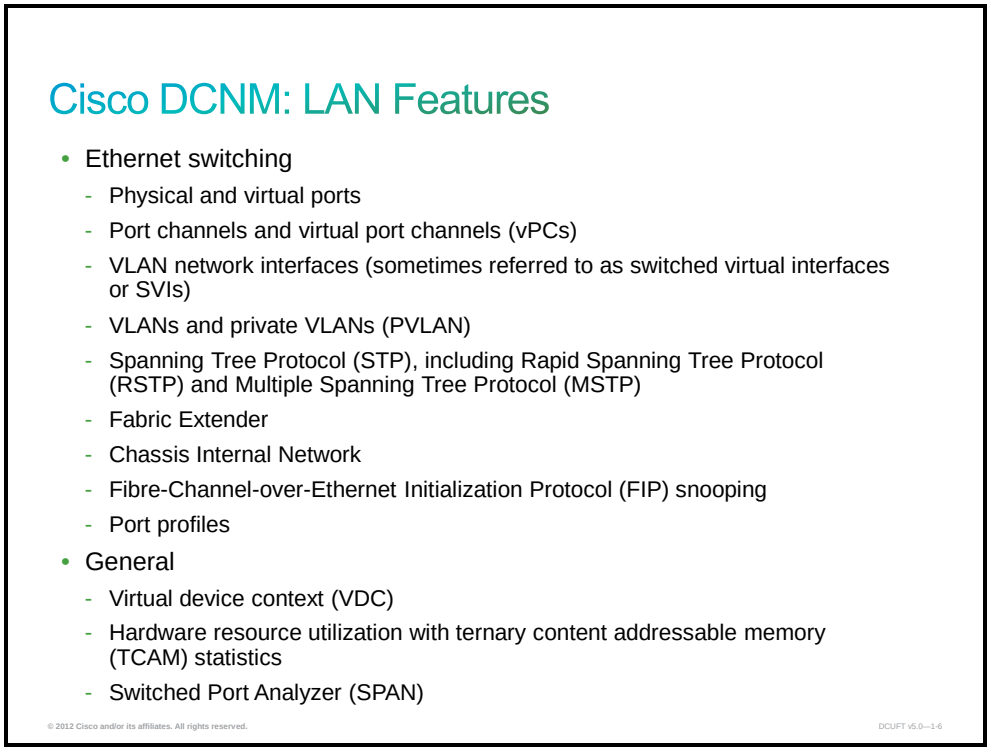
Event forwarding integrates Cisco DCNM with the enterprise network operations center (NOC) for alerts and events. Cisco DCNM uses email and alerts to notify operations staff of critical outages that may impact service.

Cisco DCNM: Converged

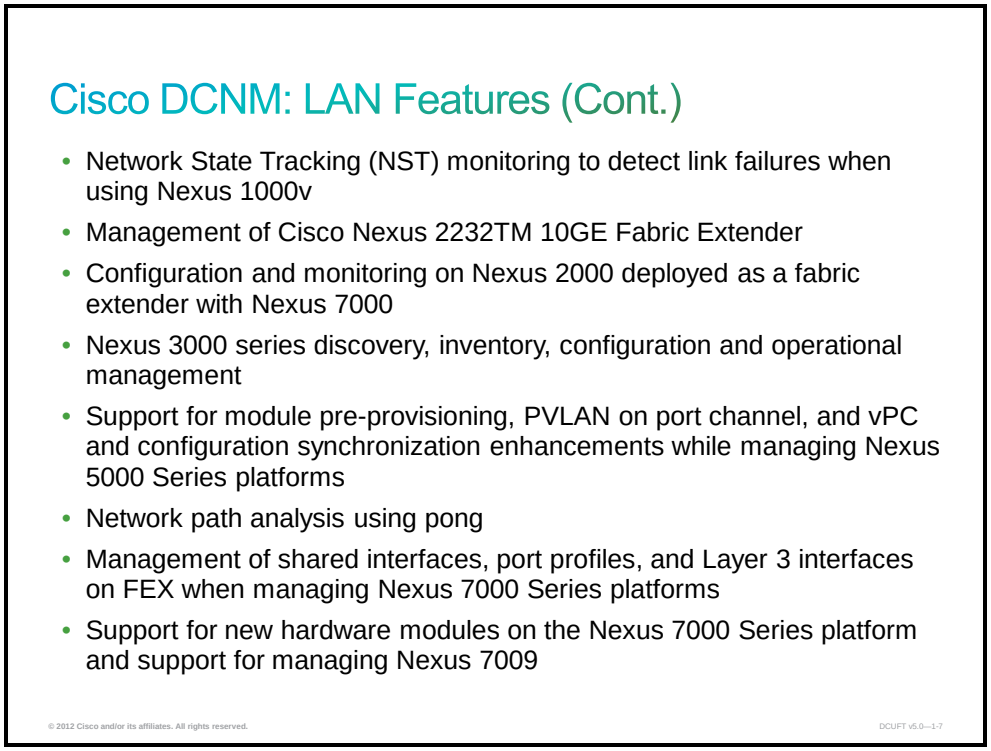
- One converged product
  - SAN and LAN health and performance dashboards
  - Can be licensed for SAN or LAN or both
  - Common operations (discovery and topology)
  - Single installer and RBAC
  - Consistent licensing model (licenses on DCNM server)



Recognizing the need to support this convergence in management, Cisco is merging two best-in-class management solutions, Cisco Fabric Manager and Cisco Data Center Network Manager for LAN, into one unified product called Cisco Data Center Network Manager (DCNM). Administrators can still maintain control and segmentation through role-based access control (RBAC), now with single-pane visibility across the network and storage access infrastructure. Cisco DCNM streamlines the provisioning of the unified fabric and proactively monitors the SAN and LAN components. Offering an exceptional level of visibility and control through a single pane for the Cisco Nexus, Cisco Unified Computing System (UCS), and Cisco MDS 9000 family of products, Cisco DCNM is the solution that Cisco recommends for managing mission-critical data center networks. Cisco DCNM can be licensed to manage a combination of SAN and LAN environments. New features are available depending on the licenses that are deployed.



The figure lists Cisco DCNM Ethernet switching and general features.



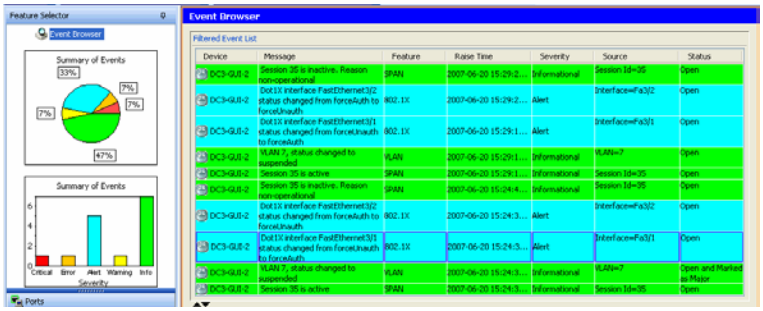
The figure lists additional features in Cisco DCNM for LAN.

Cisco DCNM supports a variety of Cisco hardware platforms, including the following:

- Cisco MDS 9500 Multilayer Directors and Cisco MDS 9200 and 9100 Series Multilayer Switches
- Cisco Nexus 7000, 5000, 4000, 3000, and 2000 Series Switches
- Cisco Nexus 1000v Virtual Switches
- Cisco Catalyst 6500 Series Switches
- Cisco UCS 6100 Series Fabric Interconnects

Cisco DCNM: Fault Management

- Industry standard Event Browser
- Event collection and normalization
- Per network feature correlation
- Noise filtering for root cause isolation
- Event propagation
  - Actionable tasks
  - Integration in the SMF



The Cisco DCNM client includes the Event Browser and feature-specific Events tabs that appears in the Details pane for features that can have events. The Event Browser shows all recent status events while a feature-specific Events tab shows recent status events that pertain to the feature.

The Cisco DCNM client updates the Event Browser and Events tabs dynamically when it receives new events from the server. A user can use the Event Browser to view recent events and a summary chart of those events.

A user can filter events in the Event Browser by the following criteria:

- **Event date and time:** By default, the Cisco DCNM client displays all events that are received after you started the Cisco DCNM client and for a configurable number of hours before starting the Cisco DCNM client
- **Event severity:** By default, the Cisco DCNM client displays events of all severities.

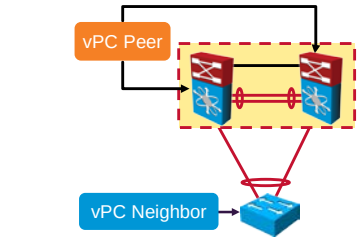
A user can change the status of an event to one of the following statuses:

- **Acknowledged:** Shown as a green check mark
- **Closed:** Shown as a yellow folder

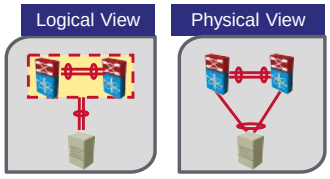
Cisco DCNM: vPC Management

vPC automation: managing both devices as one

- Configuration
  - Step-by-step wizard setup
  - Configuration audit between primary and secondary
  - Automatic resolution of configuration conflicts
  - Easy role switch
  - HSRP and STP failover settings
- vPC peer link and vPC fault-tolerant link monitoring
- Per vPC events filtering
- vPC traffic aggregation for links utilization, keepalive statistics
- Topology representation
  - Physical view
  - Logical view



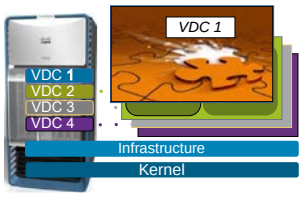
Auto Resolve Configuration Conflicts

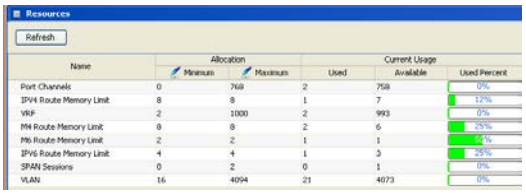


The Cisco DCNM interface allows a user to configure the parameters for the primary virtual PortChannel (vPC) peer device, automatically propagates those parameters to the secondary vPC device, and configures consistency checks.

Cisco DCNM: VDC Management

- VDCs are transparently handled throughout the application wizard-based configuration
  - Interface allocation across VDC
  - Resource limit enforcement with templates
  - Resource consumption monitoring
  - IPv4- and IPv6-capable
- VDC-aware fault and performance monitoring
- VDC-aware RBAC
- Topology representation
- VDC per chassis
- VDC-to-VDC connectivity
- Real-time or delayed discovery



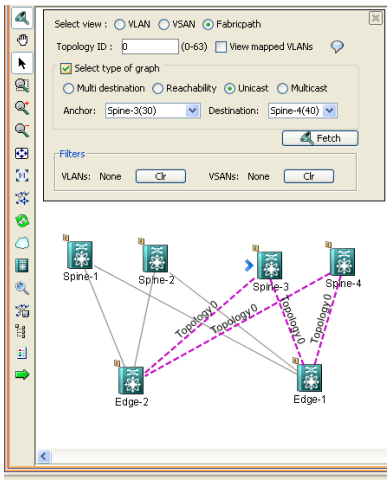


The figure describes managing virtual device contexts (VDCs) with Cisco DCNM.

The VDC Setup Wizard is used to create a new VDC.

Cisco DCNM: Cisco FabricPath Management

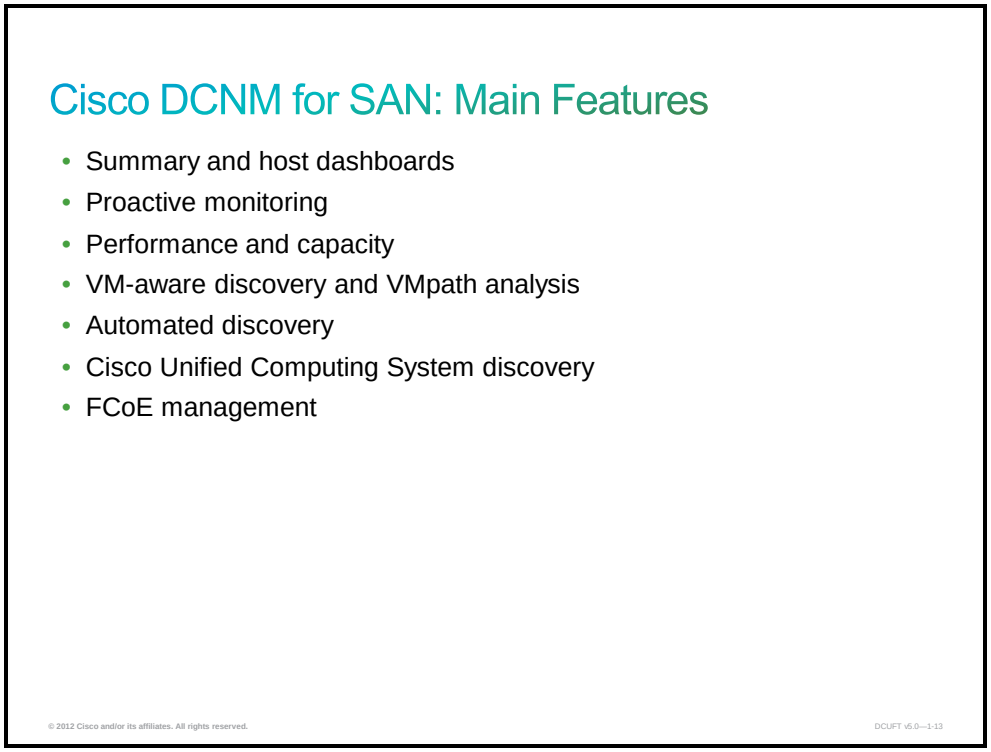
- Multipathing monitoring
  - Multipath traffic distribution: unicast, multicast, and broadcast graphs visualization
  - Cisco FabricPath link state awareness
- Troubleshooting
  - Reachability and latency between source and destination nodes
  - Threshold crossing alerts for fabric “hot spots” identification
- Configuration expert
  - One-touch Cisco FabricPath domain turn-up
  - Wizard-based control plane and Cisco FabricPath topologies settings



The figure displays information about configuring and monitoring the Cisco FabricPath feature.

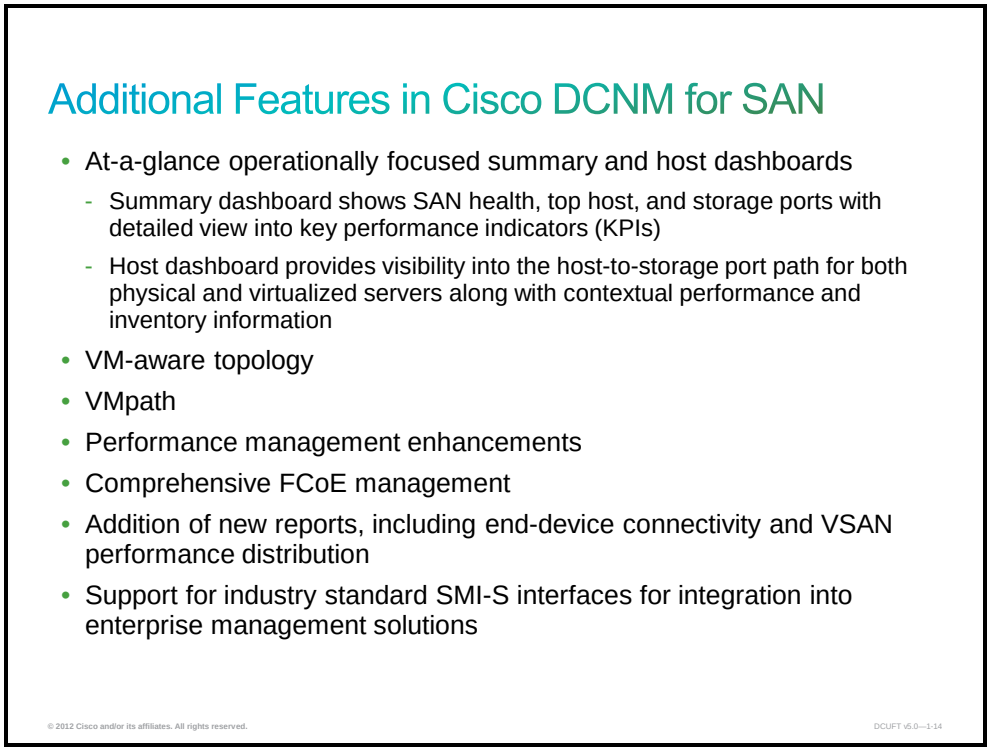
Cisco DCNM for SAN

This topic lists the features that are available in Cisco DCNM for SAN to help troubleshoot issues on the Cisco Unified Fabric.



There are several main features that are provided in Cisco DCNM for SAN:

- **Summary and host dashboards**: This feature offers a real-time fabric and network health summary with detailed views of individual network components, enabling operations staff to respond quickly to events based on their severity.
- **Proactive monitoring**: This feature facilitates early detection and prevention of outages, increasing network availability.
- **Performance and capacity**: This feature provides detailed visibility into real-time and historical performance statistics in the data center.
- **Virtual Machine-aware (VM-aware) discovery and VMpath analysis**: This feature provides a view of the virtual path through the physical fabric out to the storage array and to the data store, offering the capability to view performance for every switch hop to the individual VMware ESX server and virtual machine (VM).
- **Automated discovery**: This feature provides up-to-date physical and logical inventory information.
- **Cisco Unified Computing System discovery**: This feature discovers and monitors the Cisco UCS fabric interconnect.
- **Fibre Channel over Ethernet (FCoE) management**: This feature provides the capability to discover, provision, and monitor the FCoE path.



Depending on the licenses that are deployed, the following additional features in Cisco DCNM for SAN are available:

- At-a-glance operationally focused summary and host dashboards
  - Summary dashboard shows SAN health, top host, and storage ports with detailed view into key performance indicators (KPIs)
  - Host dashboard provides visibility into the host-to-storage port path for both physical and virtualized servers, along with contextual performance and inventory information
- VM-aware topology view showing all the dependencies from the VM to the physical host, to the switch, and to storage, with one-click access to their attributes; VM-aware views increase service availability by identifying performance bottlenecks on a VM and on a VMware ESX server, and extend visibility to the rest of the switch fabric
- VMpath offers detailed topology views providing end-to-end path information across multiple fabric clouds, shortest path, and all possible paths to available storage ports
- Performance management enhancements including the capability to view trends and correlation across multiple performance indicators on a single chart, interactive zooming options, and predictive analysis features
- Comprehensive FCoE management, including provisioning, discovery, and operation monitoring across a wide variety of Cisco platforms (Cisco Nexus 5000 and 7000 Series Switches and Cisco MDS 9000 family directors)
- Addition of new reports, including end-device connectivity and VSAN performance distribution
- Support for industry standard Storage Management Initiative Specification (SMI-S) interfaces for integration into enterprise management solutions

VM-Aware Path Management

- The VMpath views increase service availability by identifying bottlenecks in VM and VMware ESX performance and extending visibility to the physical fabric.
- The VM-aware dashboard displays all the information needed to manage the virtual environment.

Cisco DCNM simplifies management of the virtual infrastructure by enabling management of the entire path through the physical to the virtual network across the whole data center environment.

The VMpath views increase service availability by identifying bottlenecks in VM and VMware ESX performance and extending visibility to the physical fabric. The VM-aware topology view shows all the dependencies from the VM out to the physical host, through the fabric, and to the storage array with easy access to a detailed view of the path attributes. The VM-aware dashboard displays all the information that is needed to manage the virtual environment, including performance charts, inventory information, events, and VM and VMware ESX utilization information. Cisco DCNM maps paths all the way from the server to storage, enabling tracking of mission-critical workloads across the entire network.

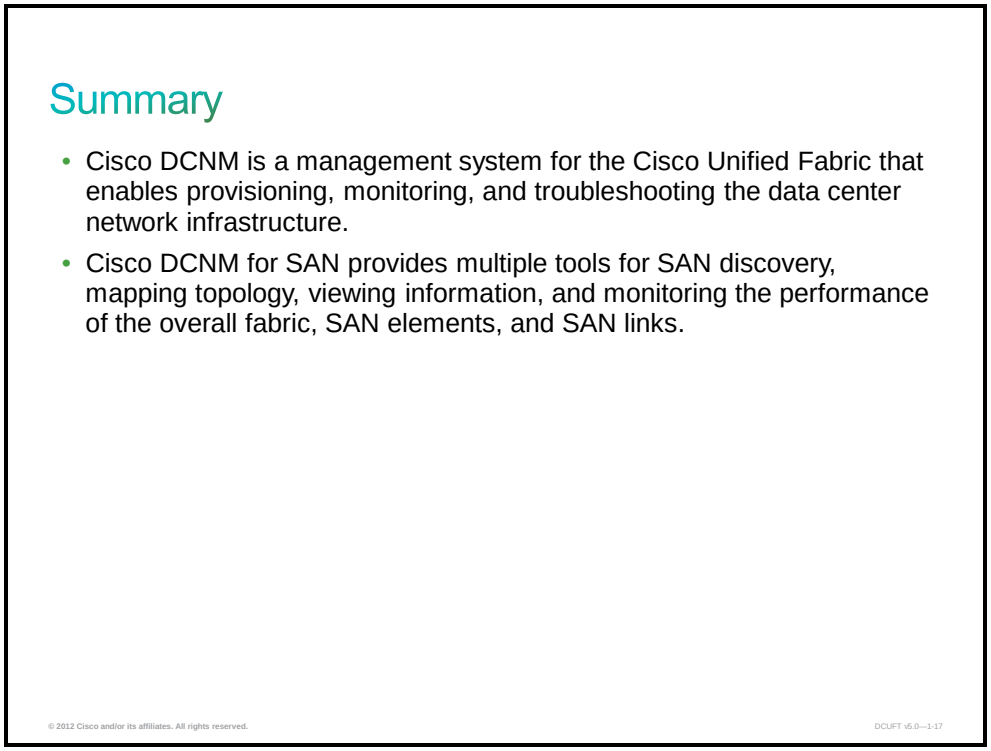
Performance and Troubleshooting

- Monitoring and providing alerts for fabric availability and performance
  - Near-real-time monitoring of fabrics
  - Visibility into traffic spikes
  - Establishing baseline traffic patterns
  - Generating alerts when predefined thresholds are breached

Cisco DCNM helps IT ensure the resiliency of the Cisco SAN infrastructure by monitoring and providing alerts for fabric availability and performance. It provides near-real-time monitoring of fabrics with visibility into traffic spikes and establishes baseline traffic patterns. When predefined thresholds are breached, appropriate alerts are generated and can be forwarded to operations staff and enterprise operations consoles for incident management.

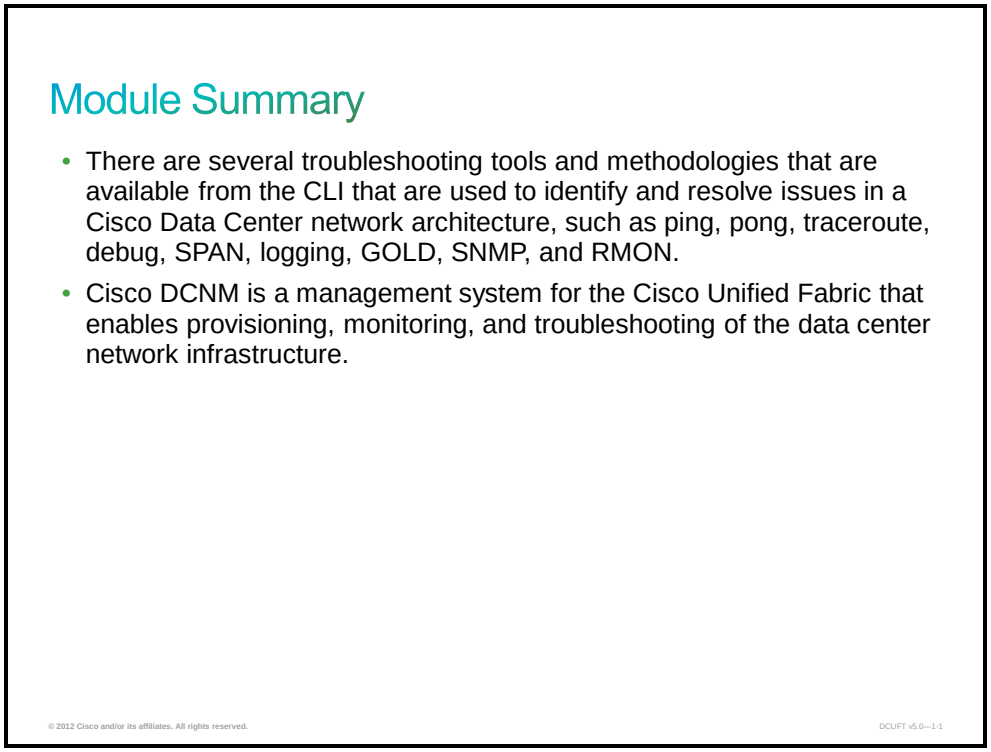
Summary

This topic summarizes the key points that were discussed in this lesson.



Module Summary

This topic summarizes the key points that were discussed in this module.



The CLI allows you to configure and monitor the Cisco Nexus Operating System (NX-OS) using a local console or remotely using a Telnet or Secure Shell (SSH) session. The CLI provides a command structure similar to Cisco IOS Software, with context-sensitive help, **show** commands, multiuser support, and role-based access control (RBAC).

Cisco Data Center Network Manager (DCNM) increases overall data center infrastructure uptime and reliability, thereby improving business continuity. It provides a robust framework and comprehensive feature set that meets the routing, switching, and storage administration needs of data centers. Cisco DCNM streamlines the provisioning for the Cisco Unified Fabric and monitors the SAN and LAN components. Cisco DCNM provides a high level of visibility and control through a single web-based management console for Cisco Nexus, Cisco MDS, and Cisco Unified Computing System (UCS) products.

Module Self-Check

Use the questions here to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) Which two commands are used to troubleshoot problems with connectivity and path choices? (Choose two.) (Source: Understanding CLI Troubleshooting Tools)

A) **debug**
B) **ping**
C) **pong**
D) **traceroute**

Q2) Whereas the **debug** commands show real-time information, the **show** commands can be used to list historical and real-time information. (Source: Understanding CLI Troubleshooting Tools)

A) True
B) False

Q3) Which feature allows network traffic for analysis by a network analyzer to be transported across an IP network? (Source: Understanding CLI Troubleshooting Tools)

A) SPAN
B) ERSPAN
C) EEM
D) GOLD

Q4) Which three options are default locations to where the Cisco NX-OS Software sends log messages? (Choose three.) (Source: Understanding CLI Troubleshooting Tools)

A) log file in flash
B) NVRAM
C) console
D) syslog

Q5) What is the difference between SNMP traps and informs? (Source: Understanding CLI Troubleshooting Tools)

_____

_____

_____

_____

Q6) Cisco DCNM provides various real-time traffic statistics. (Source: Understanding Cisco DCNM Troubleshooting Tools)

A) True
B) False

Q7) Which three protocols can be used on file servers that are supported by the Cisco DCNM? (Choose three.) (Source: Understanding Cisco DCNM Troubleshooting Tools)

A) FTP
B) HTTP
C) SFTP
D) HTTPS
E) TFTP

Q8) How many managed ports per server does Cisco DCNM for SAN support? (Source: Understanding Cisco DCNM Troubleshooting Tools)

A) 3000
B) 5000
C) 10,000
D) 15,000

Q9) What does the FCoE management feature enable? (Source: Understanding Cisco DCNM Troubleshooting Tools)

A) discovers and monitors the Cisco UCS fabric interconnect
B) provides capability to discover, provision, and monitor the FCoE path
C) notifies operations staff through email and pages of critical outages that may affect service
D) provides a view of the virtual path through the physical fabric out to the storage array and to the data store

Module Self-Check Answer Key

Q1) B, D

Q2) A

Q3) B

Q4) A, B, C

Q5) A trap is an asynchronous, unacknowledged message sent from the agent to the SNMP managers listed in the host receiver table. Informs are asynchronous messages sent from the SNMP agent to the SNMP manager, which the manager must acknowledge. Traps are less reliable than inform messages because the SNMP manager does not send any acknowledgments of receipt of the trap. SNMP managers that receive an inform request acknowledge receipt; if the Cisco NX-OS Software does not receive that receipt, it sends the inform message again.

Q6) A

Q7) A, C, E

Q8) D

Q9) B

Module 2

Layer 2 Issue Troubleshooting

Overview

This module identifies common issues that are related to Layer 2 switching. It also presents methods for troubleshooting these issues. Topics include port channels, virtual port channels (vPCs), VLANs, private VLANs (PVLANs), and Cisco FabricPath.

Module Objectives

Upon completing this module, you will be able to identify and resolve issues that are related to Layer 2 switching in the Cisco Data Center architecture. This ability includes being able to meet these objectives:

- Identify and resolve issues that are related to VLANs and PVLANs
- Identify and resolve issues that are related to port channels and vPCs
- Identify and resolve issues that are related to Cisco FabricPath
- Identify and resolve issues that are related to OTV

Lesson 1

Troubleshooting VLANs and PVLANs

Overview

This lesson is designed to provide you with some examples of VLAN- and private VLAN (PVLAN)-related issues and show you how to resolve them.

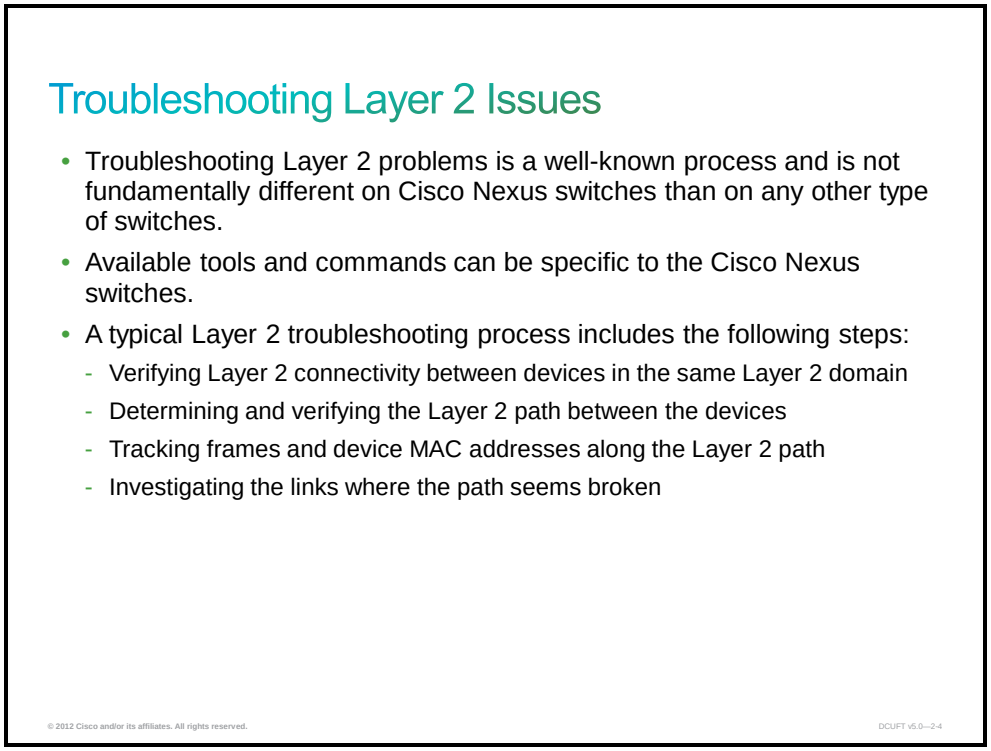
Objectives

Upon completing this lesson, you will be able to identify and resolve issues that are related to VLANs and PVLANs. You will be able to meet these objectives:

- Explain how to troubleshoot VLANs and PVLANs on the Cisco Unified Fabric
- Explain how to troubleshoot the VTP

Troubleshooting VLANs and PVLANs

This topic explains how to troubleshoot VLANs and PVLANs on the Cisco Unified Fabric.



Troubleshooting Layer 2 problems should be a well-known process to a network engineer. The way to approach this process is not fundamentally different on Cisco Nexus switches than it is on any other type of Layer 2 switch.

Usually, a Layer 2 troubleshooting process consists of finding out if the problem is specific to two devices, or more generally affecting multiple devices. This will determine the second step. Two specific devices would not normally indicate a Spanning Tree Protocol (STP) problem, for example. Next, you track the path between two devices in the same Layer 2 domain, after you have verified that they cannot communicate. Once you have determined the expected path, and the actual path according to the spanning-tree topology, you can start tracking the flow of the frames through the switches. Once you discover a point where the path seems to be broken, you investigate the suspected link or links to find the cause of the problem.

Typical Layer 2 Troubleshooting Process

- A typical Layer 2 troubleshooting process includes the following steps and commands:
  - Check if the port is up and receiving and forwarding traffic—use the **show interface eth** *x/y* command
  - Look for any errors on the interfaces—use the **show interface eth** *x/y* **counters** command

```
N7K1-POD1# show interface ethernet 1/1 counters
--------------------------------------------------------------------------------
Port                                InOctets                    InUcastPkts
--------------------------------------------------------------------------------
Eth1/1                              15203506454                 13583856
--------------------------------------------------------------------------------
Port                                InMcastPkts                 InBcastPkts
--------------------------------------------------------------------------------
Eth1/1                              3302651                     10819449
--------------------------------------------------------------------------------
Port                                OutOctets                   OutUcastPkts
--------------------------------------------------------------------------------
Eth1/1                              9462790                     59
--------------------------------------------------------------------------------
Port                                OutMcastPkts                OutBcastPkts
--------------------------------------------------------------------------------
Eth1/1                              40651                       13
```


A typical Layer 2 troubleshooting process starts with checking the interfaces. Use the **show interface ethernet** *x/y* and **show interface ethernet** *x/y* **counters** commands to check if the port is up and receiving and forwarding traffic, and look for any errors on the interfaces.

Typical Layer 2 Troubleshooting Process (Cont.)

- A typical Layer 2 troubleshooting process includes the following steps and commands:
  - Understand the topology—use the **show cdp neighbors** command

```
N7K1-POD1# show cdp neighbors
Capability Codes: R - Router, T - Trans-Bridge, B - Source-Route-Bridge
                  S - Switch, H - Host, I - IGMP, r - Repeater,
                  V - VoIP-Phone, D - Remotely-Managed-Device,
                  s - Supports-STP-Dispute

Device-ID             Local Intrfce Hldtme Capability  Platform      Port ID
UCS-RL-MGMT           mgmt0         153    R S I     WS-C3550-48   Fas0/7

UCS-RL-MGMT           Eth1/1        125    R S I     WS-C3550-48   Fas0/25

N7K2-pod2(JAF1602BGDB)
                      Eth1/41       125    R S s     N7K-C7009     Eth3/41
```


The next step in the process of troubleshooting Layer 2 problems is determining and verifying the path between devices. Use the **show cdp neighbors** command to verify the Layer 2 path between devices.

Typical Layer 2 Troubleshooting Process (Cont.)

- A typical Layer 2 troubleshooting process includes the following steps and commands:
  - Check if the MAC addresses are being learned—use the **show mac address-table** command

```
N7K1-POD1# show mac address-table
Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link
   VLAN     MAC Address      Type      age     Secure NTFY Ports/SWID.SSID.LID
---------+-----------------+--------+---------+------+----+------------------
G     -    0026.51c9.78c3    static       -       F    F  sup-eth1(R)
* 1         0000.0c07.ac01    dynamic   0          F    F  Eth1/1
* 1         0000.cd29.255d    dynamic   30         F    F  Eth1/1
* 1         0000.cd29.256c    dynamic   30         F    F  Eth1/1
* 1         0006.53cd.0801    dynamic   30         F    F  Eth1/1
* 1         000c.296f.c3de    dynamic   30         F    F  Eth1/1
* 1         000c.29aa.c96a    dynamic   150        F    F  Eth1/1
* 1         000d.5d08.35f4    dynamic   270        F    F  Eth1/1
* 1         000d.ecfa.64c1    dynamic   60         F    F  Eth1/1
* 1         000f.2486.9e00    dynamic   0          F    F  Eth1/1
* 1         000f.2486.9e01    dynamic   1770       F    F  Eth1/1
* 1         000f.24fb.eb80    dynamic   1560       F    F  Eth1/1
```


Next, track frames and device MAC addresses along the Layer 2 path. Use the **show mac address-table** to verify if the MAC addresses are being learned.

Typical Layer 2 Troubleshooting Process (Cont.)

- A typical Layer 2 troubleshooting process includes the following steps and commands:
  - Check if the interface is in STP forwarding state (if it is blocking, the MAC address will not be learned)—use the **show spanning-tree vlan** *vlan-id* command

```
N7K1-POD1# show spanning-tree vlan 1

VLAN0001
  Spanning tree enabled protocol rstp
  Root ID    Priority    8193
             Address     0019.300e.aa00
             Cost        38
             Port        129 (Ethernet1/1)
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

  Bridge ID  Priority    32769  (priority 32768 sys-id-ext 1)
             Address     0023.04ee.be0c
             Hello Time  2  sec  Max Age 20 sec  Forward Delay 15 sec

Interface        Role Sts Cost      Prio.Nbr Type
---------------- ---- --- --------- -------- --------------------------------
Eth1/1           Root FWD 19        128.129  P2p Peer(STP)
```

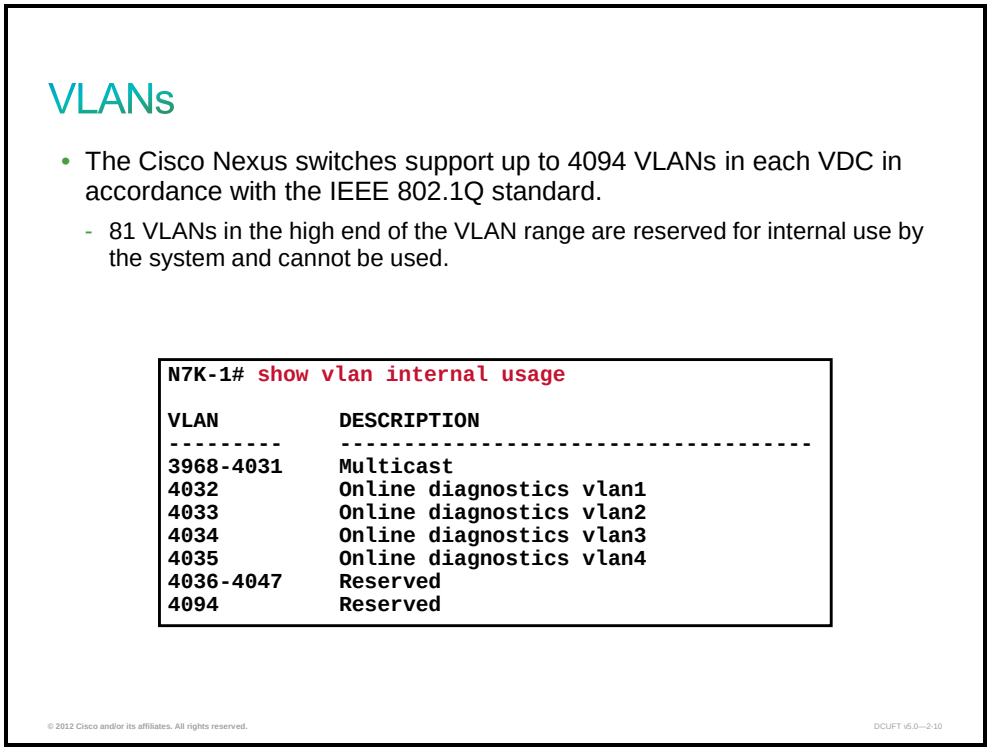

Investigate the link where the path seems to be broken. Check if the interface is in STP forwarding state by using the **show spanning-tree vlan** *vlan-id* command. If the interface is in the blocking state, the MAC address will not be learned.

Typical Layer 2 Troubleshooting Process (Cont.)

- A typical Layer 2 troubleshooting process includes the following steps and commands:
  - Check which links are bundled in the port channel—use the **show port-channel summary** command
  - Check if the ARP is learned on the SVI—use the **show ip arp vlan** *vlan-id* command
  - Configure notifications of MAC flapping – use the **mac address-table notification mac-move** command

```
switch# show ip arp vlan 900

IP ARP Table for context default
Total number of entries: 1
Address           Age       MAC Address     Interface
90.10.10.2        00:03:11  000d.ece7.df7c  Vlan900
```


You can also check if any of the links are bundled in the port channel by using the **show port-channel summary** command. By using the **show ip arp vlan** *vlan-id* command, you can also check the proper operation of the Address Resolution Protocol (ARP). To configure a log message notification of the MAC address table if the MAC address is moved, use the **mac address-table notification mac-move** command.



A switch access port belongs to a VLAN. Unicast, broadcast, and multicast packets are forwarded and flooded only to end stations in that VLAN. Each VLAN is considered a logical network. Packets that are destined for a station that does not belong to the same VLAN must be forwarded via Layer3.

The Cisco Nexus 7000 and 5000 Series Switches support up to 4094 VLANs, which are organized into ranges for specific tasks:

- **VLAN 1:** The default VLAN cannot be modified or deleted.
- **VLAN 2–1005:** Normal VLANs that can be created, used, modified, and deleted.
- **VLAN 1006–4094:** Extended VLANs that can be created, named, and used. The state of these VLANs is always active, and the VLANs are always enabled and cannot be shut down.
- **VLAN 3968–4047 and 4094:** These VLANs are allocated for internal use only.

For Cisco Nexus 7000 Series Switches, VLANs 3968–4047 and 4094 are reserved for internal use in each virtual device context (VDC) for features that need to use internal VLANs for their operation—for example, multicast and diagnostics. Due to the use of VDCs, a VLAN number can be reused in different VDCs, because each VDC is a separate virtual device. The maximum number of VLANs that can be supported across all VDCs is 16,000. The same VLANs are reserved for internal use of Cisco Nexus 5000 Series Switches, as well.

Deploying PVLANs in an enterprise data center environment provides an effective means of sparing IP address space and controlling Layer 2 access to servers and devices residing within the server farm. The Layer 2 isolation that is provided by PVLANs is an excellent way to supplement additional Layer 3 security that is already used to protect a particular server farm subnet. The procedure for troubleshooting PVLANs is the same as troubleshooting VLANs.

VLAN Significance

- Each VDC on the Cisco Nexus 7000 switch can support the full range of 4094 VLANs.
- VLANs within each VDC are isolated from VLANs in other VDCs on the same switch.

```
N7K-1(config)# vlan 20
N7K-1(config-vlan)# end

N7K-1# switchto vdc Red

N7K-1-Red# config
N7K-1-Red(config)# vlan 20
N7K-1-Red(config-vlan)#
```


When configuring VLANs, ensure that you are in the correct VDC, or issue the **switchto vdc** command. VLAN names and IDs can be repeated in different VDCs. It is very important that you confirm the VDC in which you are working.

VLANs have VDC local significance within the system, and interfaces have local significance within a VDC.

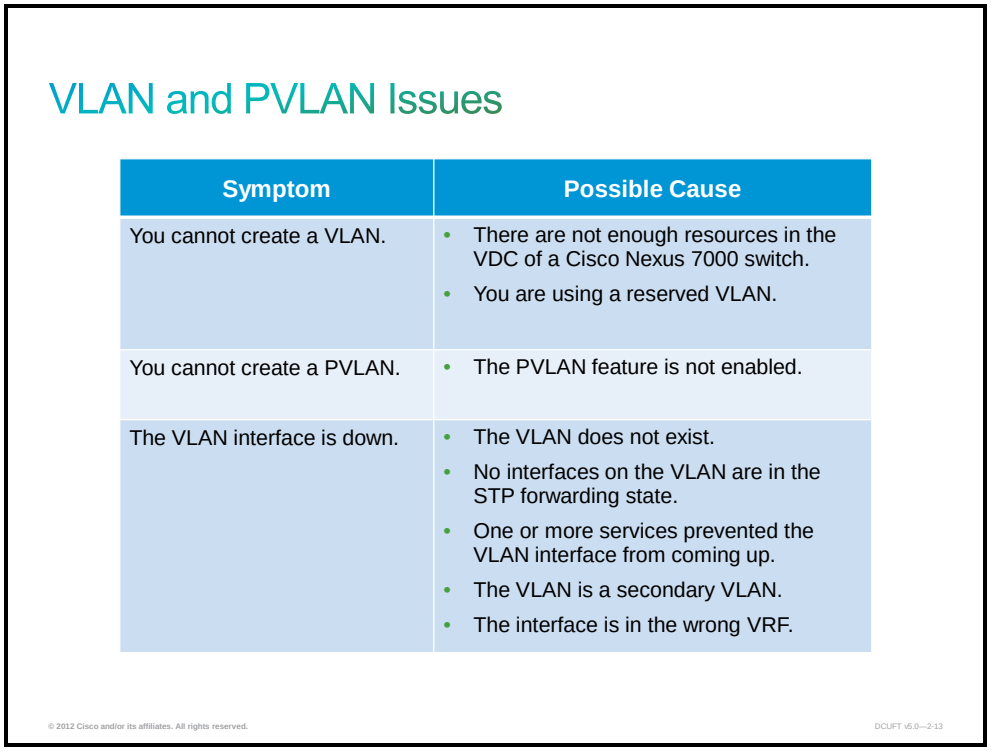
Initial Troubleshooting Checklist

- Verify the physical connectivity for any problem ports or VLANs.
- Verify that you have both end devices in the same VLAN.
- The following CLI commands are used to display VLAN information:
  - **show vlan** *vlan-id*
  - **show vlan all-ports**
  - **show tech-support vlan**

```
Switch# show vlan id 21

VLAN Name                             Status    Ports
---- -------------------------------- --------- -------------------------------

21   VLAN0021                         active

VLAN Type  SAID       MTU   Parent RingNo BridgeNo Stp  BrdgMode Trans1 Trans2
---- ----- ---------- ----- ------ ------ -------- ---- -------- ------ ------
21   enet  100021     1500  -      -      -        -    -        0      0
```


Once the VLAN configuration is completed, you can verify the VLAN parameters that are configured on the switch with the **show** commands. The fields in the **show vlan** command output are described in the following table.

Field	Description
VLAN	VLAN number
Name	Name of the VLAN, if configured
Status	Status of the VLAN (active or suspended)
Ports	Ports that belong to the VLAN
Type	Media type of the VLAN
SAID	Security association ID value for the VLAN
MTU	Maximum transmission unit size for the VLAN
Parent	Parent VLAN, if one exists
RingNo	Ring number for the VLAN, if applicable
BrdgNo	Bridge number for the VLAN, if applicable
STP	STP type used on the VLAN
BrdgMode	Bridging mode for this VLAN
Trans1	Translation bridge 1
Trans2	Translation bridge 2



The figure displays possible symptoms and causes of VLAN and PVLAN issues. The following provides more detail:

Symptom: You cannot create a VLAN.

Possible Causes:

- There are not enough resources in the VDC of a Cisco Nexus 7000 switch.
  - **Solution:** Use the **show vdc resource vlan** command to determine how many unused VLANs you can configure. If this value is 0, log in as network-admin and use the **limit-resource** command in VDC configuration mode to add more VLAN resources to this VDC. On a Nexus 5000 switch, use the similar command **show resource vlan**.
- You are using a reserved VLAN ID.
  - **Solution:** VLANs 3968 to 4047 and 4094 are reserved for internal use in each VDC; you cannot change or use these reserved VLANs.

Symptom: You cannot create a PVLAN.

Possible Cause: The PVLAN feature is not enabled.

Solution: Use the **feature private-vlan** command to enable the PVLAN feature.

Symptom: The VLAN interface is down (a problem when configuring the VLAN interface).

Possible Causes:

- The VLAN does not exist.
  - **Solution:** Use the **show vlan** command to determine if the VLAN exists. Use the **vlan** command to create the VLAN.

- No interfaces on the VLAN are in the STP forwarding state.
  - **Solution:** Use the **show vlan internal info** command to check the operating state of the STP. Configure STP so that at least one interface goes into the STP forwarding state.
- One or more services prevented the VLAN interface from coming up.
  - **Solution:** Use the **show vlan internal info** command to determine the state of the VLAN interface. If the state is “oper-es”, use the **show tech-support interface-vlan** command to gather more information.
- The VLAN is a secondary VLAN.
  - **Solution:** Use the **show vlan internal info** command to determine the state of the VLAN interface. Change the VLAN to a primary VLAN.
- The interface is in the wrong virtual routing and forwarding (VRF).
  - **Solution:** Use the **show vrf interface** command to determine the interface that the VLAN interface is assigned to.

show Commands for PVLANs

- Display the PVLAN configuration using the **show vlan private-vlan** command:

```
switch# show vlan private-vlan

Primary  Secondary  Type             Ports
-------  ---------  ---------------  ----------------------------------------
5        100        community
5        101        community        Eth1/12, veth1/1
5        102        community
5        103        community
5        109        isolated         Eth1/2
```

- Display the PVLAN type using the **show vlan private-vlan type** command:

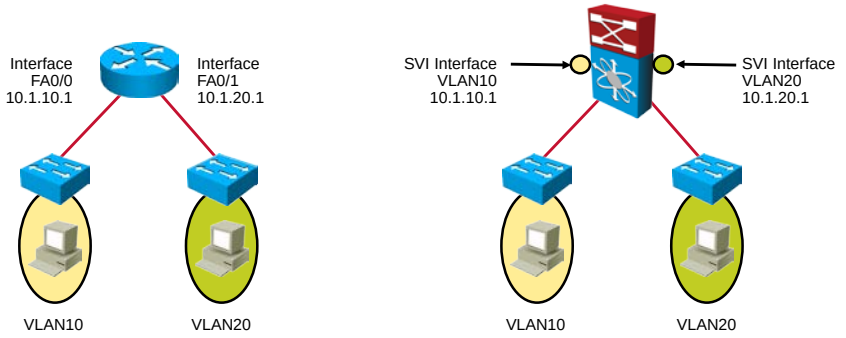
```
switch# show vlan private-vlan type

Vlan Type
---- -----------------
5    primary
100  community
101  community
102  community
103  community
109  isolated
```


To display the status of the PVLAN, use the **show vlan private-vlan** and **show vlan private-vlan type** commands.

Switch Virtual Interfaces

Routers use interfaces or subinterfaces to interconnect multiple VLANs.

Cisco Nexus 7000 switches use SVIs for routing between VLANs.



A switch virtual interface (SVI) is a virtual interface that is configured within a multilayer switch. You can create an SVI for any VLAN that exists on the switch. Only one SVI can be associated with any one VLAN. An SVI can be configured to operate at Layer 2 or Layer 3.

An SVI is “virtual” in that there is no physical port that is dedicated to the interface, yet it can perform the same functions for the VLAN as a router interface would, and can be configured in much the same way as a router interface (IP address, inbound or outbound access control lists [ACLs], and so on). The SVI for the VLAN provides Layer 3 processing for packets to and from all switch ports that are associated with that VLAN.

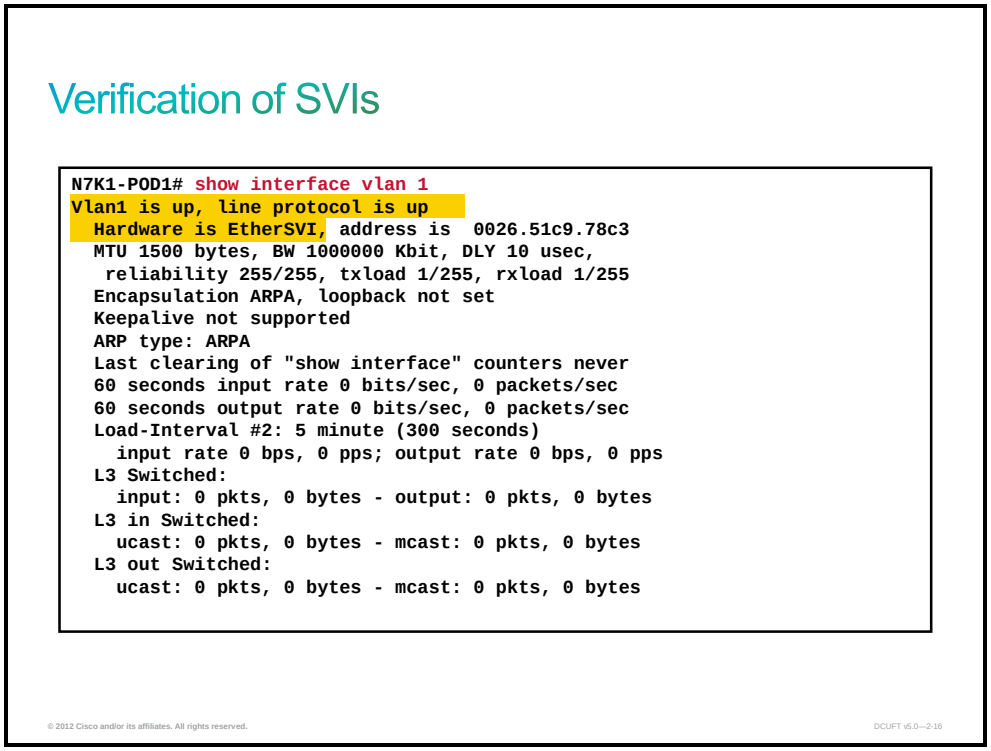
You configure an SVI for a VLAN for several reasons:

- To provide a gateway for a VLAN so that traffic can be routed into or out of that VLAN
- To provide fallback bridging if it is required for nonroutable protocols
- To provide Layer 3 IP connectivity to the switch
- To support routing protocol and bridging configurations

The **interface-vlan** feature must be enabled before configuring the SVI. Use the **show feature** command to determine which features are enabled.

Example:

```
switch(config)# feature interface-vlan
switch(config)# show feature
Feature Name Instance State
-------------------- -------- --------
tacacs 1 disabled
lacp 1 enabled
interface-vlan 1 enabled
private-vlan 1 enabled
udld 1 enabled
vpc 1 enabled
fcoe 1 disabled
fex 1 enabled
```



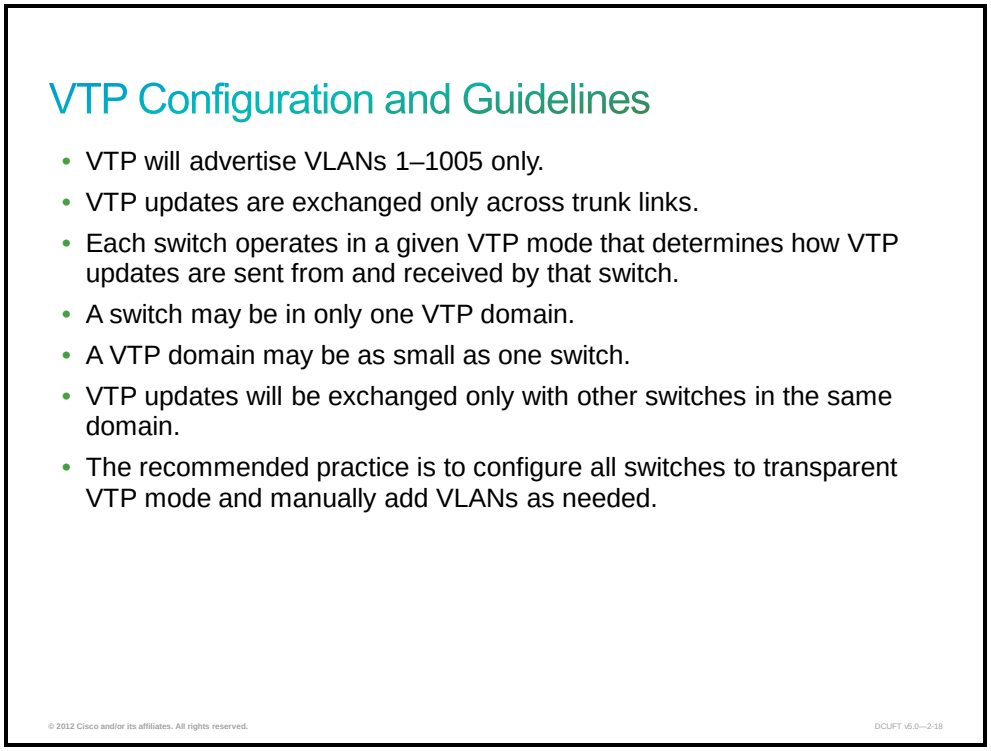
You can use the **show interfaces** command to display the interface IP address configuration and status of a Layer 3 SVI. The following is an example:

```
N7K1-POD1# show interface vlan 1
Vlan1 is up, line protocol is up
  Hardware is EtherSVI, address is  0026.51c9.78c3
  MTU 1500 bytes, BW 1000000 Kbit, DLY 10 usec,
   reliability 255/255, txload 1/255, rxload 1/255
  Encapsulation ARPA, loopback not set
  Keepalive not supported
  ARP type: ARPA
  Last clearing of "show interface" counters never
  60 seconds input rate 0 bits/sec, 0 packets/sec
  60 seconds output rate 0 bits/sec, 0 packets/sec
  Load-Interval #2: 5 minute (300 seconds)
    input rate 0 bps, 0 pps; output rate 0 bps, 0 pps
  L3 Switched:
    input: 0 pkts, 0 bytes - output: 0 pkts, 0 bytes
  L3 in Switched:
    ucast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes
  L3 out Switched:
    ucast: 0 pkts, 0 bytes - mcast: 0 pkts, 0 bytes
```

The SVI interface for VLAN 1 shows a status of up/up, because at least one port is active in VLAN 1. Note that the hardware is reported as “EtherSVI,” indicating the virtual nature of the interface. The remainder of the output is similar to what you would see on any router interface.

Troubleshooting VTP

This topic explains how troubleshoot the VLAN Trunking Protocol (VTP).



VTP allows each router or LAN device to transmit advertisements in frames on its trunk ports. These frames are sent to a multicast address where they can be received by all neighboring devices. They are not forwarded by normal bridging procedures. An advertisement lists the VTP management domain of the sending device, its configuration revision number, the VLANs that it knows about, and certain parameters for each known VLAN. By hearing these advertisements, all devices in the same management domain learn about any new VLANs that are configured in the transmitting device. This process allows you to create and configure a new VLAN only on one device in the management domain, and then that information is automatically learned by all the other devices in the same management domain.

Once a device learns about a VLAN, the device receives all frames on that VLAN from any trunk port by default, and, if appropriate, forwards them to each of its other trunk ports, if any. This process prevents unnecessary VLAN traffic from being sent to a device. An extension of VTP called VTP pruning has been defined to limit the scope of broadcast traffic and save bandwidth. Beginning with Release 5.1(1), the Cisco Nexus Operating System (NX-OS) Software supports VTP pruning on Cisco Nexus 7000 Series Switches. Cisco Nexus 5000 Series Switches do not support VTP pruning.

VTP also publishes information about the domain and the mode in a shared local database that can be read by other processes such as Cisco Discovery Protocol.

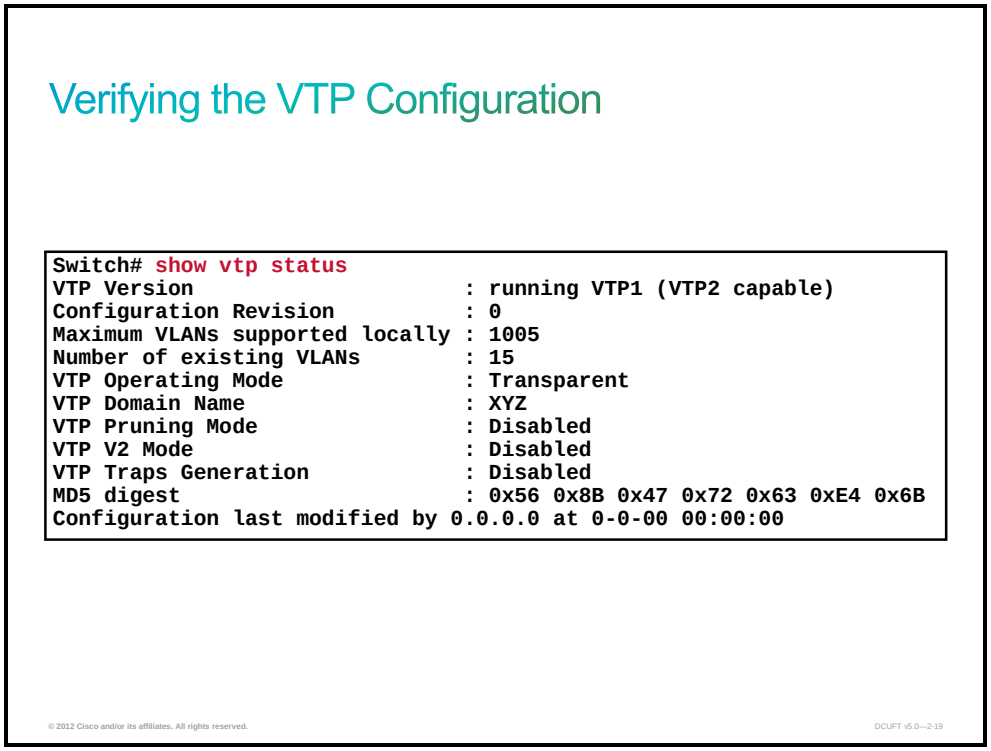
VTP is supported in the following modes: transparent, server, client, and off.

You can use a VTP client/server mode to automatically propagate VLAN definitions across the switched network. This mode is often used in a new network to facilitate the implementation of new VLANs. However, as the network grows larger, this benefit can turn into a liability. If a VLAN is deleted by accident on one server, it is deleted throughout the network. If a switch that already has a VLAN database defined is inserted into the network, it can hijack the VLAN database by deleting added VLANs. For this reason, the recommended practice is to configure all switches to transparent VTP mode and manually add VLANs as needed.

VTP has the following configuration guidelines and limitations:

- When a switch is configured as a VTP client, you cannot create VLANs on the switch in the range of 1 to 1005.
- VLAN 1 is required on all trunk ports that are used for switch interconnects if VTP is supported in the network. Disabling VLAN 1 from any of these ports prevents VTP from functioning properly.
- If you enable VTP, you must configure either version 1 or version 2. On the Cisco Nexus 5010 and Nexus 5020 switches, 512 VLANs are supported. If these switches are in a distribution network with other switches, the limit remains the same. If a Nexus 5010 switch or Cisco Nexus 5020 switch client/server receives additional VLANs from a VTP server, they transition to transparent mode.
- The **show running-configuration** command does not show VLAN or VTP configuration information for VLANs 1 to 1000. Use **show vtp status** instead.
- When deployed with a virtual port channel (vPC), both vPC switches must be configured identically. vPC performs a Type 2 consistency check for VTP configuration parameters.
- VTP advertisements are not sent out on Cisco Nexus 2000 Series Fabric Extender ports.
- When a switch is configured in VTP client or server mode, VLANs 1002 to 1005 are reserved VLANs.

On each switch, you may configure VTP to operate in one of three modes: server, client, or transparent. The default VTP mode is server. The mode will determine whether VLANs can be created on the switch and how the switch will participate in sending and receiving VTP advertisements. The number of VLANs that can be configured on a switch will vary by mode.



The **show vtp status** command is the key command for verifying VTP implementation. When initially configuring switches in a VTP domain, pay close attention to the configuration revision number. Check to see that it increases only when changes are made at intended VTP servers.

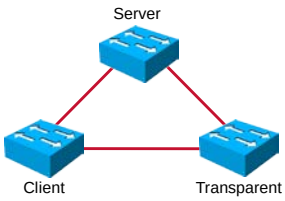
In the figure, *Configuration last modified by* specifies the IP address of the switch that last updated the VLAN database of this switch.

Note	In this example, VTP version 2 is available (as shown by the “VTP Version” line of the output) but not enabled (as shown by the “VTP V2 Mode” line of the output).

It is always advisable to use the **show vlan** command on all switches. When a switch is in client mode, it will show new VLANs only if they were created and transmitted correctly from the VTP server, because creating VLANs on a switch in client mode is not possible.

Common Problems with VTP Configuration

- Missing VLANs
  - Configuration has been overwritten by another VTP device
- Updates not received as expected
  - VTP domain and password must match
- Too many VLANs
  - Consider making VTP domain smaller

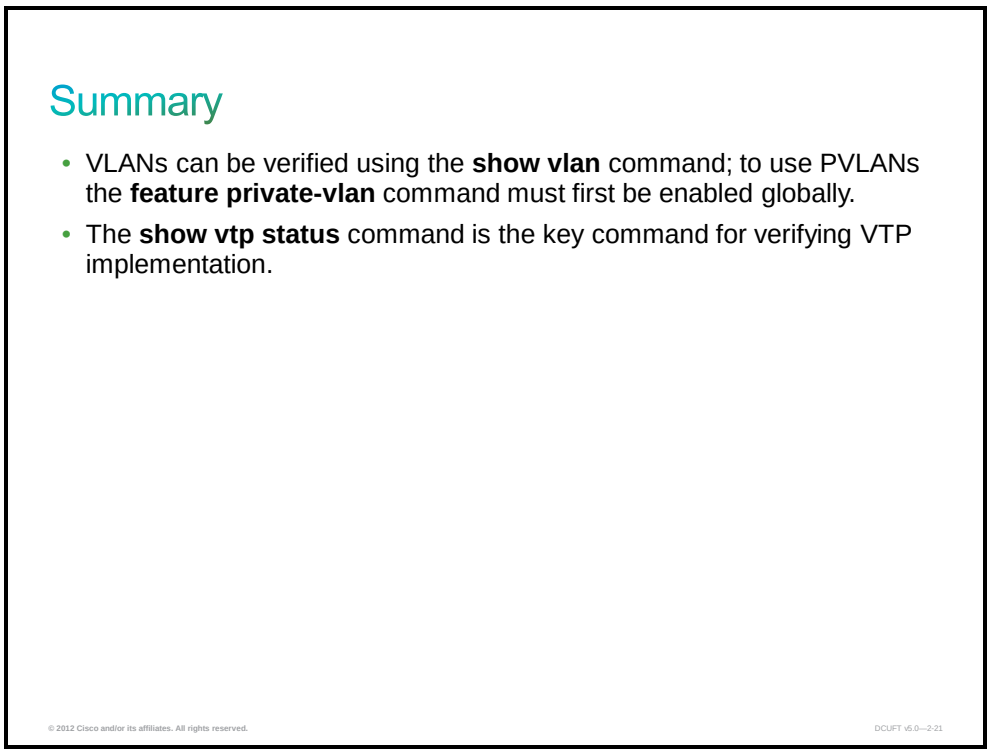


The table describes some unexpected results that can occur after VTP configuration.

Problem	Possible Causes	Action to Take
Missing VLANs	Upon initial configuration, the VTP server may have had a partial VLAN database, and it overwrote the existing, more complete database on the existing switch.	Ensure that any switch becoming a VTP server has a complete VLAN list.
	VLANs were deleted individually at the VTP server, and those deletions will be propagated in the domain.	Ensure that any switch becoming a VTP server has a complete VLAN list.
	Not all Cisco switches support the same extended-range VLANs (those numbered higher than 1005). This information is not learned or propagated through VTP, so it may vary in a switched network.	Ensure that all Cisco switches are able to use the same VLANs.
Updates not received as expected	The VTP domain name and password do not match on a given switch to receive updates from a VTP server.	The VTP domain name and password must match on a given switch to receive updates from a VTP server. The domain name is case-sensitive.
	The VTP version is incompatible with other switches in the domain.	The VTP version must be compatible with other switches in the domain.
	There is no server in the domain.	There must be at least one server in the domain.
	The link to the VTP server is not trunk.	There must be a trunk link to the VTP server.
Too many VLANs	The VTP server has a VLAN list that is more complete than the list needed by other switches in the domain.	Consider making the VTP domain smaller.

Summary

This topic summarizes the key points that were discussed in this lesson.



Lesson 2

Troubleshooting Port Channels and vPCs

Overview

This lesson is designed to provide you with some examples of issues that are related to port channels and virtual port channels (vPCs) and show you how to identify and resolve these issues.

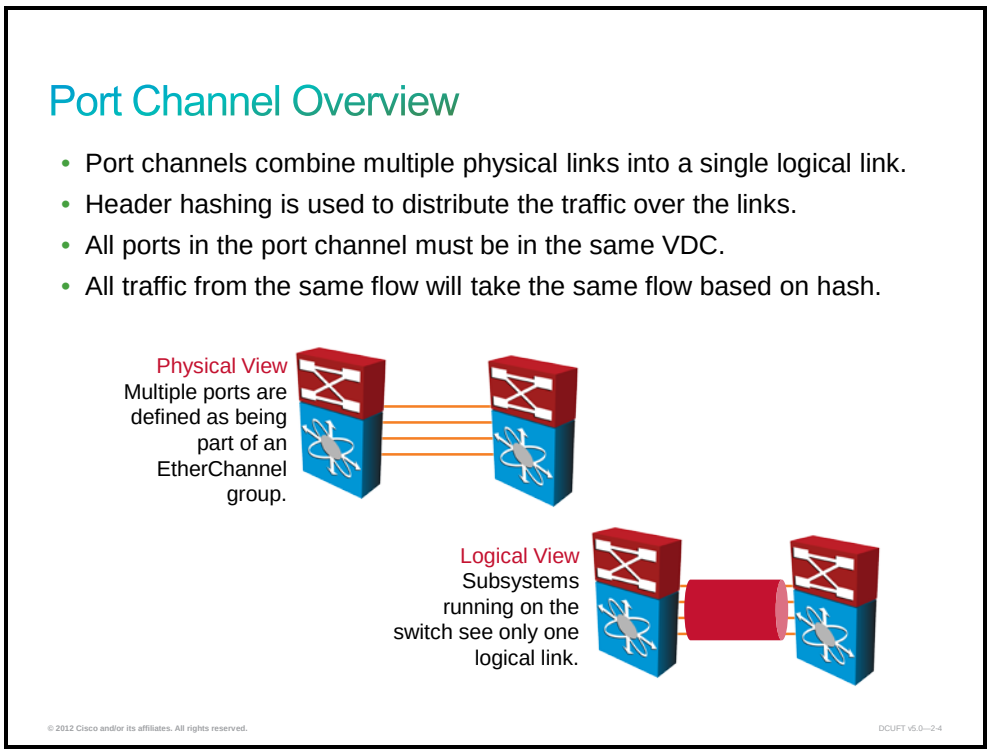
Objectives

Upon completing this lesson, you will be able to identify and resolve issues that are related to port channels and vPCs. You will be able to meet these objectives:

- Explain how to troubleshoot Ethernet port channels on a Cisco Nexus switch
- Explain how to troubleshoot LACP on a Cisco Nexus switch
- Explain how to troubleshoot vPCs on a Cisco Nexus switch

Troubleshooting Port Channels

This topic explains how to troubleshoot Ethernet port channels on a Cisco Nexus switch.



Port channels are one of the core technologies that are used in Ethernet-based networks. To add resiliency against link failures and to increase the available bandwidth between two devices, multiple physical links can be provisioned between the devices. However, without port channels, control plane protocols, such as Spanning Tree Protocol (STP) or routing protocols, treat the links as individual links. In the case of STP, this will result in blocked ports and, although the additional links add resiliency, the available bandwidth between the two devices is not increased. With routing protocols, the additional links could be used for load balancing. However, this requires a routing adjacency to be formed for every link, which increases routing protocol overhead.

The maximum number of ports in a channel depends on the exact switch hardware and software combination. On the M-series modules on the Cisco Nexus 7000 Series Switches, the maximum is eight active links per port channel. Beginning with Cisco Nexus Operating System (NX-OS) Release 5.1, you can bundle up to 16 active ports simultaneously into a port channel on the F-series modules on the Nexus 7000 Series. On the Cisco Nexus 5000 Series Switches, you can bundle up to 16 active links into a port channel.

Port channels can either be Layer 2 interfaces, or, in the case of the Cisco Nexus 5500 and 7000 switches, they can be Layer 3 interfaces. When virtual device contexts (VDCs) are used, all ports in a port channel must be in the same VDC. A port channel can either be defined statically or negotiated dynamically using Link Aggregation Control Protocol (LACP). The Cisco NX-OS Software performs a compatibility check when adding ports to a port channel to ensure that the port can participate in the port channel aggregation. Therefore, it is important that all physical ports that participate in a port channel are configured identically. All traffic from the same flow will take the same flow based on hash.

Verifying Port Channels

- Use the **show port-channel summary** command to verify port channel operation:

```
N7K-1# show port-channel summary
Flags:  D - Down        P - Up in port-channel (members)
        I - Individual  H - Hot-standby (LACP only)
        s - Suspended   r - Module-removed
        S - Switched    R - Routed
        U - Up (port-channel)
--------------------------------------------------------------------------------
Group Port-       Type     Protocol  Member Ports
      Channel
--------------------------------------------------------------------------------
1   Po1(SU)     Eth      NONE      Eth1/25(P)   Eth1/27(P)
2   Po2(SU)     Eth      LACP      Eth1/29(P)   Eth1/31(P)
3   Po3(RU)     Eth      NONE      Eth2/1(P)    Eth2/3(P)
4   Po4(RU)     Eth      LACP      Eth2/5(P)    Eth2/7(P)
```


To display port-channel configuration information, use one of the following commands:

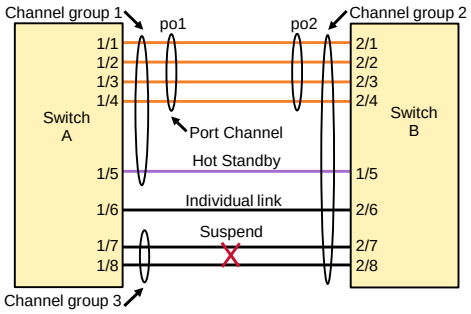
Command	Description
show interface port-channel *channel-number*	Displays the status of a port-channel interface
show feature	Displays enabled features
show port-channel compatibility-parameters	Displays the parameters that must be the same among the member ports in order to join a port channel
show port-channel database [**interface port-channel** *channel-number*]	Displays the aggregation state for one or more port-channel interfaces
show port-channel load-balance [forwarding path]	Displays the type of load balancing in use for port channels
show port-channel summary	Displays a summary for the port-channel interfaces
show port-channel traffic	Displays the traffic statistics for port channels
show port-channel usage	Displays the range of used and unused channel numbers
show lacp {**counters** [**interface port-channel** *channel-number*] | [**interface** *type*/*slot*] | **neighbor** [**interface port-channel** *channel-number*] | **port-channel** [**interface port-channel** *channel-number*] | **system-identifier**]]}	Displays information on LACP
show running-config interface port-channel *channel-number*	Displays information on the running configuration of the port-channel, and, optionally, displays the port-channel min-links configuration if there is any.

Troubleshooting LACP

This topic explains how to troubleshoot LACP on a Cisco Nexus switch.

LACP Overview

- LACP supports the automatic creation of port channels by exchanging LACP packets between LAN ports.
- The protocol learns the capabilities of LAN port groups dynamically and informs the other LAN ports.
- After LACP identifies correctly matched Ethernet links, it facilitates grouping the links into a port channel.
- LACP allows you to configure up to 16 interfaces into a port channel.



LACP supports the automatic creation of port channels by exchanging LACP packets between LAN ports. The protocol learns the capabilities of LAN port groups dynamically and informs the other LAN ports. After LACP identifies correctly matched Ethernet links, it facilitates grouping the links into a port channel.

The figure shows how individual links can be combined into LACP port channels and channel groups, as well as function as individual links.

With LACP, you can bundle up to 16 interfaces in a channel group. If the channel group has more than eight interfaces, the remaining interfaces are in hot standby for the port channel that is associated with this channel group.

Beginning with Cisco NX-OS Release 5.1, you can bundle up to 16 active links into a port channel on the F-series module.

Note	When you delete the port channel, the software automatically deletes the associated channel group. All member interfaces revert to their original configuration.

You cannot disable LACP while any LACP configurations are present.

LACP Channel Modes

Channel Mode	Description
Passive	LACP mode that places a port into a passive negotiating state in which the port responds to LACP packets that it receives but does not initiate LACP negotiation.
Active	LACP mode that places a port into an active negotiating state in which the port initiates negotiations with other ports by sending LACP packets.
On	All static port channels (that are not running LACP) remain in this mode. If you attempt to change the channel mode to active or passive before enabling LACP, the device displays an error message. The default port channel mode is on.

Individual interfaces in port channels are configured with channel modes. When you run static port channels with no aggregation protocol, the channel mode is always set to on.

After you enable LACP globally on the device, you enable LACP for each channel by setting the channel mode for each interface to active or passive. You can configure either channel mode for individual links in the LACP channel group when you are adding the links to the channel group.

Both the passive and active modes allow LACP to negotiate between ports to determine if they can form a port channel that is based on criteria such as port speed and the trunking state. The passive mode is useful when you do not know whether the remote system, or partner, supports LACP.

Ports can form an LACP port channel when they are in different LACP modes if the modes are compatible as in the following examples:

- A port in active mode can form a port channel successfully with another port that is in active mode.
- A port in active mode can form a port channel with another port that is in passive mode.
- A port in passive mode cannot form a port channel with another port that is also in passive mode, because neither port will initiate negotiation.
- A port in on mode is not running LACP and cannot form a port channel with another port that is in active or passive mode.

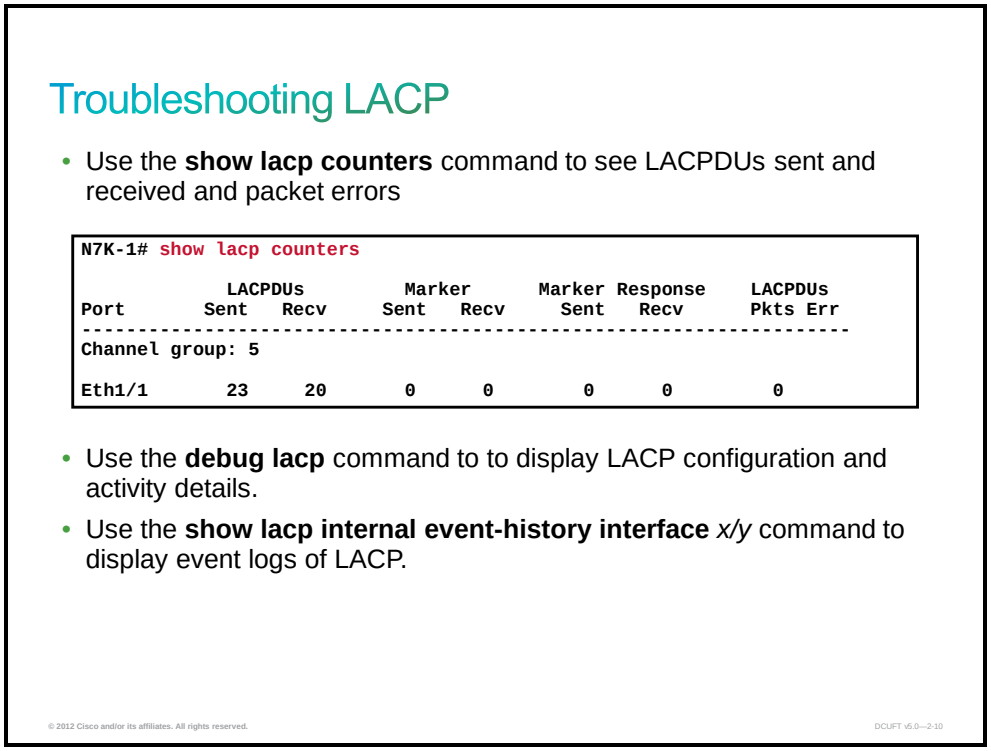
Monitoring LACP Status

- LACP is disabled by default; you must enable LACP (**feature lacp** command) before you begin any LACP configuration.
- Use the **show lacp** command to monitor LACP activity in the network.

```
N7K-1# show lacp internal

Flags:  S - Device is requesting Slow LACPDUs
        F - Device is requesting Fast LACPDUs
        A - Device is in Active mode
        P - Device is in Passive mode

Channel group 5
                              LACP port     Admin     Oper    Port      Port
Port      Flags   State       Priority      Key       Key     Number    State

Eth1/1     SA      bndl        32768       0x5       0x5     0x42       0x3D
```


LACP is disabled by default; you must enable LACP before you begin LACP configuration. You cannot disable LACP while any LACP configuration is present.

LACP learns the capabilities of LAN port groups dynamically and informs the other LAN ports. Once LACP identifies correctly matched Ethernet links, it facilitates grouping the links into a port channel.



Use the **show lacp counters** command to see Link Aggregation Control Protocol data units (LACPDUs) sent and received, as well as packet errors.

```
N7K-1# show lacp counters

              LACPDUs         Marker      Marker Response    LACPDUs
Port        Sent   Recv     Sent   Recv     Sent   Recv      Pkts Err
--------------------------------------------------------------------
Channel group: 5

Eth1/1       23     20       0      0        0      0         0
```

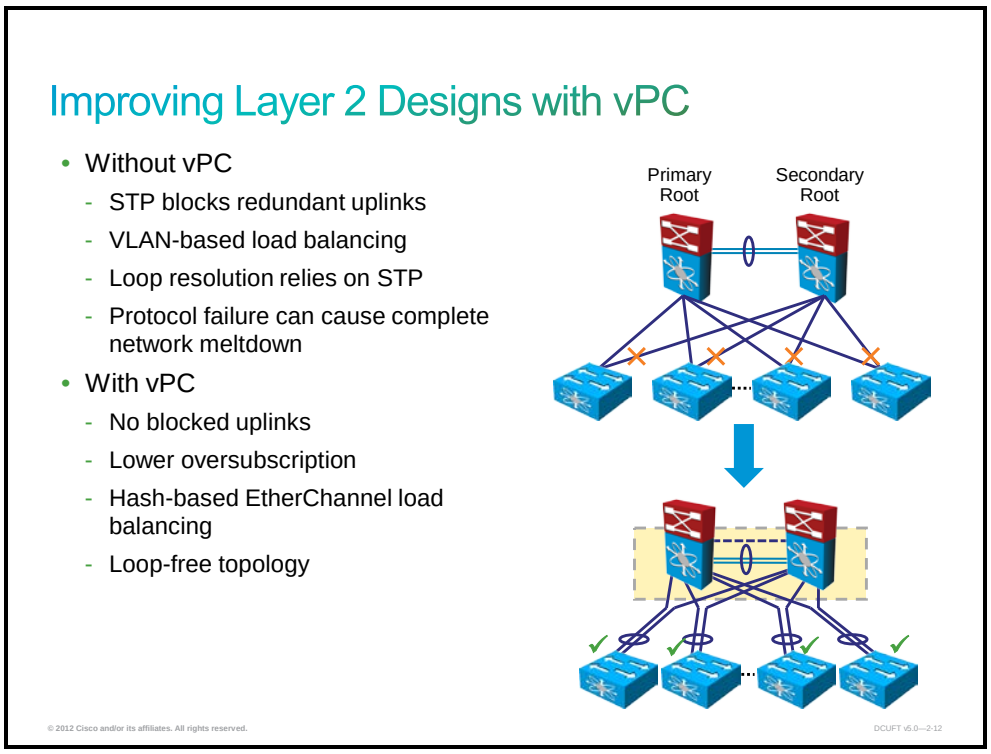
Use the **debug lacp** command to display LACP configuration and activity details. Use the **show lacp internal event-history interface** *x/y* command to display event logs of LACP on a specific interface.

The following example shows how to add an interface to a bundle and activity details:

```
N7K-1(config)# interface ethernet 1/2
N7K-1(config-if)# channel-group 5 mode active
N7K-1(config-if)#

*Mar 20 17:10:19.057: %LINK-3-UPDOWN: Interface Ethernet1/1, changed
state to down
*Mar 20 17:10:19.469: %C10K_ALARM-6-INFO: ASSERT CRITICAL Eth1/1
Physical Port Link Down
*Mar 20 17:10:19.473: %C10K_ALARM-6-INFO: CLEAR CRITICAL Eth1/1
Physical Port Link Down
*Mar 20 17:10:21.473: %LINK-3-UPDOWN: Interface Ethernet1/1, changed
state to up
*Mar 20 17:10:23.413: Ethernet1/1 added as member-1 to port-channel5
*Mar 20 17:10:23.473: %LINK-3-UPDOWN: Interface Port-channel5,
changed state to up
```

Troubleshooting vPCs

This topic explains how to troubleshoot vPCs on a Cisco Nexus switch.



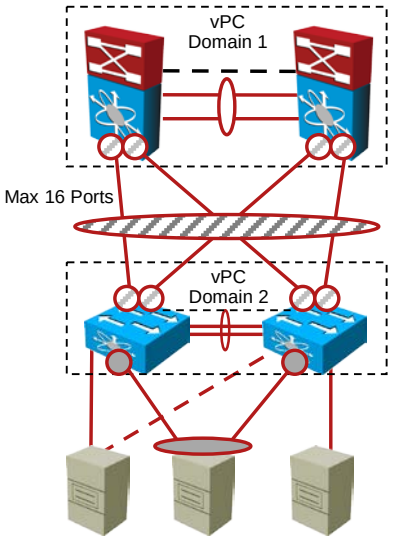
Virtualization technologies such as VMware ESX Server and clustering solutions such as Microsoft Cluster Service currently require Layer 2 Ethernet connectivity to function properly. With the increased use of these types of technologies in data centers, and now even across data center locations, organizations are shifting from a highly scalable Layer 3 network model to a highly scalable Layer 2 model. This shift is causing changes in the technologies that are used to manage large Layer 2 network environments. These changes include migration away from STP as a primary loop management technology toward new technologies such as vPC and Cisco FabricPath.

An early enhancement to Layer 2 Ethernet networks was port channel technology. This enhancement meant that multiple links between two participating devices could use all the links between the devices to forward traffic. Traffic is forwarded by using a load-balancing algorithm that equally balances traffic across the available interswitch links (ISLs), while also managing the loop problem by bundling the links as one logical link.

The biggest limitation in classic port channel communication is that the port channel operates only between two devices. In large networks, the support of multiple devices together is often a design requirement to provide some form of hardware failure alternate path. This alternate path is often connected in a way that would cause a loop, limiting the benefits that are gained with port channel technology to a single path. To address this limitation, the Cisco NX-OS Software platform provides a technology called virtual port channel (vPC). Although a pair of switches acting as a vPC peer endpoint looks like a single logical entity to port-channel-attached devices, the two devices that act as the logical port channel endpoint are still two separate devices. This environment combines the benefits of hardware redundancy with the benefits of port channel loop management. The other main benefit of migration to an all-port-channel-based loop management mechanism is that link recovery is potentially much faster. STP can recover from a link failure in approximately 6 seconds, while an all-port-channel-based solution has the potential for failure recovery in less than a second.

Double-Sided vPC

- vPC is supported on both the Cisco Nexus 5000 and Cisco Nexus 7000 Series Switches.
- vPC can be deployed in multiple layers of the data center simultaneously.
  - Server to access
  - Access to aggregation
- Double-sided vPC enables a unique 16-way port channel.
  - Can be scaled to 32-way port channels with F-series modules

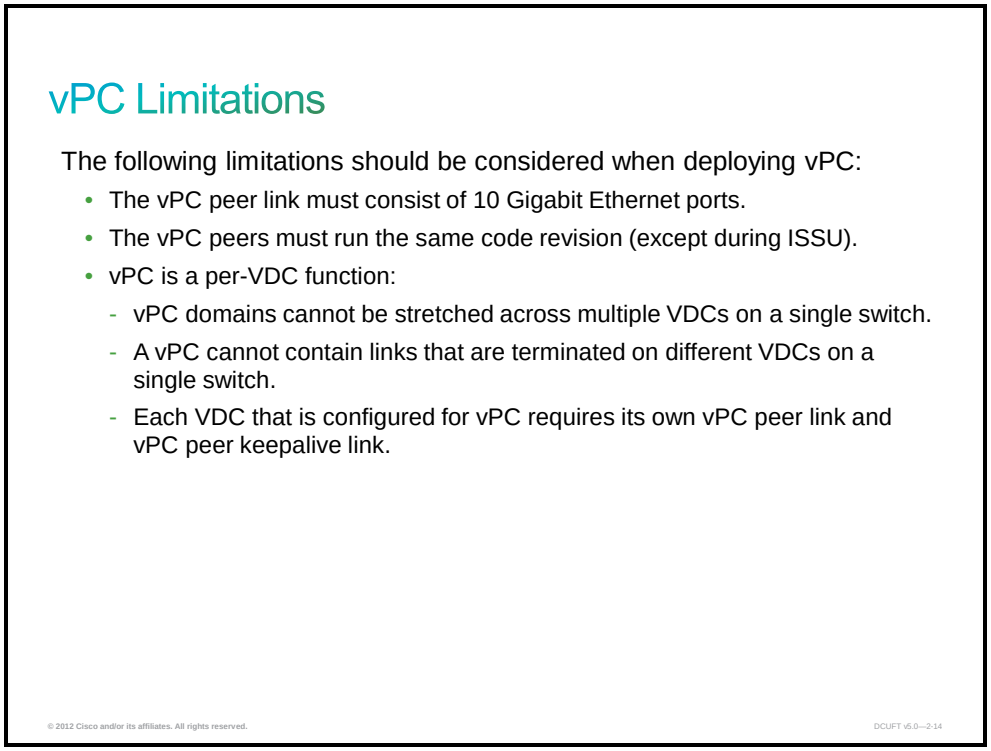


vPC is supported on both the Cisco Nexus 7000 and 5000 Series Switches. The benefits that are provided by the vPC technology apply to any Layer 2 switched domain. Therefore, vPC is commonly deployed in both the aggregation and access layers of the data center.

vPC can be used to create a loop-free logical topology between the access and aggregation layer switches, which increases the bisectional bandwidth and improves network stability and convergence. vPC can also be used between servers and the access layer switches to enable server dual-homing with dual-active connections.

When the switches in the access and aggregation layers both support vPC, a unique 16-way port channel can be created between the two layers. This scenario is commonly referred to as dual-sided vPC. This design provides up to 160 Gb/s of bandwidth from a pair of access switches to the aggregation layer.

Note	If Cisco Nexus 7000 Series Switches with F1-series modules are used on both sides of a dual-sided vPC, a 32-way port channel can be created to support up to 320 Gb/s of bandwidth between the access and aggregation layers.



When deploying vPC, there are some limitations that must be considered.

Only 10 Gigabit Ethernet ports can be used for the vPC peer link. It is recommended to use at least two 10 Gigabit Ethernet ports in dedicated mode on two different I/O modules.

The vPC peers must run the same code revision except during the nondisruptive upgrade (In-Service Software Upgrade [ISSU]).

vPC is a per-VDC function on the Cisco Nexus 7000 Series Switches. vPC can be configured in multiple VDCs, but the configuration is entirely independent. A separate vPC peer link and vPC peer keepalive link are required for each of the VDCs. vPC domains cannot be stretched across multiple VDCs on the same switch, and all ports for a given vPC must be in the same VDC.

vPC Limitations (Cont.)

The following limitations should be considered when deploying vPC:

- A vPC domain cannot consist of more than two peer switches or VDCs.
- You cannot configure more than one vPC domain per switch or VDC.
- A vPC is a Layer 2 port channel.
  - Dynamic routing to vPC peers across a vPC or across the vPC peer link is not supported.
  - Static routing across a vPC to an FHRP addresses is supported.
  - Dynamic routing across a vPC between two Layer 3 switches that are not participating in vPC is supported.
  - Dynamic routing adjacency between vPC members across a dedicated Layer 3 link between the two vPC members is supported on Cisco Nexus 5000 Series Switches.

A vPC domain, by definition, consists of a pair of switches that are identified by a shared vPC domain ID. It is not possible to add more than two switches or VDCs to a vPC domain.

Only one vPC domain ID can be configured on a single switch or VDC. It is not possible for a switch or VDC to participate in more than one vPC domain.

A vPC is a Layer 2 port channel. vPC does not support the configuration of Layer 3 port channels. Dynamic routing from the vPC peers to routers connected on a vPC is not supported. It is recommended that routing adjacencies are established on separate routed links.

Static routing to First Hop Redundancy Protocol (FHRP) addresses is supported. The FHRP enhancements for vPC enable routing to a virtual FHRP address across a vPC.

A vPC can be used as a Layer 2 link to establish a routing adjacency between two external routers. The routing restrictions for vPC only apply to routing adjacencies between the vPC peer switches and routers that are connected on a vPC on a Cisco Nexus 7000 switch. Dynamic routing adjacency between vPC members across dedicated Layer 3 link between the two VPC members is supported on a Cisco Nexus 5000 switch.

Verifying vPC

To verify vPC operation, use the **show vpc brief** command:

```
N7K-1# show vpc brief
Legend:
                (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 10
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status: success
Type-2 consistency status         : success
vPC role                          : primary
Number of vPCs configured         : 1
Peer Gateway                      : Enabled
Dual-active excluded VLANs        : -

vPC Peer-link status
---------------------------------------------------------------------
id   Port   Status Active vlans
--   ----   ------ --------------------------------------------------
1    Po20   up     100-105

vPC status
----------------------------------------------------------------------------
id   Port   Status Consistency Reason                     Active vlans
--   ----   ------ ----------- -------------------------- ------------
7    Po7    up     success     success                    100-105
```


Several commands can be used to verify the operation of vPC. The primary command to be used in initial verification is the **show vpc brief** command. This command displays the vPC domain ID, the peer-link status, the keepalive message status, whether the configuration consistency is successful, and whether a peer link is formed. It also displays the status of the individual vPCs that are configured on the switch, including the result of the consistency checks.

Verifying vPC (Cont.)

To check for potential vPC configuration consistency problems, use the **show vpc consistency-parameters** command:

```
N7K-1# show vpc consistency-parameters global

Legend:
        Type 1 : vPC will be suspended in case of mismatch

Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
STP Mode                    1     Rapid-PVST             Rapid-PVST
STP Disabled                1     None                   None
STP MST Region Name         1     ""                     ""
STP MST Region Revision     1     0                      0
STP MST Region Instance to  1
 VLAN Mapping
STP Loopguard               1     Disabled               Disabled
STP Bridge Assurance        1     Enabled                Enabled
STP Port Type, Edge         1     Normal, Disabled,      Normal, Disabled,
BPDUFilter, Edge BPDUGuard        Disabled               Disabled
STP MST Simulate PVST       1     Enabled                Enabled
Interface-vlan admin up     2     10,100-101             10,100-101
Allowed VLANs               -     100-105                100-105
Local suspended VLANs       -     -
```


If the **show vpc brief** command displays failed consistency checks, you can use the **show vpc consistency-parameters** command to find the specific parameters that caused the consistency check to fail. The **global** option on this command allows you to verify the consistency of the global parameters between the two peer switches. The **vpc** or **interface** option can be used to verify consistency between the port channel configurations for vPC member ports.

After you enable the vPC feature and configure the peer link on both vPC peer devices, Cisco Fabric Services messages provide a copy of the configuration on the local vPC peer device configuration to the remote vPC peer device. The system then determines whether any of the crucial configuration parameters differ on the two devices.

Verifying vPC (Cont.)

To check for potential vPC configuration consistency problems, use the **show vpc consistency-parameters** command:

```
N7K-1# show vpc consistency-parameters vpc 21

    Legend:
        Type 1 : vPC will be suspended in case of mismatch

Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------------------
STP Port Type               1     Default                Default
STP Port Guard              1     None                   None
STP MST Simulate PVST       1     Default                Default
lag-id                      1     [(7f9b,                [(7f9b,
                                  0-23-4-ee-be-a, 8007,  0-23-4-ee-be-a, 8007,
                                  0, 0), (8000,          0, 0), (8000,
                                  0-5-9b-1f-89-fc, 0, 0, 0-5-9b-1f-89-fc, 0, 0,
                                   0)]                    0)]
mode                        1     active                 active
Speed                       1     10 Gb/s                10 Gb/s
Duplex                      1     full                   full
Port Mode                   1     trunk                  trunk
Native Vlan                 1     1                      1
MTU                         1     1500                   1500
Allowed VLANs               -     100-104                100-104
Local suspended VLANs       -     105                    -
```


The configuration parameters in this section must be configured identically on both devices of the vPC peer link or the vPC moves into suspend mode. The devices automatically check for compatibility for some of these parameters on the vPC interfaces. The per-interface parameters must be consistent per interface, and the global parameters must be consistent globally:

- Port-channel mode: on, off, or active
- Link speed per channel
- Duplex mode per channel
- Trunk mode per channel, including native VLAN, VLANs allowed on trunk, and the tagging of native VLAN traffic
- STP mode
- STP region configuration for Multiple Spanning Tree (MST)
- Enabled or disabled state per VLAN
- STP global settings, including bridge assurance setting, port type setting, and loop guard settings
- STP interface settings, including port type setting, loop guard, and root guard
- Maximum transmission unit (MTU)

Troubleshooting vPC Issues

- vPC allows the creation of MEC between an EtherChannel-capable device and a pair of Cisco Nexus switches or VDCs.
- When you deploy vPC, you can encounter specific problems related to the use of vPC, in addition to regular port-channel or Layer 2 problems.
- Most of the problems with vPC are caused by configuration inconsistencies:
  - Global configuration inconsistencies between the pair of switches that form the vPC domain
  - vPC specific configuration inconsistencies between the pair of switches that form the vPC domain
  - Regular port channel inconsistencies

vPC allows EtherChannel-capable network devices to build a Multichassis EtherChannel (MEC) that is terminated on two different Cisco Nexus 5000 or 7000 Series Switches. The pair of vPC peer switches present themselves as a single device to the device connected on the MEC.

One of the most important aspects in troubleshooting port channels, in general, is to ensure that the configuration between the ports in the channel is consistent. This principle also applies to vPC and is complicated by the fact that the configuration now does not need to match only between ports on a single device, but also between the two vPC peer devices.

Many common problems with vPCs are caused by configuration inconsistencies. Therefore, this is always one of the primary areas to focus on when troubleshooting.

It is also possible that problems that seem to be vPC issues are actually caused by underlying Layer 1 or Layer 2 problems.

vPC Troubleshooting Checklist

Before you start any detailed examination of the vPC configuration and operation, perform these basic checks:

- Verify that the vPC peer link is configured as a port channel using ports on M-series or F-series 10 Gigabit Ethernet modules, but not mixing the types.
- Verify that all vPC member links are using ports on M-series or F-series, but not mixing the types.
- Verify that you have the OOB management interfaces connected to a management switch if you are not using a separate VRF for the peer-keepalive link.
- Verify that both the source and destination IP addresses used for the peer-keepalive messages are reachable from the VRF that is associated with the vPC peer-keepalive link.
- Verify that the peer-keepalive link is up before bringing up the vPC peer-link.
- Verify that the vPC peer-link is configured as a Layer 2 trunk, which only allows vPC VLANs.

In any troubleshooting process, it is good to perform a number of basic checks before diving too deeply into a problem. When you are troubleshooting problems that are related to vPC, begin by checking the following items first:

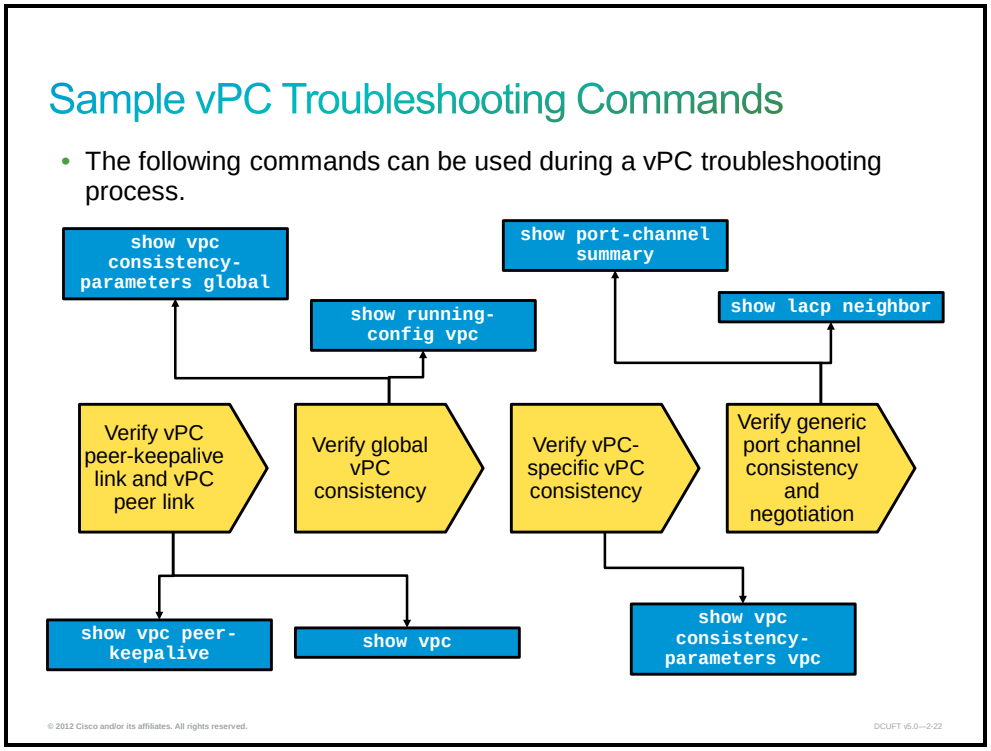
- Verify that the vPC peer link is configured as a port channel that consists either of 10 Gigabit Ethernet ports that are terminated on Cisco Nexus 7000 M-series I/O modules or of 10 Gigabit Ethernet ports that are terminated on Cisco Nexus 7000 F-series I/O modules. Either type of module can be used, but you cannot mix both types of ports in the vPC peer link port channel.
- Verify that all vPC member links are using ports on M-series or F-series on Nexus 7000 but not mixing the types.
- If you do not specify a virtual routing and forwarding (VRF) when you configure the vPC peer-keepalive link, then it will be placed in the management VRF by default. In this case, you should ensure that the out-of-band (OOB) management interfaces of the switches are connected to a separate management network.
- Verify IP connectivity between the IP addresses that are used for the vPC peer-keepalive link. When you test connectivity using the **ping** command, make sure that you specify the source IP address and VRF that are used for the peer-keepalive link.
- The peer-keepalive link needs to be configured and operational before the vPC peer link can be brought up. Ensure that the peer-keepalive link is operational before configuring the vPC peer link.
- Verify that the vPC peer link is configured as an 802.1Q trunk and that only the vPC VLANs are allowed on this trunk.

vPC Troubleshooting Checklist (Cont.)

Before you start any detailed examination of the vPC configuration and operation perform these basic checks:

- Verify that the vPC number that you assigned to the port channel that connects to the downstream device from the vPC peer device is identical on both vPC peer devices.
- If you manually configured the system priority, verify that you assigned the same priority value on both vPC peer devices.
- Use the **show vpc consistency-parameters** command to verify that both vPC peer devices have identical type 1 parameters.
- Verify that the primary vPC is the primary STP root and the secondary vPC is the secondary STP root. If you enabled the vPC peer-switch option, verify that both vPC peers are configured identically and see themselves as the root.
- Check if there are any "orphan ports;" an orphan port is any port not configured as a vPC, but carries a vPC VLAN.

- The vPC number that is used for a specific vPC should match on both vPC peer devices. Verify that the same vPC number is configured on both the port channel interfaces that are part of a single vPC. The port channel number does not need to match on both switches, but it is recommended to keep it the same and preferably make it the same as the vPC number.
- The vPC system priority needs to be the same on both vPC peer switches. If you change it on one switch, you should change it in an identical manner on the other switch.
- There are a large number of global configuration parameters that need to be the same on the two vPC peer switches for vPC to work. Use the **show vpc consistency-parameters global** command to verify that all the listed type 1 parameters match.
- Verify that the primary vPC peer switch is the spanning-tree root for all vPC VLANs and that the secondary vPC peer switch is the backup root for all vPC VLANS. If you configured the vPC peer-switch option, then both switches should be configured with the exact same spanning-tree priority. They should both see themselves as the root of the spanning tree for all vPC VLANs.
- Check if there are any "orphan ports." An orphan port is any port not configured as a vPC, but carries a vPC VLAN.

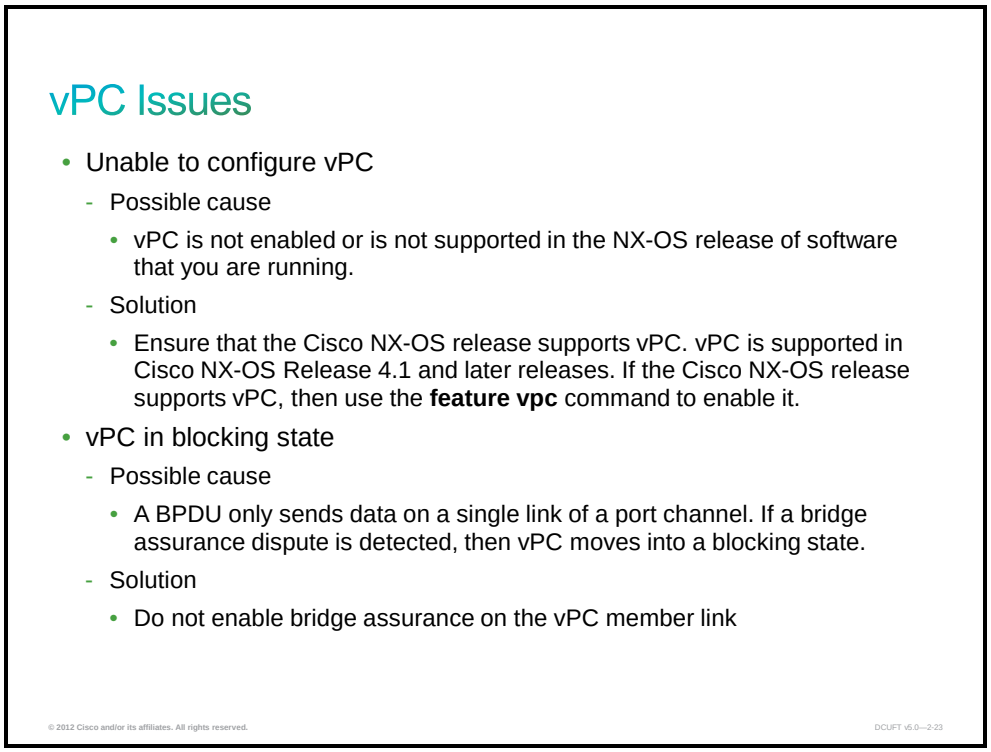


The figure illustrates a generic vPC troubleshooting process and some of the Cisco NX-OS CLI commands that can be used during that process. This list is not exhaustive and there are many more commands that could be useful during the vPC troubleshooting process.

The most commonly used Cisco NX-OS vPC troubleshooting commands are the following:

- **show vpc peer-keepalive:** This command displays the state of the peer-keepalive link, which must be operational before the vPC peer link can come up.
- **show vpc:** This command can be used to verify that the vPC peer link is operational in addition to viewing the global vPC parameters.
- **show vpc consistency-parameters global:** This command displays all the relevant global parameters that need to match on the vPC peers. If there are global configuration inconsistencies, you should be able to spot them in the output of this command. This command will not yield any useful results until the peer-link has been established and consistency checks have been performed.
- **show running-config vpc:** This command is useful to compare the vPC configuration on both peer switches and spot potential configuration errors and inconsistencies. By adding the **vpc** keyword to the **show running-config** command, any non-vPC-related commands are left out of the configuration.
- **show vpc consistency parameters vpc:** This command can be used to verify that the configuration on the vPC port-channels is consistent on both vPC peer switches. If the configuration on both peers is not consistent, the vPC will be suspended.
- **show port-channel summary:** This command can be used to verify the state of the port channel interfaces, both on the vPC peer switches and on the connected downstream device.
- **show lacp neighbors:** This command can help to troubleshoot the negotiation of port channels that are using LACP. Verifying the LACP system identifiers can help to verify that the individual links in the port channel are connected to the correct switches.

In addition to these vPC specific troubleshooting commands, you may need to use Layer 1 and Layer 2 troubleshooting commands to troubleshoot problems with individual links.

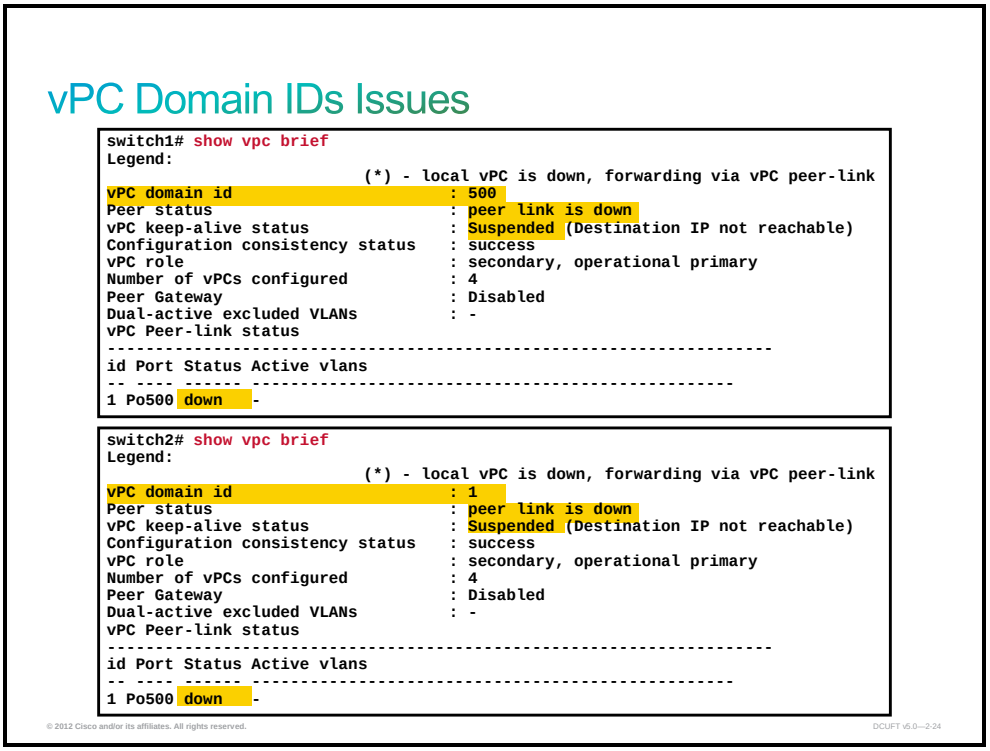


If you are unable to configure vPC, a possible cause is that vPC is not enabled or is not supported in the Cisco NX-OS release of software that you are running.

Ensure that the Cisco NX-OS release supports vPC. vPC is supported in Cisco NX-OS Release 4.1 and later. If the Cisco NX-OS release supports vPC, then use the command feature of vPC to enable it.

A bridge protocol data unit (BPDU) only sends data on a single link of a port channel. If a bridge assurance dispute is detected, then vPC moves into a blocking state. To solve this, do not enable bridge assurance on the vPC link because of the following:

- Bridge assurance cannot be used on a spanning tree port type network.
- Bridge assurance prevents you from encountering ISSU issues. It should only be enabled on the vPC peer link.



The vPC peer link is down if the vPC domain IDs are configured improperly. Compare the vPC domain IDs of the two switches and ensure that they match.

Example:

```
switch1# show vpc brief
Legend:     (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id : 500
Peer status : peer link is down
vPC keep-alive status : Suspended (Destination IP not reachable)
Configuration consistency status: success
vPC role : secondary, operational primary
Number of vPCs configured : 4
Peer Gateway : Disabled
Dual-active excluded VLANs : -
vPC Peer-link status
--------------------------------------------------------------------------
id Port Status Active vlans
-- ---- ------ --------------------------------------------------

1 Po500 down -

switch2# show vpc brief
Legend:     (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id : 1
Peer status : peer link is down
vPC keep-alive status : Suspended (Destination IP not reachable)
Configuration consistency status: success
vPC role : primary
Number of vPCs configured : 4
Peer Gateway : Disabled
Dual-active excluded VLANs : -
vPC Peer-link status
--------------------------------------------------------------------------
id Port Status Active vlans
-- ---- ------ --------------------------------------------------

1 Po500 down -
```


vPC Domain IDs Issues (Cont.)

The vPC domain IDs of these Cisco Nexus switches must be changed to match.

```
switch2(config)# vpc domain 500
Changing domain id will flap peer-link and vPCs. Continue (yes/no)? [no] yes
Note:
--------:: Re-init of peer-link and vPCs started ::--------

switch2# show vpc brief
Legend:
                        (*) - local vPC is down, forwarding via vPC peer-link

vPC domain id                     : 500
Peer status                       : peer adjacency formed ok
vPC keep-alive status             : peer is alive
Configuration consistency status  : success
vPC role                          : primary, operational secondary
Number of vPCs configured         : 4
Peer Gateway                      : Disabled
Dual-active excluded VLANs        : -

vPC Peer-link status
---------------------------------------------------------------------
id Port Status Active vlans
-- ---- ------ --------------------------------------------------
1 Po500 up     1,19,91,99,757
```


The two switches in this example have different vPC domain IDs. The vPC domain IDs of these Cisco Nexus switches must be changed to match. This can be done by entering configuration commands, one per line, and ending each by pressing Ctrl-Z.

```
switch2(config)# vpc domain 500
Changing domain id will flap peer-link and vPCs. Continue (yes/no)?
[no] yes
Note:
--------:: Re-init of peer-link and vPCs started ::--------
switch2# show vpc brief
Legend:     (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id : 500
Peer status : peer adjacency formed ok
vPC keep-alive status : peer is alive
Configuration consistency status: success
vPC role : primary, operational secondary
Number of vPCs configured : 4
Peer Gateway : Disabled
Dual-active excluded VLANs : -
vPC Peer-link status
---------------------------------------------------------------------
-
id Port Status Active vlans
-- ---- ------ --------------------------------------------------
1 Po500 up 1,19,91,99,757
```

vPC Connectivity Issues

- Possible cause
  - vPC peer-keepalive link and connectivity issues over mgmt0 might exist.
- Solution
  - Check for the peer-keepalive mgmt0 reachability.

```
switch2# sh run int mgmt 0

!Command: show running-config interface mgmt0
!Time: Thu Mar 8 03:20:58 2012

version 5.1(3)N2(1)

interface mgmt0
  ip address 172.18.118.162/24
```

  - Ensure there is reachability from the peer switch

```
switch1# ping 172.18.118.162 vrf management
PING 172.18.118.162 (172.18.118.162): 56 data bytes
64 bytes from 172.18.118.162: icmp_seq=0 ttl=254 time=5.306 ms
64 bytes from 172.18.118.162: icmp_seq=1 ttl=254 time=3.963 ms
64 bytes from 172.18.118.162: icmp_seq=2 ttl=254 time=4.04 ms
64 bytes from 172.18.118.162: icmp_seq=3 ttl=254 time=4.077 ms
64 bytes from 172.18.118.162: icmp_seq=4 ttl=254 time=4.057 ms
```


If there are connectivity issues, you should first check for the peer keepalive mgmt0 reachability.

On the peer Cisco Nexus switch, enter the command:

```
switch2# sh run int mgmt 0
!Command: show running-config interface mgmt0
!Time: Tue Mar 8 03:20:58 2011
version 4.2(1)N2(1)
interface mgmt0
ip address 172.18.118.162/24
```

Ensure there is reachability from switch1:

```
switch1# ping 172.18.118.162 vrf management
PING 172.18.118.162 (172.18.118.162): 56 data bytes
64 bytes from 172.18.118.162: icmp_seq=0 ttl=254 time=5.306 ms
64 bytes from 172.18.118.162: icmp_seq=1 ttl=254 time=3.963 ms
64 bytes from 172.18.118.162: icmp_seq=2 ttl=254 time=4.04 ms
64 bytes from 172.18.118.162: icmp_seq=3 ttl=254 time=4.077 ms
64 bytes from 172.18.118.162: icmp_seq=4 ttl=254 time=4.057 ms
```

If the ping fails, it means that the connectivity between both mgmt0 interfaces does not exist or that they are not interconnected properly. Make sure the mgmt0 interface is unshut and that you can ping the switch mgmt0 interface.

```
switch1# sh int br | grep mgmt0
mgmt0 -- down 172.16.118.62 -- 1500
```

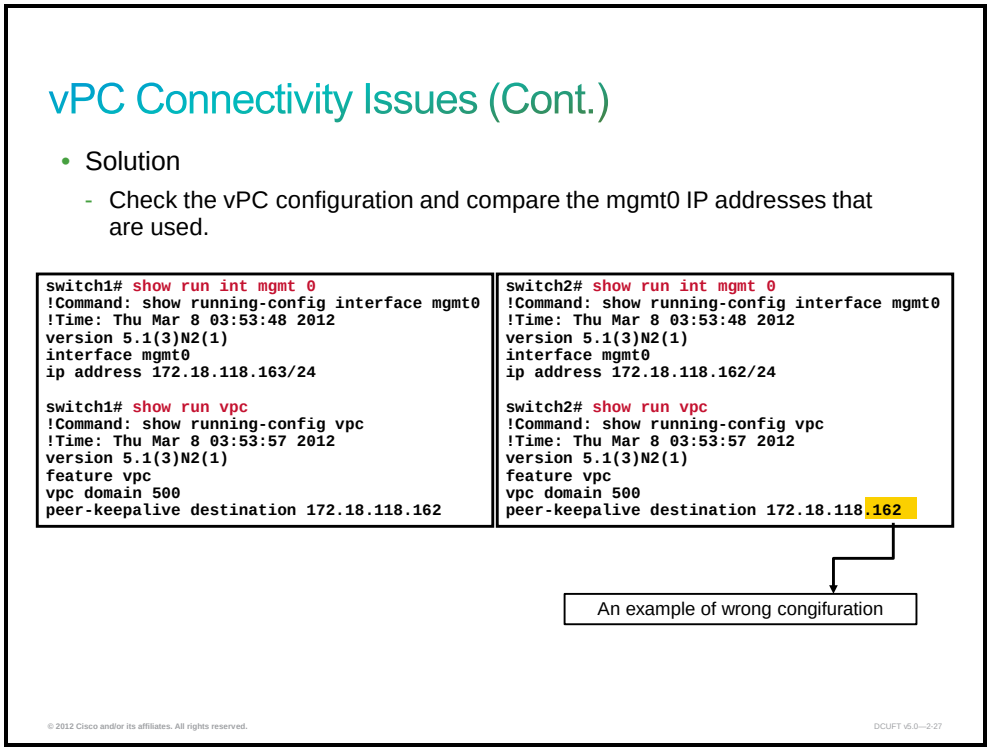
If the status shows that it is down, it means there is no physical connection to mgmt0 or that the interface is in administrative shutdown. You need to verify the physical connectivity and unshut the port:

```
switch1# config t
switch1(config)# int mgmt 0
switch1(config-if)# no shut
switch1(config-if)# show int br | grep mgmt0
mgmt0 -- up 172.16.118.62 1000 1500
```

If pinging the other switch continues to fail, then there is an interconnection issue between the two Cisco Nexus switches.

Check the networking in between the switches:

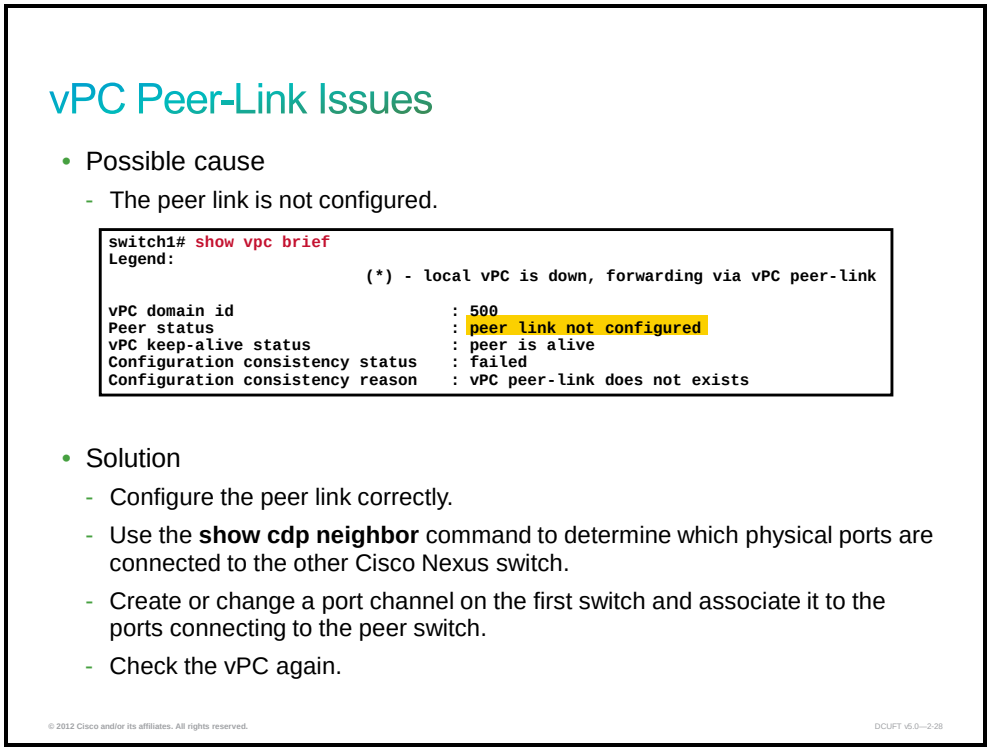
- Switch interconnecting in access VLAN mode, using the same VLAN for both Nexus switches.
- The VLAN is allowed across and between the switches.



If there are still connectivity issues, check the vPC configuration and compare the mgmt0 IP addresses that are used:

```
switch1# show run int mgmt 0
!Command: show running-config interface mgmt0
!Time: Tue Mar 8 03:53:48 2011
version 4.2(1)N2(1)
interface mgmt0
ip address 172.18.118.163/24

switch1# show run vpc
!Command: show running-config vpc
!Time: Tue Mar 8 03:53:57 2011
version 4.2(1)N2(1)
feature vpc
vpc domain 500
peer-keepalive destination 172.18.118.162

switch2# show run int mgmt 0
!Command: show running-config interface mgmt0
!Time: Tue Mar 8 03:53:53 2011
version 4.2(1)N2(1)
interface mgmt0
ip address 172.18.118.162/24

switch2# sh run vpc
!Command: show running-config vpc
!Time: Tue Mar 8 03:54:01 2011
version 4.2(1)N2(1)
feature vpc
vpc domain 500
peer-keepalive destination 172.18.118.162
```

In this example, the destination IP is not correct. The correct IP is 172.18.118.163, which is the peer IP address.



If you encounter peer-link issues, check and verify the peer link configuration. In the following example, the problem is that the vPC peer link does not exist:

```
switch1# show vpc brief
Legend:     (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id : 500
Peer status : peer link not configured
vPC keep-alive status : peer is alive
Configuration consistency status: failed
Configuration consistency reason: vPC peer-link does not exists
```

You can use the **show cdp neighbor** command to determine which physical ports are connected to the other Cisco Nexus switch.

```
switch1# show cdp neighbor
Capability Codes: R - Router, T - Trans-Bridge, B - Source-Route-
Bridge
S - Switch, H - Host, I - IGMP, r - Repeater,
V - VoIP-Phone, D - Remotely-Managed-Device,
s - Supports-STP-Dispute
Device-ID Local Intrfce Hldtme Capability Platform Port ID
switch2(SSI1324033X)Eth1/25 128 S I s N5K-C5020P-BF Eth1/25
switch2(SSI1324033X)Eth1/26 128 S I s N5K-C5020P-BF Eth1/26
```

In this example, ports 25 and 26 connect to the other Cisco Nexus switch and should be configured as a peer link. Run the same command on the other Cisco Nexus switch and observe the ports.

```
switch2# show cdp neighbor
Capability Codes: R - Router, T - Trans-Bridge, B - Source-Route-
Bridge
S - Switch, H - Host, I - IGMP, r - Repeater,
V - VoIP-Phone, D - Remotely-Managed-Device,
s - Supports-STP-Dispute
Device-ID Local Intrfce Hldtme Capability Platform Port ID
```

```
switch1(SSI14150768)Eth1/25 168 S I s N5K-C5020P-BF Eth1/25
switch1(SSI14150768)Eth1/26 168 S I s N5K-C5020P-BF Eth1/26

switch2# show run int e1/25
!Command: show running-config interface Ethernet1/25
version 5.1(3)N2(1)
interface Ethernet1/25
switchport mode trunk
channel-group 500

switch2# show run int e1/26
!Command: show running-config interface Ethernet1/26
version 5.1(3)N2(1)
interface Ethernet1/26
switchport mode trunk
channel-group 500
```

In this example, you can see that port channel 500 is used on the connection to switch1 on switch2. You now need to determine how port channel 500 is configured on switch2.

```
switch2# show run int po 500
!Command: show running-config interface port-channel500
!Time: Tue Mar 8 04:10:38 2011
version 4.2(1)N2(1)
interface port-channel500
switchport mode trunk
vpc peer-link
spanning-tree port type network
speed 10000
```

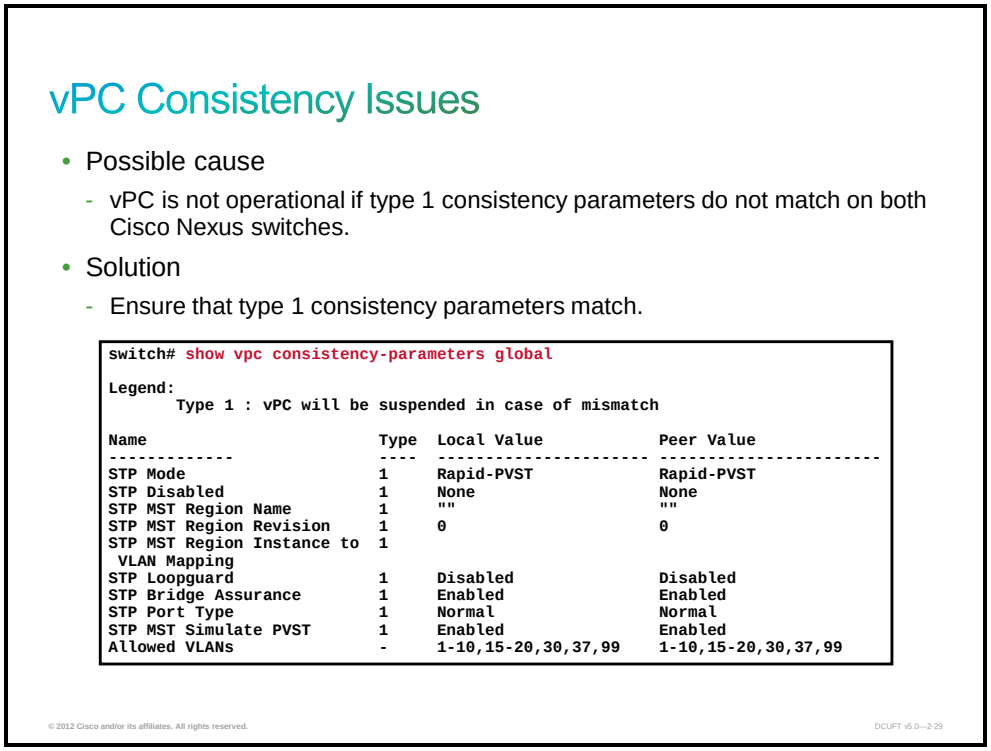
Create a port channel 500 on switch1 and associate it to the ports connecting to e1/25 and e1/26 on switch2.

```
switch1(config)# int po 500
switch1(config-if)# int e1/25-26
switch1(config-if-range)# channel-group 500
switch1(config-if-range)# int po 500
switch1(config-if)# vpc peer-link
```

Notice that the spanning tree port type has changed to a network port type on the vPC peer link. This enables spanning tree bridge assurance on the vPC peer link, provided that STP bridge assurance is not disabled. (STP bridge assurance is enabled by default.) Check the vPC again.

```
switch1(config-if)# show vpc brief
Legend:     (*) - local vPC is down, forwarding via vPC peer-link
vPC domain id : 500
Peer status : peer adjacency formed ok
vPC keep-alive status : peer is alive
Configuration consistency status: success
vPC role : primary
Number of vPCs configured : 4
Peer Gateway : Disabled
Dual-active excluded VLANs : -
vPC Peer-link status
---------------------------------------------------------------------
id Port Status Active vlans
-- ---- ------ --------------------------------------------------
1 Po500 up 1,19,91,99,757
```

Port channel 500 and the peer link are now up. The vPC is successful.



vPC is not operational if type 1 consistency parameters do not match on both Cisco Nexus 5000 switches. You have to ensure that type 1 consistency parameters match.

The possible values for type are 1, 2, or -. Items that are type 1 must match on both Cisco Nexus switches. If they do not match, then vPC is suspended. Starting with Release 5.0 on the Nexus 5000 Series, a type 2 was introduced. Items that are type 2 do not have to match on both Nexus 5000 switches for the vPC to be operational.

The command in the following example displays local and peer values. Run the command on both switches to ensure that the type 1 items match. To check for a mismatch, display the consistency parameters.

```
switch# show vpc consistency-parameters global

Legend:
       Type 1 : vPC will be suspended in case of mismatch
Name                        Type  Local Value            Peer Value
-------------               ----  ---------------------- -----------
------------
STP Mode                    1     Rapid-PVST             Rapid-PVST
STP Disabled                1     None                   None
STP MST Region Name         1     ""                     ""
STP MST Region Revision     1     0                      0
STP MST Region Instance to  1
 VLAN Mapping
STP Loopguard               1     Disabled               Disabled
STP Bridge Assurance        1     Enabled                Enabled
STP Port Type               1     Normal                 Normal
STP MST Simulate PVST       1     Enabled                Enabled
Allowed VLANs               -     1-10,15-20,30,37,99    1-10,15-
20,30,37,99
```

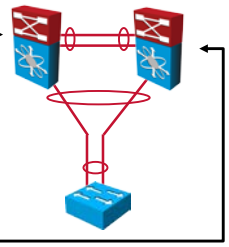
In this example, all values match and vPC will be operational.

vPC Consistency Issues (Cont.)

- Type 2 consistency checks are meant to prevent undesired forwarding.
- Solution
  - Ensure that there is no VLAN mismatch.

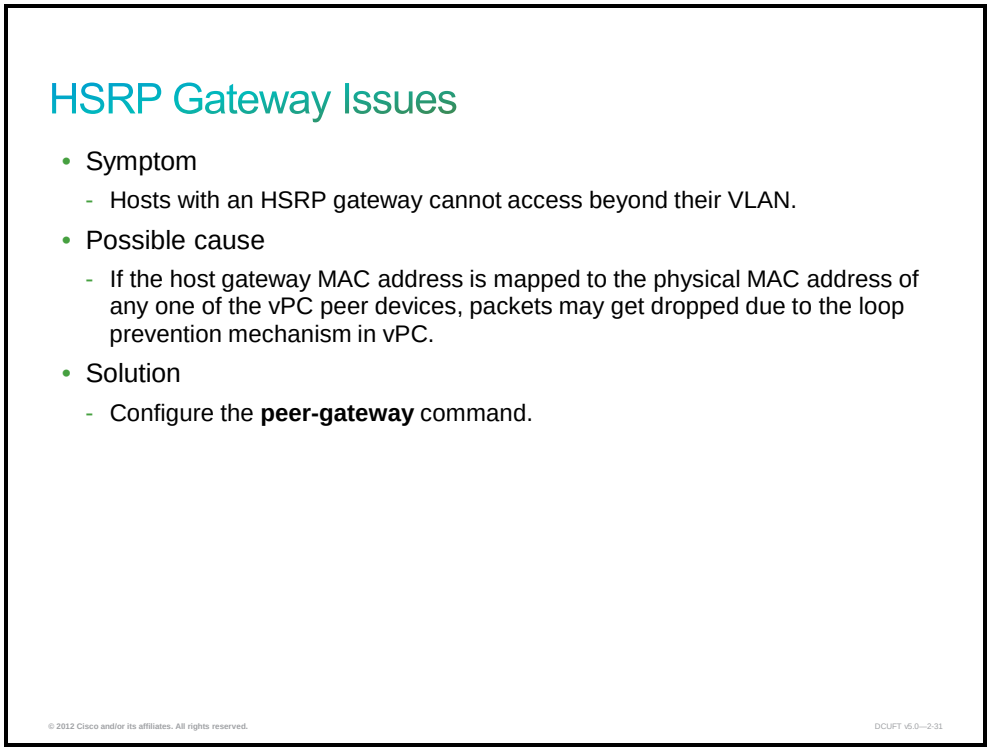
```
switch1# show run int po 201

interface port-channel201
switchport trunk allowed vlan 100-105
vpc 201
spanning-tree port type network
```

```
switch2# show run int po 201

interface port-channel201
switchport trunk allowed vlan 100-104
vpc 201
spanning-tree port type network
```



```
switch1# show log

2012 Apr 11 21:56:28 switch1 %ETHPORT-5-IF_ERROR_VLANS_SUSPENDED: VLANs 105
on Interface port-channel201 are being suspended. (Reason: Vlan is not
configured on remote vPC interface)
```


Depending on the severity of the misconfiguration, vPC may either warn the user (type 2 misconfiguration) or suspend the port channel (type 1 misconfiguration). In the specific case of a VLAN mismatch, only the VLAN that differs between the vPC member ports is going to be suspended on all the vPC port channels.



When Hot Standby Router Protocol (HSRP) is enabled on both vPC peer devices on a VLAN and hosts on that VLAN set the HSRP as their gateway, they may not able to reach anything outside their own VLAN.

If the host gateway MAC address is mapped to the physical MAC address of any one of the vPC peer devices, packets may get dropped due to the loop prevention mechanism in vPC.

The solution is to map the host gateway MAC address to the HSRP MAC address and not the physical MAC address of any one of the vPC peer devices. Using the **peer-gateway** command can be a workaround for this scenario.

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- Use the **show interface port-channel** *channel-number* command to display the status of a port channel interface and the **show port-channel load-balance** command to display the type of load balancing in use for port channels.
- Use the **show lacp** command to monitor LACP activity in the network; use the **debug lacp** command to display LACP configuration and activity details.
- Before you start any detailed examination of the vPC configuration and operation, perform some basic checks. To verify vPC operation, use the **show vpc brief** command.

Lesson 3

Troubleshooting Cisco FabricPath

Overview

This lesson is designed to provide you with some examples of issues that are related to Cisco FabricPath and show you how to identify and resolve these issues.

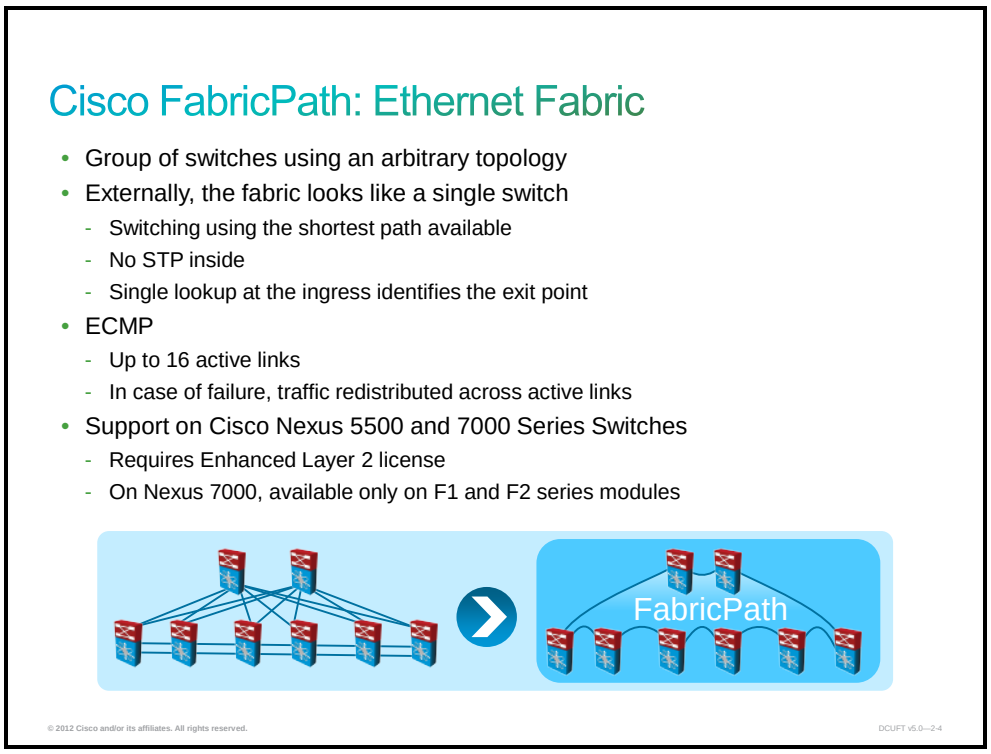
Objectives

Upon completing this lesson, you will be able to identify and resolve issues that are related to Cisco FabricPath. You will be able to meet these objectives:

- Describe the Cisco FabricPath control plane
- Describe the Cisco FabricPath data plane
- Explain how to troubleshoot Cisco FabricPath on a Cisco Nexus switch

Cisco FabricPath Control Plane

This topic describes the Cisco FabricPath control plane.



Cisco FabricPath is an innovative Cisco Nexus Operating System (NX-OS) feature designed to bring the stability and performance of routing to Layer 2. It brings the benefits of Layer 3 routing to Layer 2 switched networks to build a highly resilient and scalable Layer 2 fabric.

Cisco FabricPath switching allows multipath networking at the Layer 2 level. The Cisco FabricPath network still delivers packets on a best-effort basis (which is similar to the classic Ethernet network), but the Cisco FabricPath network can use multiple paths for Layer 2 traffic. In a Cisco FabricPath network, you do not need to run the Spanning Tree Protocol (STP). Instead, you can use Cisco FabricPath across data centers.

Externally, a fabric looks like a single switch, yet internally there is a protocol that adds fabric-side intelligence. This intelligence ties the elements of the Cisco FabricPath infrastructure together.

Frames are forwarded along the shortest path to their destination, reducing the latency of the exchanges between end stations when compared to a spanning tree-based solution.

Every interface that is involved in Cisco FabricPath switching falls into one of two categories:

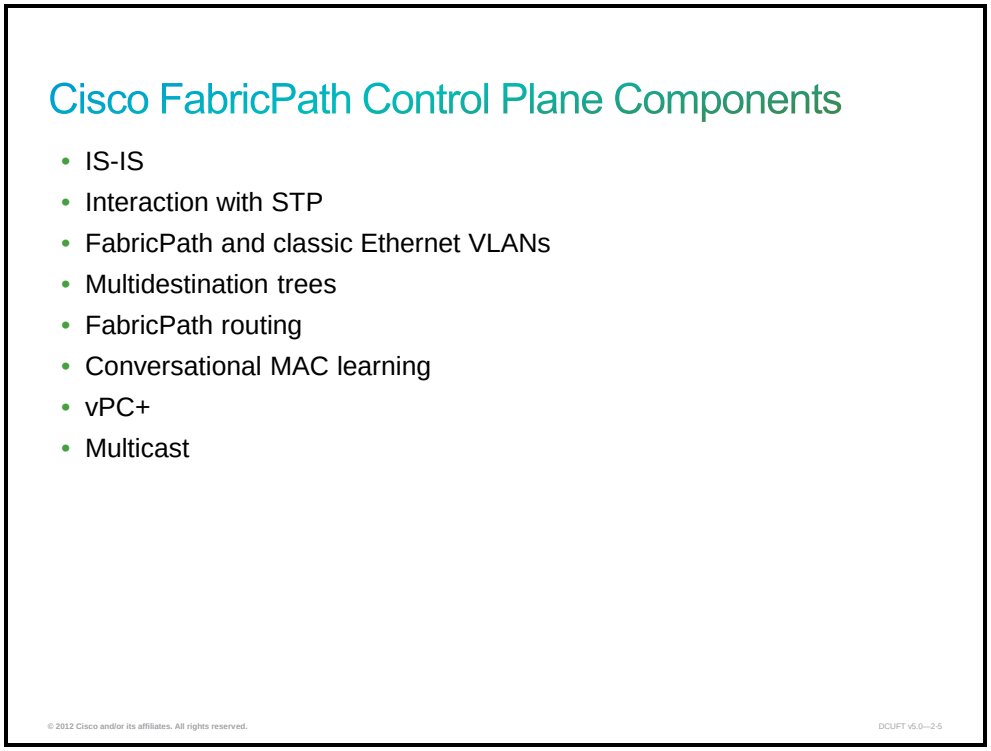
- **Cisco FabricPath edge port:** Cisco FabricPath edge ports are interfaces at the edge of the Cisco FabricPath domain. These interfaces run classic Ethernet and behave exactly like normal Ethernet ports. You can attach any classic Ethernet device to the Cisco FabricPath fabric by connecting it to a Cisco FabricPath edge port. Cisco FabricPath switches perform MAC address learning on edge ports, and frames that are transmitted on edge ports are standard IEEE 802.3 Ethernet frames. You can configure an edge port as an access port or as an IEEE 802.1Q trunk.

- **Cisco FabricPath core port:** Cisco FabricPath core ports always forward Ethernet frames encapsulated in a Cisco FabricPath header. As a rule, no MAC address learning occurs on Cisco FabricPath core ports; forwarding decisions occur based exclusively on lookups in the switch table. Ethernet frames transmitted on a Cisco FabricPath interface always carry an IEEE 802.1Q tag, and therefore the port can conceptually be considered a trunk port.

In Cisco FabricPath topologies, there are two types of "functions" (which can be performed by all Cisco FabricPath hardware):

- **Edge (or leaf) devices:** These devices have ports that are connected to Classic Ethernet devices (servers, firewalls, router ports, and so on) and ports that are connected to the Cisco FabricPath cloud (or FabricPath ports). Edge devices are able to map a MAC address to the destination switch ID.
- **Spine devices:** These devices exclusively interconnect edge devices. Spine devices switch exclusively based on the destination switch ID.

Because Equal-Cost Multipath (ECMP) can be used at the data plane, the network can use all the links that are available between any two devices. Cisco FabricPath can perform 16-way ECMP.



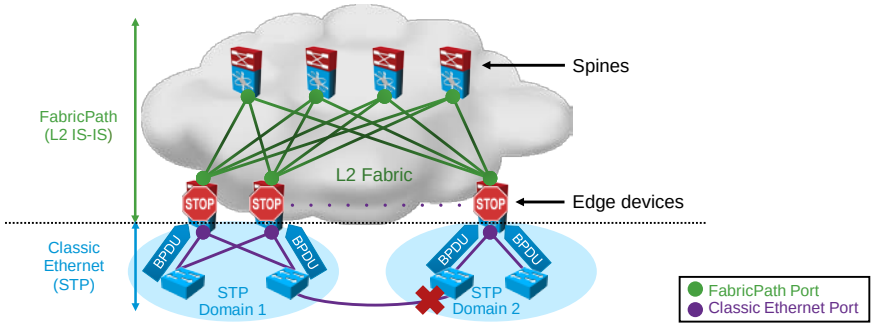
These are the Cisco FabricPath control plane components:

- **IS-IS:** Cisco FabricPath uses the Layer 2 Intermediate System-to-Intermediate System (IS-IS) protocol for a single control plane that functions for unicast, broadcast, and multicast packets.
- **Interaction with STP:** A detailed description is provided later in this lesson. The STP domains do not cross into the FabricPath network.
- **Cisco FabricPath and classic Ethernet VLANs:** To interact with the classic Ethernet network, you set VLANs to either classic Ethernet or Cisco FabricPath mode. The classic Ethernet VLANs carry traffic from the classic Ethernet hosts to the Cisco FabricPath interfaces, and the Cisco FabricPath VLANs carry traffic throughout the Cisco FabricPath topology. All VLANs that are meant to be forwarded over the Cisco FabricPath network must be created as Cisco FabricPath VLANs. By default, all VLANs are in classic Ethernet mode.
- **Multidestination trees:** When a Cisco FabricPath edge switch receives a multidestination frame on an edge port, it selects one of the available multidestination trees to forward the frame.
- **Cisco FabricPath routing:** The IS-IS protocol establishes switch ID tables that enable the routing of FabricPath frames through the Cisco FabricPath network.
- **Conversational MAC learning:** Cisco FabricPath introduces new MAC address learning rules that optimize the learning process within the fabric and help conserve MAC address table space on the edge switches. This technique, which is known as conversational learning, occurs automatically in VLANs configured for Cisco FabricPath mode.

- **vPC+:** Virtual port channel plus (vPC+) is an extension to virtual port channels (vPCs) that provides the solution by creating a unique virtual switch that appears as a separate device to the rest of the Cisco FabricPath network. A vPC+ provides active-active Layer 2 paths for dual-homed classic Ethernet devices or clouds, even though the Cisco FabricPath network allows only 1-to-1 mapping between the MAC address and the switch ID.
- **Multicast:** Cisco FabricPath uses a hash-based system to assign each of the multicast flows to one of the two designated trees to ensure that the multicast traffic is load-balanced.

Cisco FabricPath Interaction with STP

- Layer 2 fabric appears as a single bridge to all connected CE devices.
- Layer 2 fabric should be the root for all connected STP domains.
  - Classic Ethernet ports will go into blocking state when “better BPDU” is received (rootguard)
- No BPDUs are forwarded across the fabric.
  - Terminated on classic Ethernet ports



In Cisco FabricPath topologies, there are two types of functions: edge (or leaf) devices and spine devices.

The STP domains do not cross into the FabricPath network.

You must configure the Cisco FabricPath edge switches to have the lowest STP priority (the recommendation is 8192) of all the devices in the STP domain to which they are attached. This action ensures that they become root for any attached STP domains. You should also configure all the Cisco FabricPath edge switches with the same priority. The system assigns the bridge ID for the Layer 2 gateway devices from a pool of reserved MAC addresses.

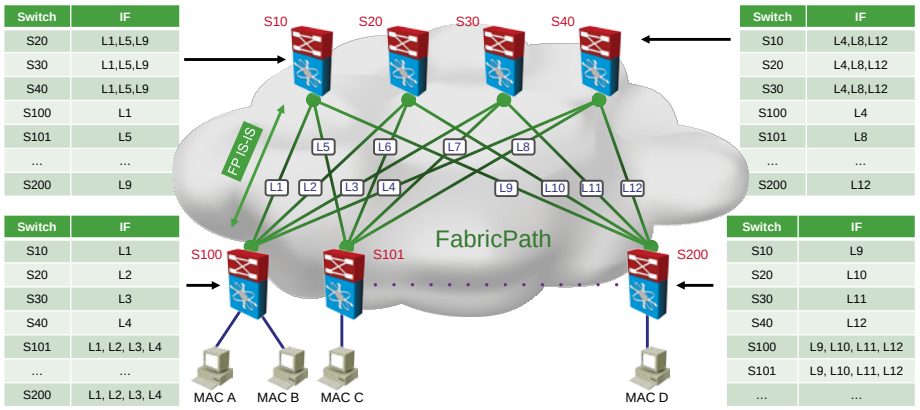
Other than configuring the STP priority on the Cisco FabricPath Layer 2 gateway switches, you do not need to configure anything for the STP to work seamlessly with the Cisco FabricPath network. Only connected classic Ethernet devices form a single STP domain. Those classic Ethernet devices that are not interconnected form separate STP domains, as shown in the figure.

All classic Ethernet interfaces should be designated ports, which occur automatically, or they are pruned from the active STP topology.

The Cisco FabricPath edge switches propagate the topology change notifications (TCNs) on all its classic Ethernet interfaces. The devices in the separate STP domains need to know the TCN information only for the domain to which they belong. You can configure a unique STP domain ID for each separate STP domain that connects to the same Cisco FabricPath network. The Layer 2 IS-IS messages carry the TCNs across the Cisco FabricPath network. Only those Cisco FabricPath Layer 2 gateway switches in the same STP domain as the TCN message need to act and propagate the message to connected classic Ethernet devices.

Cisco FabricPath Routing Table

- FabricPath IS-IS manages switch ID (routing) table
- Equal-cost path selection based on ECMP hash function
  - Maximum 16 (default) next-hop interfaces for each destination switch ID
  - Number controlled by **maximum-paths** command in FabricPath IS-IS process

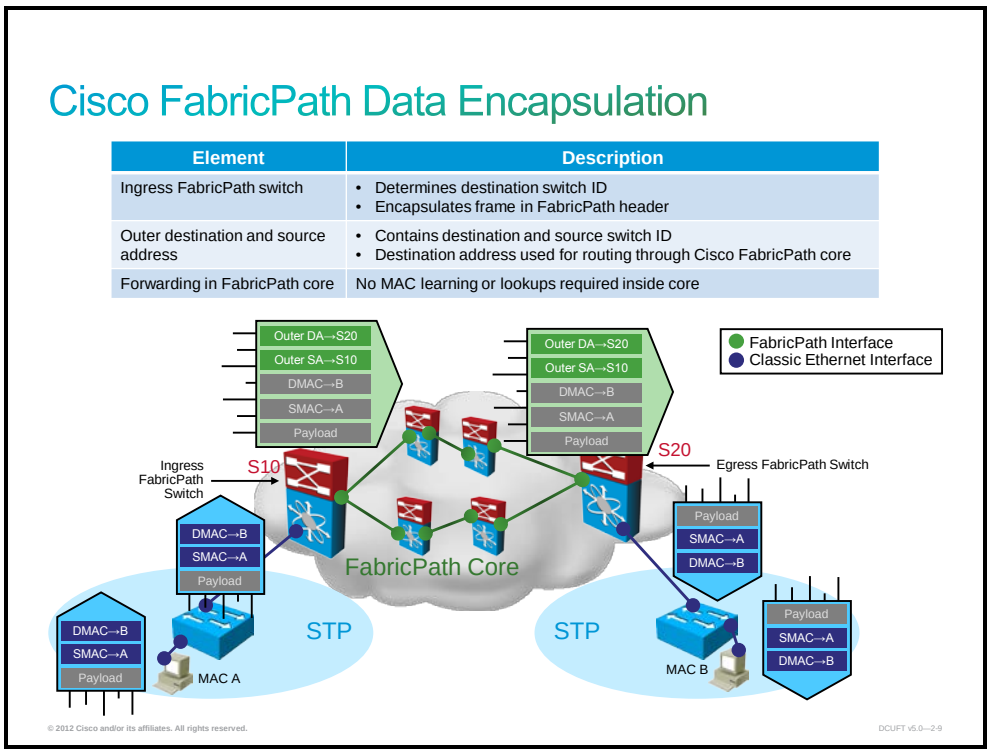


The IS-IS protocol establishes switch ID tables that enable the routing of Cisco FabricPath frames through the Cisco FabricPath network. The tables describe all available shortest paths to a given switch ID. Frames that traverse the Cisco FabricPath network carry the destination switch ID in the outer destination address. The transit switches (spines) look up the destination switch ID in the switch ID table and forward the frame along the selected multidestination tree (identified with a forwarding tag [FTag]) toward the destination edge switch.

Cisco FabricPath, using Layer 2 IS-IS, can utilize up to 16 active Layer 2 paths for forwarding known unicast packets. Forwarding of broadcast and multicast packets is constrained to a specific multidestination tree.

Cisco FabricPath Data Plane

This topic describes the Cisco FabricPath data plane.

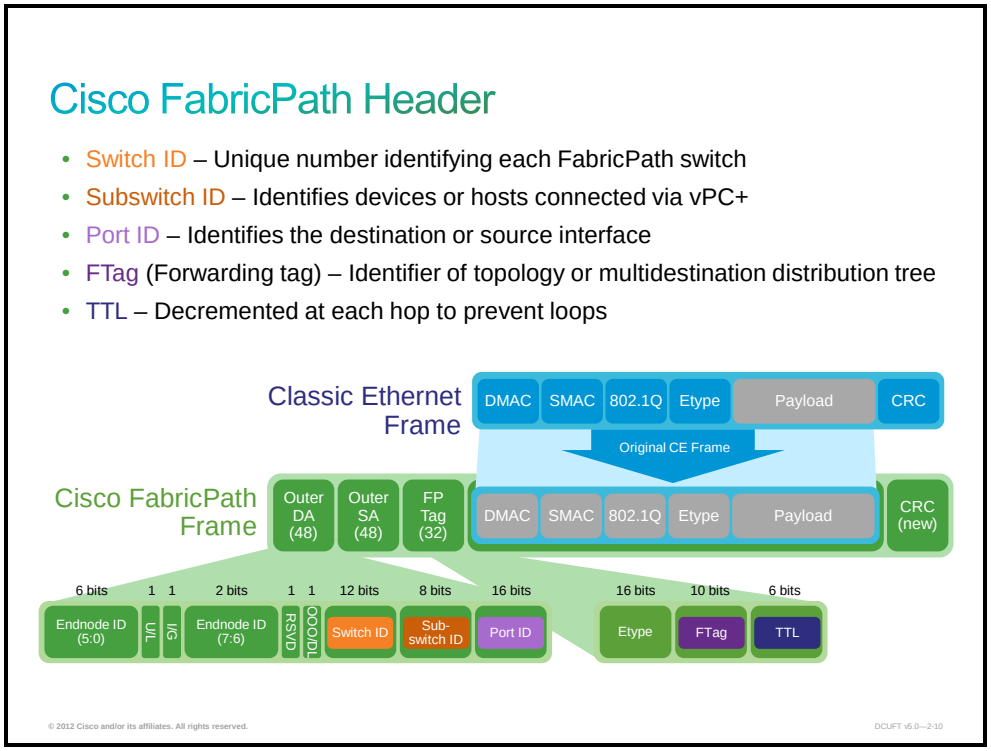


When a frame enters the Cisco FabricPath network on a Cisco FabricPath VLAN, the system encapsulates the Layer 2 frame with a new Cisco FabricPath header. The outer destination address (ODA) and outer source address (OSA) in the Cisco FabricPath header contain the switch IDs of the egress and ingress switch, respectively.

The system applies the encapsulation on the ingress edge port of the Cisco FabricPath network and de-encapsulates the frame on the egress edge port of the FabricPath network; all the ports within the FabricPath network are FabricPath ports that use only the hierarchical MAC address. This feature greatly reduces the size of the MAC tables in the core of the Cisco FabricPath network.

The system automatically assigns each device in the Cisco FabricPath network with a unique switch ID. Optionally, you can configure the switch ID for the FabricPath device.

The OSA is the Cisco FabricPath switch ID of the device where the frame ingresses the FabricPath network, and the ODA is the FabricPath switch ID of the device where the frame egresses the FabricPath network. When the frame egresses the Cisco FabricPath network, the FabricPath device strips the FabricPath header, and the original classic Ethernet frame continues on the classic Ethernet network. The Cisco FabricPath network uses only the OSA and ODA, with the Layer 2 IS-IS protocol transmitting the topology information. Both the Cisco FabricPath ODA and OSA are in a standard MAC format (xxxx.xxxx.xxxx).



The figure illustrates the encapsulation process and the outer header that is used for transporting the frame through the Cisco FabricPath cloud. The Cisco FabricPath encapsulation uses a MAC address-in-MAC address encapsulation format. The original Ethernet frame, along with an IEEE 802.1Q tag, is prepended by a 48-bit OSA, a 48-bit ODA, and a 32-bit Cisco FabricPath tag.

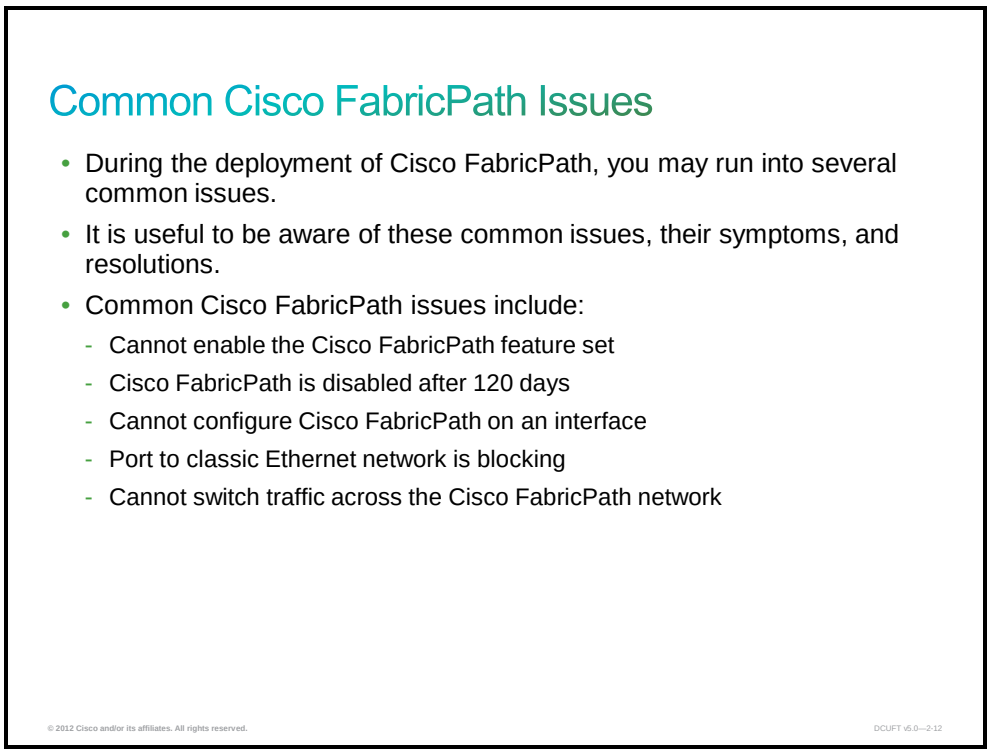
In addition to the switch ID, the Cisco FabricPath header addresses contain these fields:

- The subswitch ID field identifies the source or destination vPC+ port channel interface associated with a particular vPC+ switch pair. Cisco FabricPath switches that are running vPC+ use this field to identify the specific vPC+ port channel on which traffic is to be forwarded. The subswitch ID value is locally significant to each vPC+ switch pair. In the absence of vPC+, this field is set to 0.
- The port ID, also known as the local identifier (local ID), identifies the specific physical or logical interface on which the frame was sourced or to which it is destined. The value is locally significant to each switch. This field in the ODA allows the egress Cisco FabricPath switch to forward the frame to the appropriate edge interface without requiring a MAC address table lookup. For frames sourced from or destined to a vPC+ port channel, this field is set to a common value shared by both vPC+ peer switches, and the subswitch ID is used to select the outgoing port instead.
- The EtherType value for Cisco FabricPath encapsulated frames is 0x8903.
- The function of the FTag depends on whether a particular frame is unicast or multidestination. For unicast frames, the FTag identifies the Cisco FabricPath topology that the frame is traversing. For multidestination frames, the FTag identifies the multidestination forwarding tree that the frame should traverse.

The time-to-live (TTL) field serves the same purpose as in traditional IP forwarding: each switch hop decrements the TTL by 1, and frames with an expired TTL are discarded. It prevents Layer 2 bridged frames from looping endlessly if a transitory loop occurs. Ingress Cisco FabricPath edge switches set the TTL to 32 for all frames.

Troubleshooting Cisco FabricPath

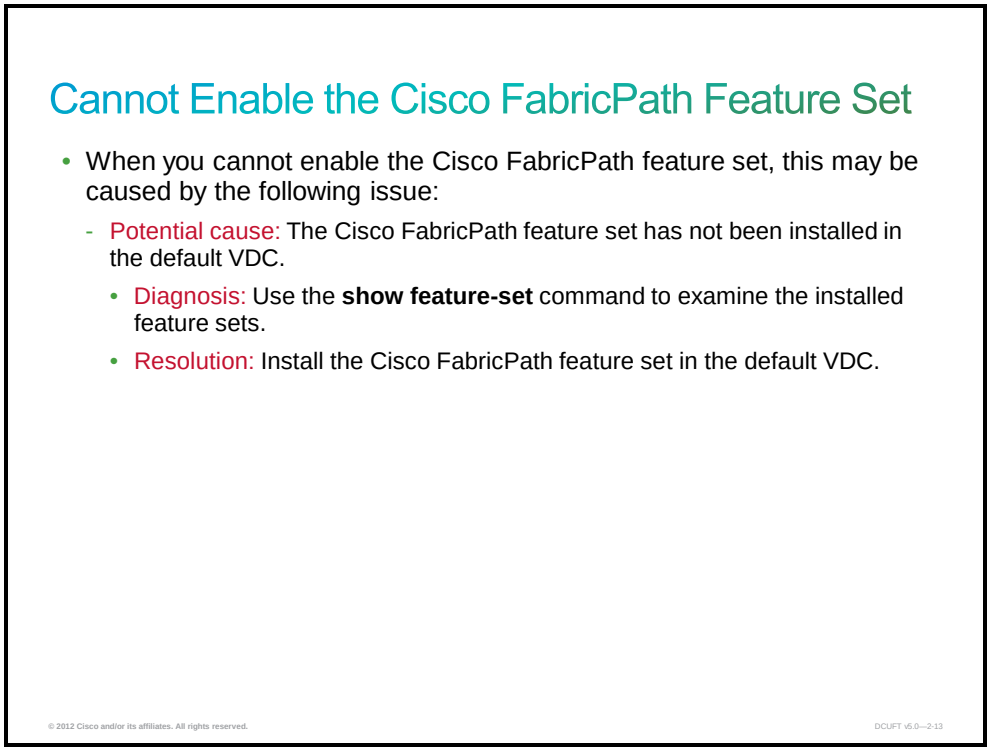
This topic explains how to troubleshoot Cisco FabricPath on a Cisco Nexus switch.



During the deployment of Cisco FabricPath in the data center network, you may encounter various problems. Assuming it is configured correctly (and the configuration is minimal), very little should go wrong with Cisco FabricPath. Usually, issues that occur are due to optional configuration commands.

It is not possible to create an exhaustive list that contains all possible Cisco FabricPath problems with their root causes and resolutions. However, certain problems occur more often than others, and it can be helpful to have a short list of common Cisco FabricPath issues, their possible causes, and their resolutions.

This topic covers some of the most common Cisco FabricPath problems.



In order to enable the Cisco FabricPath feature set in any virtual device context (VDC), including the default VDC, it is necessary to first install the Cisco FabricPath feature set in the default VDC. If the feature set has not been installed, the **fabricpath** keyword will not be available in the **feature-set** command.

It is also possible that the feature set has been disabled explicitly for the VDC that you are in. By default, a feature set is available to all VDCs once the feature set has been installed. However, a network administrator can specifically deny the use of a feature set in a VDC from the default VDC.

```
N7K1-POD3# show feature-set
Feature Set Name      ID        State
--------------------  --------  --------
fcoe                   1          disabled
fabricpath             2          enabled
fex                    3          disabled
mpls                   4          disabled
N7K1-POD3#
```

Cisco FabricPath Is Disabled After 120 Days

- When the Cisco FabricPath feature is disabled after 120 days, this may be caused by the following issue:
  - Potential cause: There is no valid Enhanced Layer 2 license installed and the grace period has expired.
    - Diagnosis: Use the **show license usage** command to examine the installed licenses.
    - Resolution: Install the Enhanced Layer 2 license in the default VDC.
- Note: If the system is using the grace period to run the Cisco FabricPath feature, you will see daily system log messages of the following type to alert you to this fact:

  "%LICMGR-2-LOG_LICAPP_NO_LIC: Application fabricpath running without ENHANCED_LAYER2_PKG license, shutdown in 119 days."

In order to use the Cisco FabricPath feature, you need to have a valid Enhanced Layer 2 license installed. The feature can be enabled without a license for a 120-day grace period. When the grace period expires, the Cisco FabricPath feature will be disabled.

If the system is using the grace period instead of a valid installed license to run the Cisco FabricPath feature, you will be notified of this fact by system log messages that are similar to the following:

```
2012 Feb 15 19:03:23 N7K-1 %LICMGR-2-LOG_LIC_NO_LIC: No
license(s) present for feature ENHANCED_LAYER2_PKG.
Application(s) shut down in 119 days.
2012 Feb 15 19:03:23 N7K-1 %LICMGR-2-LOG_LICAPP_NO_LIC:
Application fabricpath running without ENHANCED_LAYER2_PKG
license, shutdown in 119 days.
```

Cannot Configure Cisco FabricPath on an Interface

- When you cannot configure Cisco FabricPath on an interface, this may be caused by the following issue:
  - Potential cause: The interface does not belong to an F-series I/O module, such as the N7K-F132XP-15 module.
    - Symptom: The command is rejected with an error message that states *"Configuration does not match the port capability."*
    - Diagnosis: Use the **show interface capabilities** command to determine the type of module that the interface is located on and to verify whether the port is Cisco FabricPath-capable.
    - Resolution: Use a port on an F-series I/O module to configure Cisco FabricPath on.

Only ports on a Cisco Nexus 7000 F-series I/O module can be used as Cisco FabricPath interfaces. Cisco Nexus 5500 switches also support Cisco FabricPath, while 50x0 switches do not. When you attempt to configure Cisco FabricPath on a non-F-series module using the **switchport mode fabricpath** command, the command is rejected with an error message that states "Configuration does not match the port capability."

This restriction is not limited to Cisco FabricPath interfaces, but also applies to ingress interfaces that carry Cisco FabricPath VLANs. If a Cisco FabricPath VLAN is associated with a port that does not reside on an F-series module, it will not be possible to switch traffic from that port across the Cisco FabricPath network.

If you configure a port on a non-F-series module, such as the N7K-M132XP-12 module as an access port in a Cisco FabricPath VLAN, you will see the following interface state in the **show interface** command:

```
N7K-1# show interface ethernet 1/17
Ethernet1/17 is down (Inactive - M1 port not allowed in FabricPath-
mode VLAN)
```

Port to Classic Ethernet Network is Blocking

- When STP is blocking the port to the classic Ethernet network for a Cisco FabricPath VLAN, this may be caused by the following issue:
  - Potential cause: The port is receiving superior BPDUs on a classic Ethernet port for a Cisco FabricPath VLAN.
    - Diagnosis: Use the **show spanning-tree** or the **show spanning-tree summary** command to determine if you are receiving superior BPDUs on the classic Ethernet port.
    - Resolution: Configure the spanning-tree priority of the switch to be lower than any other switch in the classic Ethernet network to enforce that the Cisco FabricPath switch is the root for the classic Ethernet network.

```
N5548-2# show spanning-tree summary
Switch is in rapid-pvst mode
Root bridge for: VLAN0001, VLAN0010-VLAN0013
<…>
Name                 Blocking Listening Learning Forwarding STP Active
---------------------- -------- --------- -------- ---------- ----------
VLAN0001                    0         0        0          2          2
VLAN0010                    0         0        0          2          2
VLAN0011                    0         0        0          2          2
VLAN0012                    0         0        0          2          2
VLAN0013                    0         0        0          2          2
---------------------- -------- --------- -------- ---------- ----------
5 vlans                     0         0        0         10         10
```


You must configure the Cisco FabricPath Layer 2 gateway device to have the lowest spanning-tree priority of all the devices in the spanning-tree domain to which it is attached to ensure that the Cisco FabricPath gateway device becomes the root of the spanning tree for the classic Ethernet network. You must also configure all the Cisco FabricPath Layer 2 gateway devices that are connected to one Cisco FabricPath network to have the same priority. To have a loop-free topology for the classic Ethernet and Cisco FabricPath hybrid network, the Cisco FabricPath network automatically presents itself as a single bridge to all connected classic Ethernet devices.

Note It is recommended that you set the spanning-tree priority on all Cisco FabricPath Layer 2 gateway switches to 8192 to ensure that they are the lowest priority.

When a Cisco FabricPath Layer 2 gateway device receives a superior bridge protocol data unit (BPDU) on a classic Ethernet port, it will block that port until it stops receiving superior BPDUs.

Cannot Switch Across the Cisco FabricPath Network

- When you cannot switch Layer 2 traffic across the Cisco FabricPath network, this may be caused by the following issues:
  - Potential cause: The VLANs that the connectivity is failing for are not configured as Cisco FabricPath VLANs.
    - Diagnosis: Use the **show vlan** command to verify that the VLAN is configured as a Cisco FabricPath VLAN and use the **show fabricpath topology vlan active** command to verify that the Cisco FabricPath VLAN is active in the Cisco FabricPath topology.
    - Resolution: Configure the VLANs that need to be switched across the Cisco FabricPath network as Cisco FabricPath VLANs.

In order to allow a VLAN to be forwarded across the Cisco FabricPath network, it is necessary to configure it as a Cisco FabricPath VLAN. By default, all VLANs are treated as pure classical Ethernet VLANs and not extended across the fabric.

```
N7K1-POD1# show fabricpath topology vlan active
Topo-Description                 Topo-ID   Active VLAN List
-------------------------------- --------- -------------------
--
0                                0         10-12
N7K1-POD1#
```

When you cannot switch Layer 2 traffic across the Cisco FabricPath network, this may also be caused by the following issues:

- **Potential cause:** The ingress interface does not belong to an F-series I/O module, such as the N7K-F132XP-15 module.
- **Diagnosis:** Use the **show module** command to determine the type of module that the ingress interface is located on.
- **Resolution:** Ensure that you use a port on an F-series module to connect to the classic Ethernet network for Cisco FabricPath VLANs.

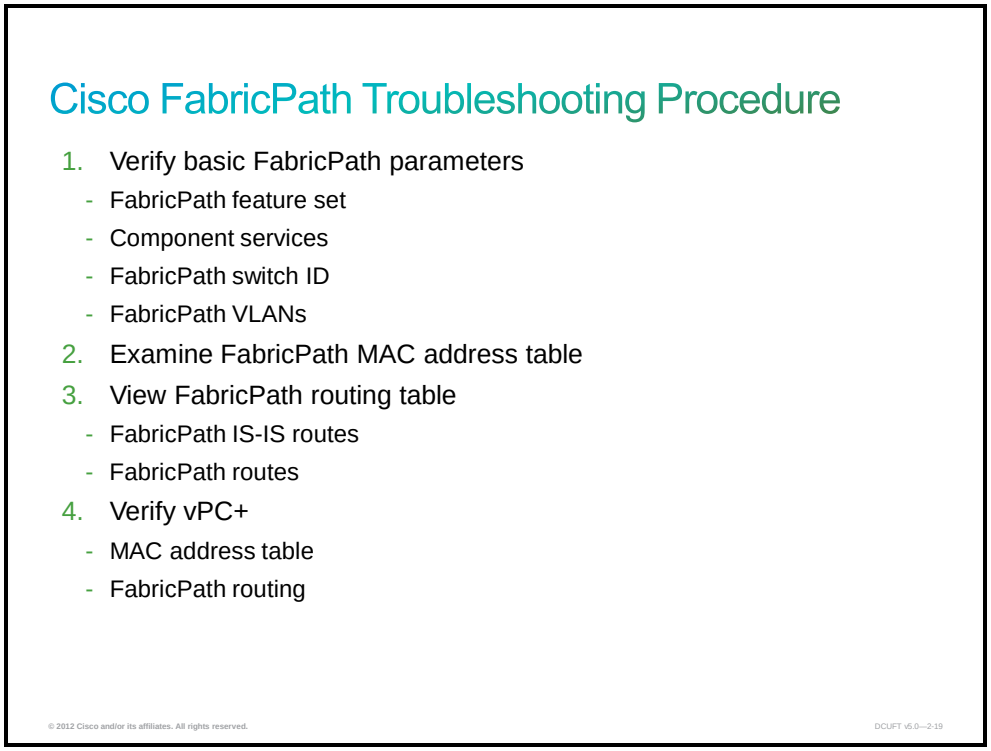
Diagnostic Tools

- It is important to understand the troubleshooting process and diagnostic tools that can be used to troubleshoot Cisco FabricPath problems with an unknown cause.
  - Typical Cisco FabricPath troubleshooting processes resemble Layer 3 troubleshooting processes due to the nature of the Cisco FabricPath control plane.
  - Typical Cisco FabricPath troubleshooting processes include Layer 2 troubleshooting processes due to the nature of the Cisco FabricPath data plane.
- Tools to support the Cisco FabricPath troubleshooting process include:
  - Cisco NX-OS CLI commands
  - Cisco DCNM

It is not possible to create an exhaustive list of all possible Cisco FabricPath problems, their root causes, and resolutions. It is important for a data center network engineer to understand the general operation and implementation of Cisco FabricPath and to be able to apply this knowledge in a structured troubleshooting process. It is equally important to know the available tools to gather information about the state of the Cisco FabricPath network. Being familiar with the Cisco FabricPath troubleshooting toolkit will make your troubleshooting processes more efficient and effective.

Troubleshooting Cisco FabricPath shares aspects with troubleshooting Layer 3 routing. The Cisco FabricPath control protocol is based on IS-IS and the FabricPath routing table resembles a Layer 3 routing table. Troubleshooting Cisco FabricPath also includes some Layer 2 troubleshooting processes. Cisco FabricPath switches Layer 2 Frames between classic Ethernet networks across the FabricPath cloud. At the edge of the Cisco FabricPath network, conversational MAC learning is used to distribute MAC address reachability information. Verification of MAC address tables and spanning-tree interaction at the edge of the Cisco FabricPath is an important step in many FabricPath troubleshooting processes.

Two important tools that can be used to support a FabricPath troubleshooting process are the Cisco NX-OS CLI on the Nexus switches and the Cisco Data Center Network Manager (DCNM) network management software.



Perform these steps to verify Cisco FabricPath operation:

Step 1 Verify basic FabricPath parameters, such as the enabled FabricPath feature, its component services, FabricPath switch ID, and FabricPath VLANs.

Step 2 Examine the FabricPath MAC address table.

Step 3 View FabricPath routing table. You can use various command options to view different information about the switch ID table.

Step 4 Verify vPC+. You can examine the MAC address table for entries that are related to vPC+ and search the switch ID table for the entries with the virtual switch ID.

The explicit CLI commands used for all these verifications and examples for demonstration purposes are shown in the following pages. When using a Cisco Nexus 7000 switch, ensure you are in the correct VDC.

Verify Basic Cisco FabricPath Settings

```
switch# show feature-set
Feature Set Name      ID        State
--------------------  --------  --------
fabricpath            2         enabled
fex                   3         disabled

switch# show feature-set services fabricpath
u2rib
drap
isis_fabricpath
3 services in feature set fabricpath

switch# show fabricpath switch-id

                      FABRICPATH SWITCH-ID TABLE
Legend: '*' - this system
========================================================================
SWITCH-ID      SYSTEM-ID      FLAGS       STATE    STATIC EMULATED
----------+---------------+------------+-----------+-------------
10          0018.bad8.12fd    Primary     Confirmed   Yes      No
*25         0018.bad8.12fe    Primary     Confirmed   Yes      No
30          0018.bad8.12ff    Primary     Confirmed   Yes      No

switch# show fabricpath topology vlan active
TPG-name TPG-  ID              Active VLAN List
-------- --------- -------------------------------------------
0            0                  10-30
```


First, you should make sure that the Cisco FabricPath feature set has been enabled. It comprises three services: Unicast Layer 2 Routing Information Base (U2RIB), Dynamic Resource Allocation Protocol (DRAP), and IS-IS FabricPath. Examine the switch IDs using the **show fabricpath switch-id** command. The switch IDs can be set manually or automatically can be provisioned by the system. View the active Cisco FabricPath VLANs using the **show fabricpath topology vlan active** command.

Verify Cisco FabricPath MAC Address Table

- Local MAC addresses denoted by attachment port
- Remote MAC addresses denoted by switch ID, subswitch ID, and local ID
- Local ID
  - Identifies the exact source and destination port on the switch
  - No need for address lookup on egress switch

```
switch# show mac address-table dynamic vlan 10
Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link
   VLAN     MAC Address      Type      age     Secure NTFY Ports/SWID.SSID.LID
---------+-----------------+--------+---------+------+----+------------------
* 10       0000.0000.0001    dynamic   0          F    F  Eth1/15
* 10       0000.0000.0002    dynamic   0          F    F  Eth1/15
* 10       0000.0000.0008    dynamic   0          F    F  Eth1/15
* 10       0000.0000.0009    dynamic   0          F    F  Eth1/15
* 10       0000.0000.000a    dynamic   0          F    F  Eth1/15
  10       0000.0000.000b    dynamic   0          F    F  200.0.30
  10       0000.0000.000c    dynamic   0          F    F  200.0.30
  10       0000.0000.000d    dynamic   0          F    F  200.0.30
  10       0000.0000.000e    dynamic   0          F    F  200.0.30
```

MAC address table in FabricPath VLAN

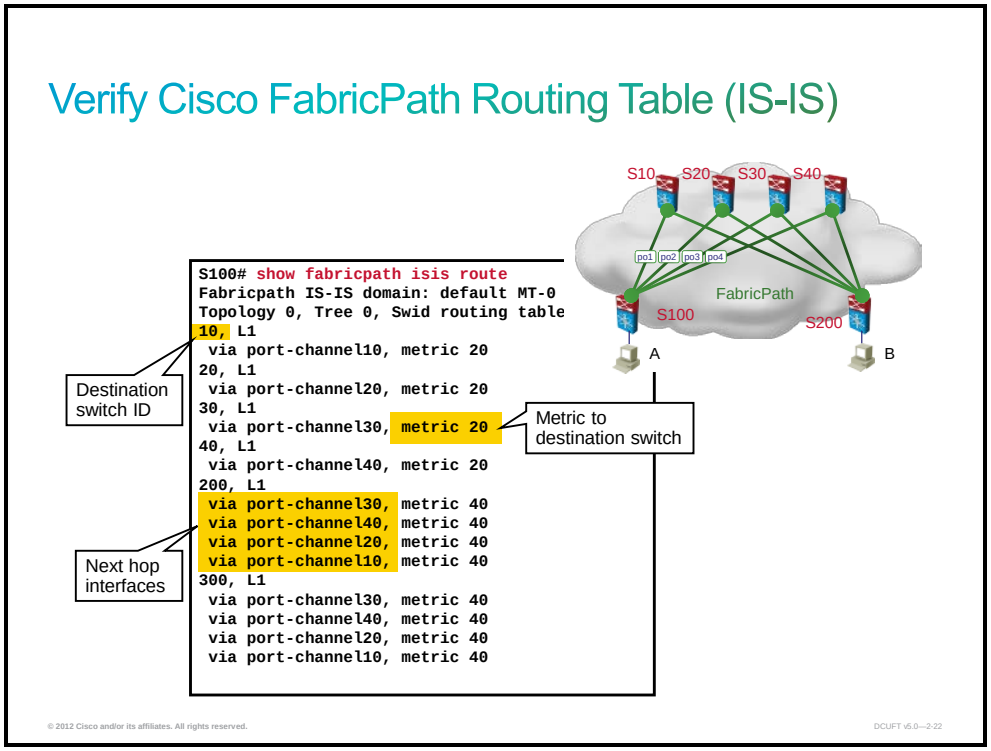
Local address

Remote address

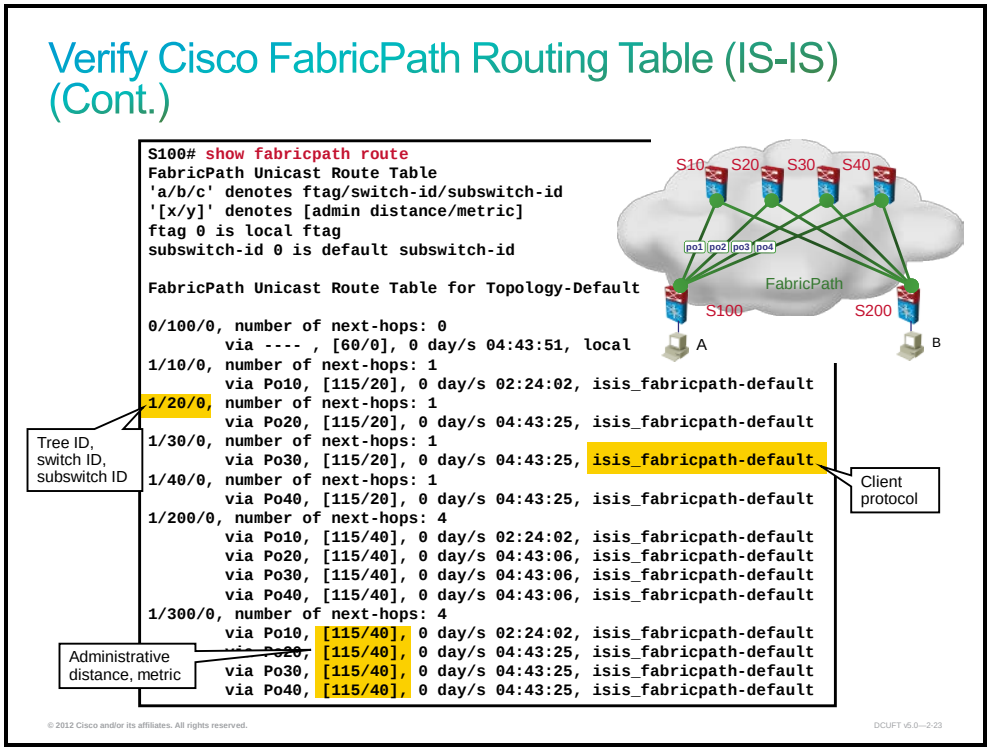
You can display the MAC address table using the **show mac address-table** command and view the MAC addresses that are learned in a VLAN. VLAN 10 has been configured as Cisco FabricPath VLAN and therefore contains two types of MAC address entries: local and remote.

Local addresses are identified by the classic Ethernet interface to which they are attached.

Remote addresses are denoted by the parameters switch ID, subswitch ID, and local ID. The remote MAC addresses that are shown in the figure are attached behind the remote switch with ID 200. The subswitch ID is set to 0 because there is no vPC+ bundle at the remote end. The remote interface has the local interface ID of 30. Local ID identifies the exact source and destination port on the switch. It simplifies forwarding by making an address lookup on the egress switch redundant.



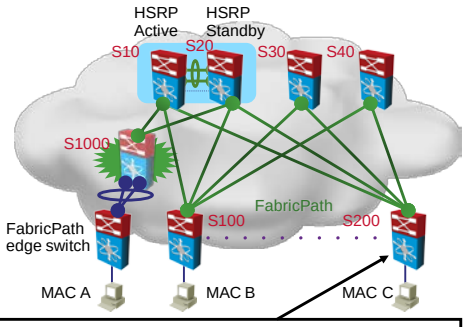
You can examine the switch ID table using multiple command options. The **show fabricpath isis route** command displays the IS-IS topology by showing the available best paths to all switch IDs in the Cisco FabricPath domain.



The **show fabricpath route** command offers more comprehensive information by adding details about the administrative distance, the age, and the client protocol.

Verify VPC+: MAC Address Table

- Notation: SWID.sSID.LID
- In vPC+:
  - SWID: Virtual switch ID (1000)
  - sSID: Subswitch identifies exact port channel
  - LID: not used



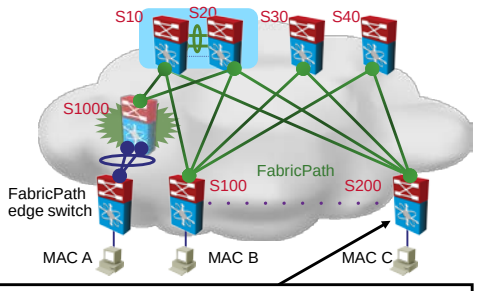

```
S200# show mac address-table dynamic
Legend:
        * - primary entry, G - Gateway MAC, (R) - Routed MAC, O - Overlay MAC
        age - seconds since last seen,+ - primary entry using vPC Peer-Link
   VLAN     MAC Address      Type      age     Secure NTFY Ports/SWID.SSID.LID
---------+-----------------+--------+---------+------+----+------------------
 10       0000.0000.000c    dynamic   1500       F    F    Eth1/30
 10       0000.0c07.ac0a    dynamic   0          F    F    1000.11.4513
```

HSRP vMAC

You can verify the vPC+ operations by examining the MAC address tables on other switches within the Cisco FabricPath network. You will see one or more MAC addresses that are related to the virtual switch ID. The output in the figure displays the HSRP virtual MAC address. The switch ID is the virtual switch ID (1000), the subswitch ID (11) identifies the vPC bundle, and the local ID is not used.

Verify vPC+: Cisco FabricPath Routing

1. Search the FabricPath routing table for the virtual switch ID (1000).
2. In this example, there are two parallel paths to the virtual switch in the default tree.



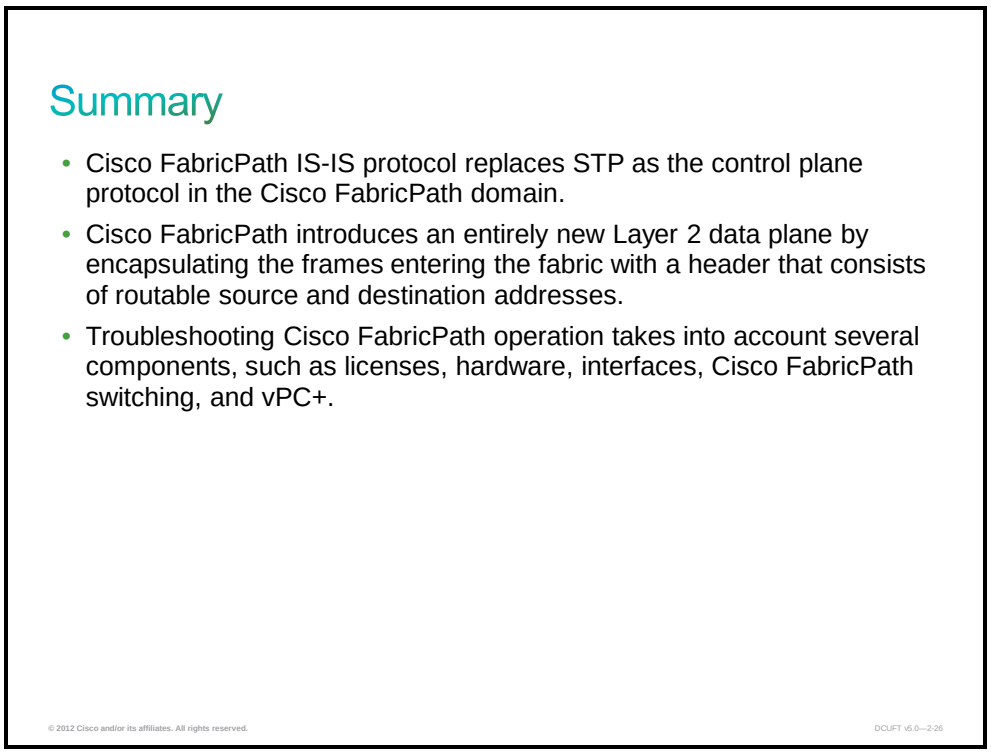

```
S200# show fabricpath route topology 0 switchid 1000
FabricPath Unicast Route Table
'a/b/c' denotes ftag/switch-id/subswitch-id
'[x/y]' denotes [admin distance/metric]
ftag 0 is local ftag
subswitch-id 0 is default subswitch-id


FabricPath Unicast Route Table for Topology-Default

1/1000/0, number of next-hops: 2
        via Po1, [115/10], 0 day/s 01:09:56, isis_l2mp-default
        via Po2, [115/10], 0 day/s 01:09:56, isis_l2mp-default
```

1 2

Further, you can verify vPC+ operation by examining the switch ID table on other switches. You can narrow down the contents by selecting a specific topology (FTag) and the desired switch ID, as in the command in the figure. The output displays two paths to the virtual switch ID 1000, in the default topology (ID 1).

Summary

This topic summarizes the key points that were discussed in this lesson.



Lesson 4

Troubleshooting OTV

Overview

This lesson is designed to provide the student with some examples of common issues that are related to Overlay Transport Virtualization (OTV) and how to resolve them.

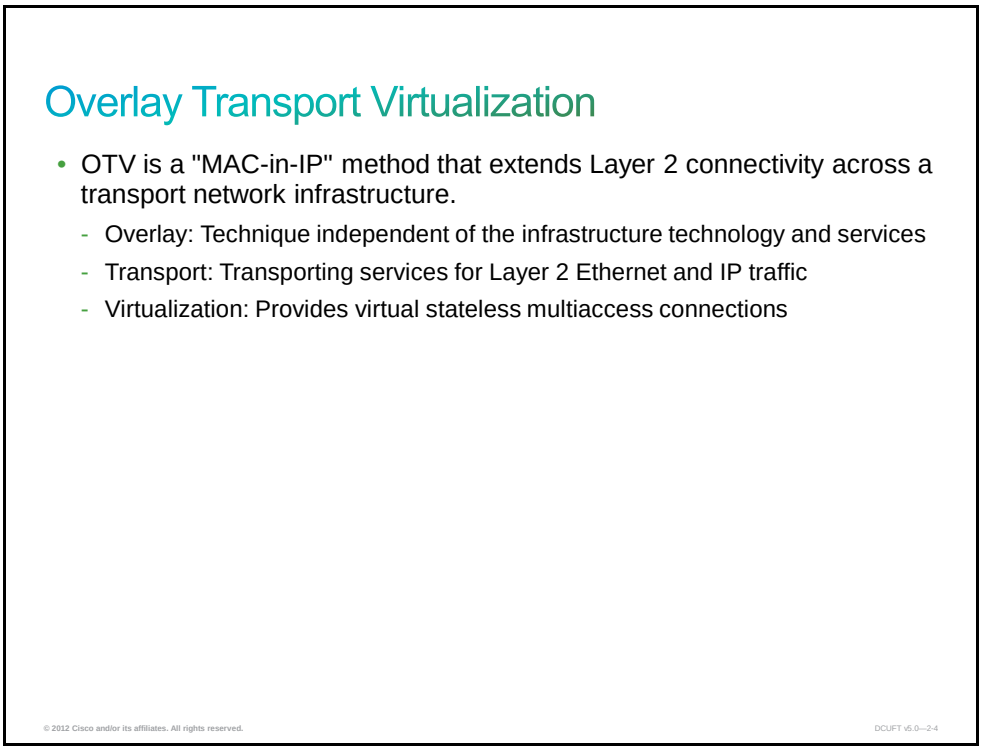
Objectives

Upon completing this lesson, you will be able to identify and resolve issues that are related to OTV. You will be able to meet these objectives:

- Describe the OTV on the Cisco Nexus switch
- Explain how to troubleshoot issues that are related to OTV on the Cisco Nexus switch
- Describe the HSRP isolation between data centers using OTV

OTV Review

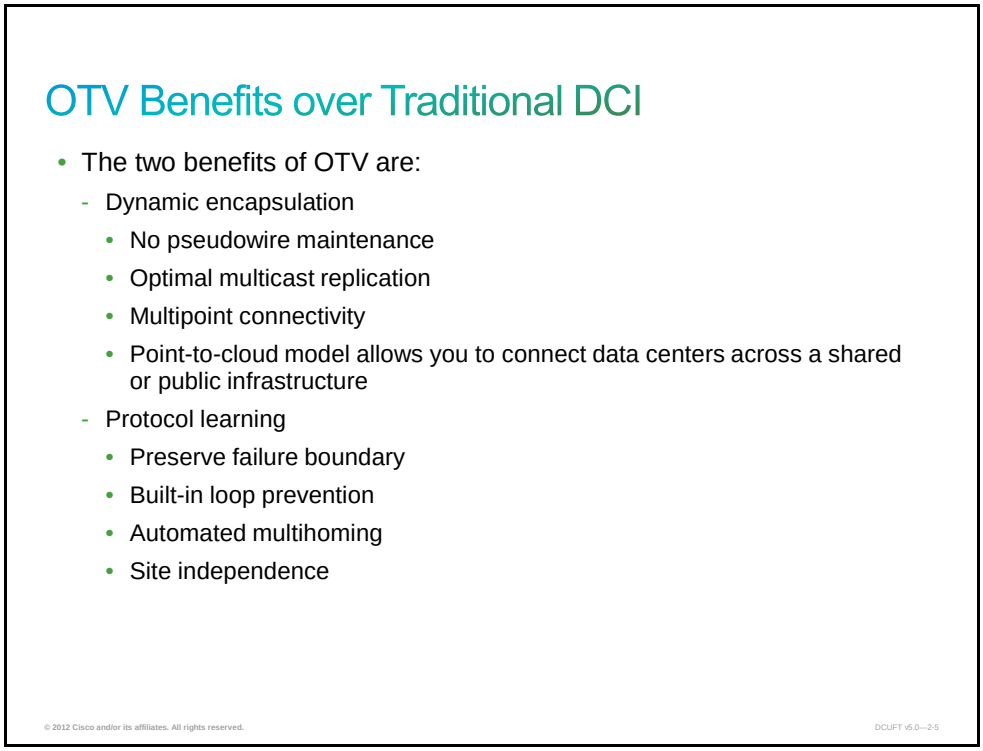
This topic describes the OTV on the Cisco Nexus switch.



OTV overcomes the challenges inherent to traditional Layer 2 data center interconnect technologies. OTV consists of two main components. There is a control plane component, the OTV Intermediate System-to-Intermediate System (IS-IS) control protocol, which is used to advertise MAC reachability between sites. In addition, there is a data plane component, which encapsulates and de-encapsulates the packets as they are sent or received on the overlay.

The name of the technology describes its key characteristics:

- **Overlay:** OTV provides an overlay VPN on top of an IP network. It is independent of the underlying infrastructure technologies and services. It does not impose any restrictions on the underlying infrastructure, as long as it is capable of transporting IP packets.
- **Transport:** OTV provides a Layer 2 transport across a Layer 3 network. It can leverage all the underlying capabilities of the underlying transport network, such as fast convergence, load balancing, and multicast replication.
- **Virtualization:** OTV provides a virtual multiaccess Layer 2 network that supports the efficient transport of unicast, multicast, and broadcast traffic. Sites can be added to an OTV overlay without a need to provision additional point-to-point connections to the other sites. There are no virtual circuits, pseudowires, or other point-to-point connections to maintain between the sites. Packets are routed independently from site to site without a need to establish stateful connections between the sites.

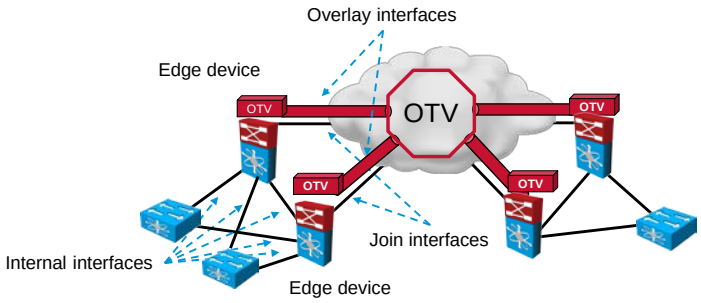


OTV is a "MAC-in-IP" encapsulation technique, which rests on two technology pillars:

- **Dynamic encapsulation:** OTV uses a point-to-cloud model, which encapsulates Layer 2 frames in IP packets that can be independently routed to the destination site or sites. There is no need to provision point-to-point Layer 2 circuits between the sites. If the underlying IP network is multicast-enabled, then multicast frames can be encapsulated in IP multicast packets. The frames can be replicated and routed to the appropriate sites by the multicast IP transport network.
- **Protocol learning:** OTV uses MAC address routing that is based on the IS-IS protocol to dynamically learn which MAC addresses are available in each of the sites. The use of a routing protocol eliminates the need to depend on Layer 2 flooding for MAC address learning. By eliminating unnecessary flooding, failure domains are bounded. Also, there are loop prevention mechanisms that are built into the routing protocol, which eliminates the need to extend Spanning Tree Protocol (STP) to break forwarding loops. Sites remain independent spanning-tree domains. Finally, the OTV neighbor discovery mechanism provides for automated multihoming of sites to the overlay network.

OTV Components

- Edge device: Encapsulation and de-encapsulation between Layer 2 and OTV and all OTV functions
- Internal interfaces: Connects to the VLANs that are to be extended
- Join interface: Joins an overlay network
- Overlay interface: Encapsulates the Layer 2 frames in IP packets



In order to understand the operation of OTV, it is important to establish some of the key terms.

OTV is an edge function. Layer 2 traffic is received from the switched network. For all VLANs that need to be extended to remote locations, the Ethernet frames are dynamically encapsulated into IP packets that are then sent across the transport infrastructure. A device that performs the OTV encapsulation and de-encapsulation functions between the Layer 2 network and the transport network is an "OTV edge device."

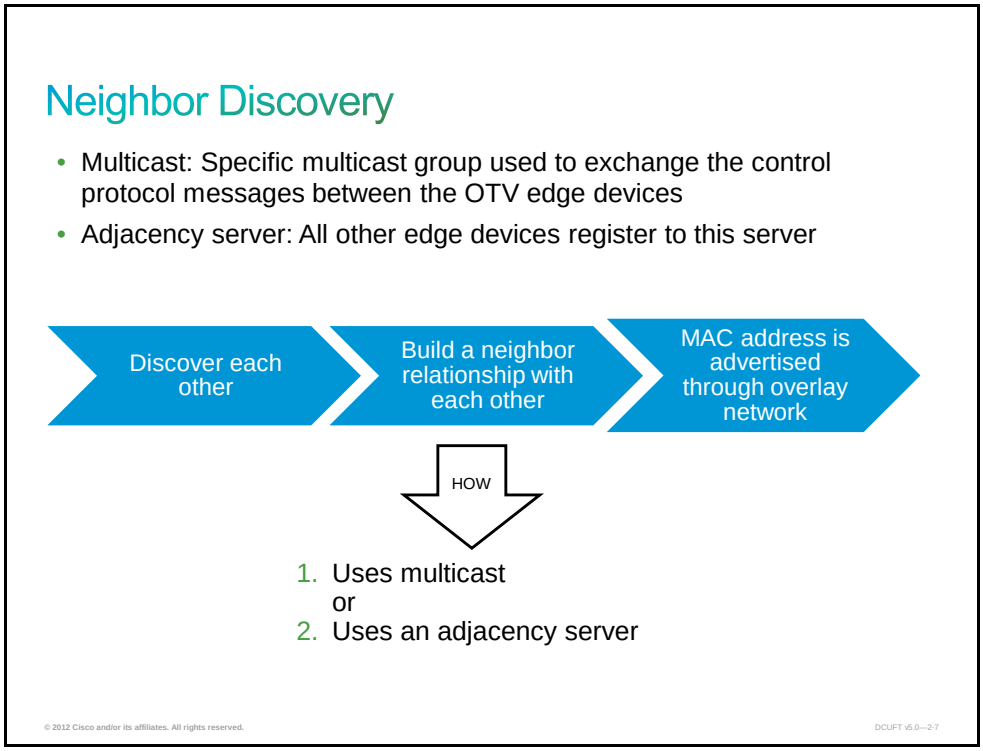
The OTV edge device can either be in the core or aggregation layer on the Cisco Nexus 7000 of the data center. A site can have multiple edge devices to provide additional resiliency. This scenario is commonly referred to as multihoming.

By definition, an OTV edge device has interfaces that connect to the transport network and interfaces that connect to the Layer 2 switched network. The Layer 2 interfaces that receive the traffic from the VLANs that are to be extended across OTV are named the "internal interfaces." These interfaces are regular Layer 2 interfaces, usually 802.1Q trunks. No OTV-specific configuration is required on the internal interfaces.

The "join interface" is used to source the OTV-encapsulated traffic and send it to the Layer 3 domain of the data center network. The join interface is a Layer 3 entity. The join interface can only be defined as a routed point-to-point physical or logical (port channel) interface.

An OTV overlay can only have a single join interface per edge device. Multiple overlays can share the same join interface.

The "overlay interface" is a logical multiaccess and multicast-capable interface that must be explicitly defined by the user. The entire OTV overlay configuration is applied on this logical interface. Every time the OTV edge device receives a Layer 2 frame that is destined for a remote data center site, the frame is logically forwarded to the overlay interface. This causes the edge device to perform the dynamic OTV encapsulation on the Layer 2 frame and send it to the join interface toward the routed domain.



In order to advertise MAC address reachability, the OTV edge devices need to discover all the other edge devices on the overlay and establish adjacencies with each other.

The neighbor discovery and adjacency creation process can be achieved in one of two ways, depending on the nature of the transport network interconnecting the various sites:

- If the transport is multicast-enabled, a specific multicast group can be used to exchange the control protocol messages between the OTV edge devices.
- If the transport is not multicast-enabled, an OTV edge device is configured as an adjacency server to which all other edge devices register. The adjacency server communicates the list of devices belonging to a given overlay to all other OTV edge devices.

In order to take advantage of all OTV benefits, the use of a multicast enabled transport is recommended when extending VLANs between more than two data centers. Not only does this simplify the neighbor discovery and adjacency building process, but it also allows for more efficient forwarding of multicast traffic between the sites.

Neighbor Discovery (Cont.)

- When the transport network is multicast-enabled, OTV uses a multicast group for neighbor discovery.
  - Edge devices join the multicast group using IGMP.
  - PIM configuration is not required on edge devices.
  - OTV hellos and updates are encapsulated in the multicast group.
  - Core multicast replication is used to send updates to all OTV neighbors.
- To be able to unicast replicate, each OTV node must know a list of neighbors to replicate to.
  - OTV can be deployed with unicast-only transport.
  - Choose which OTV edge will be the adjacency server and configure it.
  - OTV edge devices must first register to the adjacency server.
  - When an OTV edge device is configured with the adjacency server address, it is added to the unicast replication list.

Assuming that the transport network is multicast-enabled, all OTV edge devices can be configured to join a specific any-source multicast (ASM) group on which they simultaneously act as a receiver and source. The specific ASM group to be used for OTV control is not prescribed and any unused ASM group can be used. If the transport is not under your own control, but provided by a service provider, you will have to coordinate the use of this ASM group with the provider.

The OTV edge devices do not act as multicast routers, but as multicast endpoints. Protocol Independent Multicast (PIM) does not need to be enabled on the OTV join interface.

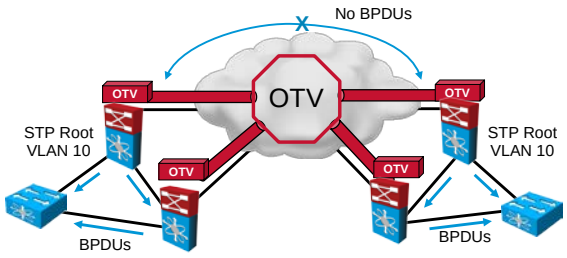
Edge devices use Internet Group Management Protocol (IGMP) to join the ASM control group on the join interface in order to be able to receive the OTV control protocol packets from the remote edge devices. Edge devices encapsulate their OTV protocol hellos and updates through the dynamic OTV encapsulation process, using the ASM control group as the destination IP address. The multicast routing protocols that are deployed in the transport network replicate the packets and deliver them to all the remote edge devices.

OTV can be deployed with unicast-only transport. In order to communicate with all the remote OTV devices, each OTV node needs to know a list of neighbors to replicate the control packets to. Rather than statically configuring in each OTV node the list of all neighbors, one (or more) OTV edge device to perform a specific role called the adjacency server. Every OTV device that has to join a specific OTV logical overlay needs first to register itself with the adjacency server by sending OTV hello messages to it. All other OTV neighbor addresses are discovered dynamically through the adjacency server.

The receipt of the hello messages from all the OTV edge devices helps the adjacency server build a list of all the OTV devices that should be part of the same overlay (named a unicast replication list). This list is periodically sent in unicast fashion to all the listed OTV devices, so that they can dynamically be aware of all the OTV neighbors in the network.

Spanning Tree and OTV

- OTV is site transparent for STP:
  - Each site maintains its own STP topology.
  - An OTV edge device only sends and receives BPDUs on internal interfaces.
  - This mechanism is built into OTV and requires no additional configuration.

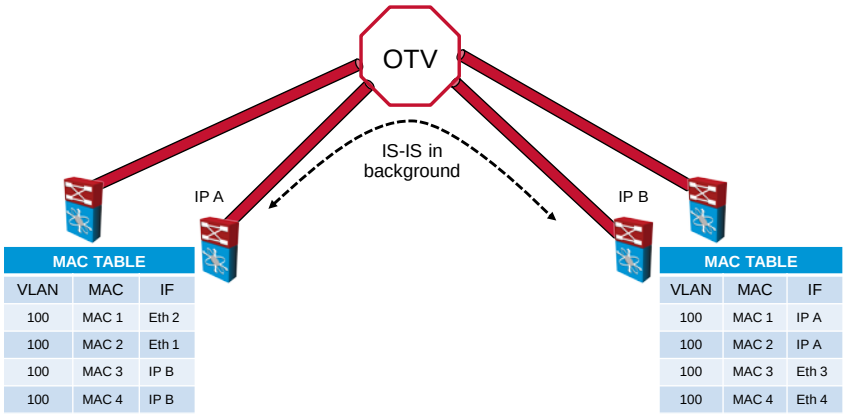


OTV by default does not transmit STP bridge protocol data units (BPDUs) across the overlay. This is a native OTV function that does not require the use of any explicit configuration, such as BPDU filtering. This allows every site to remain an independent spanning-tree domain: Spanning-tree root configuration, parameters, and the spanning-tree protocol flavor can be decided on a per-site basis.

The separation of spanning-tree domains fundamentally limits the fate sharing between data center sites. A spanning-tree problem in the control plane of a given site would not produce any effect on the remote data centers.

Building the MAC Address Table

- OTV periodically advertises MAC address reachability.
- MAC addresses are advertised in the background and no specific configuration is required.
- IS-IS runs in the background—there is no need to configure or understand it.



Before any encapsulation and de-encapsulation of frames across the overlay can be performed, it is necessary to exchange MAC address reachability information between the sites in order to create corresponding OTV MAC address table entries.

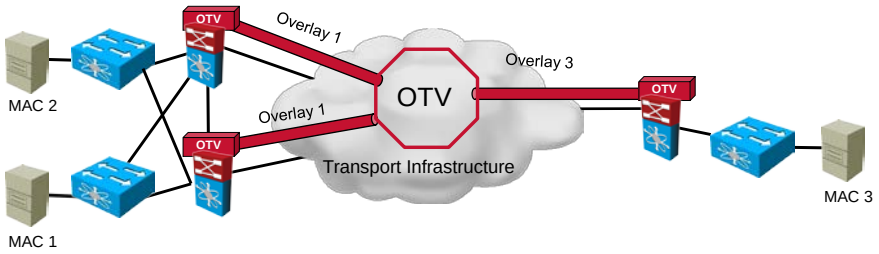
OTV does not depend on flooding to propagate MAC address reachability information. Instead, OTV uses a control plane protocol to distribute MAC address reachability information to remote OTV edge devices. This protocol runs as an overlay control plane between OTV edge devices. Therefore, there is no dependency with the routing protocol used in the Layer 3 domain of the data center or in the transport infrastructure.

The OTV control plane is transparently enabled in the background after creating the OTV overlay interface and does not require explicit configuration. Tuning parameters such as timers for the OTV protocol is possible, but is not expected to be a common requirement.

Note The routing protocol that is used to implement the OTV control plane is IS-IS. It was selected because it is a standards-based protocol, originally designed with the capability of carrying MAC address information in type, length, value (TLV) triplets. Despite the fact that IS-IS is used, the control plane protocol will be generically called "OTV protocol" to differentiate it from IS-IS being used as an interior gateway protocol (IGP) for IPv4 or IPv6. It is not necessary to have a working knowledge of IS-IS configuration to implement OTV. However, some background in IS-IS can be helpful when troubleshooting OTV.

Multihomed Sites and Load Balancing

- Multihome site: More then one edge device is present in overlay network.
- The site VLAN is a local VLAN and should not be extended across the overlay.
- OTV elects an AED to be the forwarder for a subset of the extended VLANs.
- The AED is responsible for advertising the MAC addresses and forwarding traffic to and from the overlay for its set of VLANs.
- Load balancing is achieved through designation of a subset of all VLANs for each edge device.



One key function that is built in the OTV protocol is multihoming, where two or more OTV edge devices provide LAN extension services to a given site. This redundant node deployment, which is combined with the fact that STP BPDUs are not sent across the OTV overlay, could potentially lead to the creation of a bridging loop between sites. To prevent this loop, OTV has a built-in mechanism to ensure that only one of the edge devices forwards traffic for a given VLAN. The edge device that has the active forwarding role for the VLAN is called an authoritative edge device (AED) for that VLAN.

The AED has two main tasks:

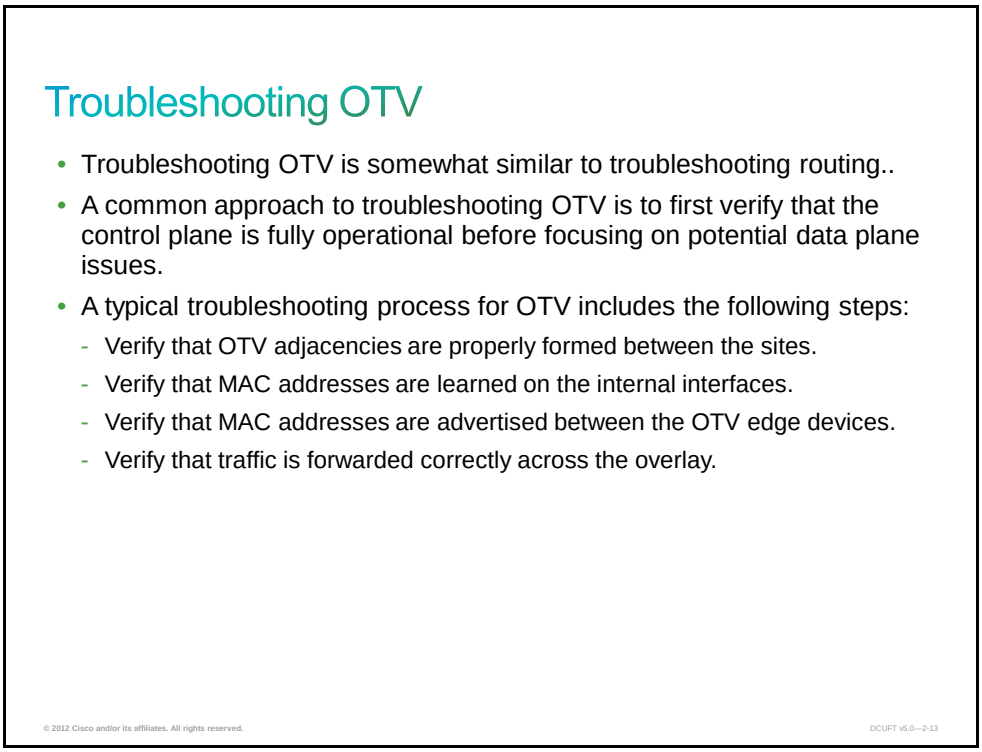
- Forwarding Layer 2 unicast, multicast, and broadcast traffic between the site and the overlay and vice versa
- Advertising MAC address reachability information to the remote edge devices

The AED role is negotiated, on a per-VLAN basis, between all the OTV edge devices belonging to the same site. To decide which device should be elected as an AED for a given site, the OTV edge devices establish an internal OTV control protocol peering.

The internal adjacency is established on a dedicated VLAN, named the site VLAN, and is used to negotiate the AED role. The site VLAN should be carried on multiple Layer 2 paths internal to a given OTV site to increase the resiliency of this internal adjacency.

Troubleshooting OTV

This topic explains how to troubleshoot issues that are related to OTV on the Cisco Nexus switch.



Troubleshooting OTV is somewhat similar to troubleshooting routing. The OTV data plane can only function correctly if the control plane is operating properly and the MAC address reachability information is exchanged between the sites. Therefore, a common approach to troubleshooting OTV is to first verify that the control plane is fully operational before focusing on potential data plane issues.

An initial troubleshooting plan could contain the following steps:

- Verify that OTV control protocol adjacencies are properly established between the sites. Without an adjacency, MAC addresses cannot be learned, and without MAC address entries, traffic cannot be forwarded, because OTV does not flood unknown unicast traffic on the overlay.
- Verify that the MAC addresses of the devices that are experiencing problems are learned on the internal interfaces of the OTV edge device. If the MAC addresses are not learned on the internal interfaces, they cannot be advertised on the overlay. If the MAC addresses are not learned on the internal interfaces, then this is likely to be a generic Layer 2 problem instead of an OTV problem.
- If the MAC addresses are learned on the internal interfaces, you should verify that they are advertised on the overlay. If MAC address reachability information is not properly communicated between sites, data traffic will not be possible.
- If the control plane looks correct, you should attempt to establish whether data is actually flowing on the overlay.

Verifying OTV Adjacencies

- Use the **show otv adjacency** command to confirm that a neighbor relationship has been established across the transport network:

```
N7K-1# show otv adjacency
Overlay Adjacency database

Overlay-Interface Overlay1  :
Hostname                        System-ID      Dest Addr       Up Time
State
N7K-2-pod6                      0026.9804.a944 10.7.7.6        12:36:42  UP
```

- OTV adjacencies are established using the configured multicast control group.
  - If adjacencies fail to establish, verify the following:
    - State and configuration of the overlay interface
    - IP multicast forwarding in the transport network

The **show otv adjacency** and **show otv data-group** commands can be used to verify whether the OTV control protocol adjacencies have been properly established. The OTV adjacencies are formed using the OTV control group for the overlay.

If adjacencies fail to establish, it is worth verifying that the configured control group is the same on all the overlay interfaces and that the overlay interface itself is operational.

If the overlay interface is operational and the control group is correctly configured, you may need to perform multicast troubleshooting in the core. You would do this to verify that the sites have joined the multicast tree for the control group, that they are sourcing packets for the control group, and that a multicast tree has been built from source to receiver.

Tip	To test multicast forwarding, you can configure the **ip igmp join-group** command for the control group on the join interface of an OTV edge device and then ping this multicast group from the remote OTV edge device.

OTV can also be deployed with unicast-only transport (adjacency-server mode). The OTV control plane over a unicast-only transport works exactly the same way as OTV with multicast mode. The only difference is that each OTV device would need to create multiple copies of each control plane packet and unicast them to each remote OTV device that is part of the same logical overlay.

Two pieces of configuration are required to deploy OTV across a unicast-only transport infrastructure. First, you must define the role of adjacency server (usually enabled on a generic OTV edge device). The other piece of configuration is required in each OTV edge device that is not acting as an adjacency server (that is, acting as a client). All client OTV edge devices are configured with the address of the adjacency server. All other adjacency addresses are discovered dynamically. Thereby, when a new site is added, only the OTV edge devices for the new site need to be configured with the adjacency server addresses. No other sites need additional configuration.

Verifying OTV Adjacencies (Cont.)

- Use the **show otv overlay** command to verify that the overlay interface is enabled and that the essential parameters match on both sites:

```
N7K-1# show otv overlay 1

OTV Overlay Information

Overlay interface Overlay1

 VPN name            : Overlay1
 VPN state           : UP
 Extended vlans      : 10-12 (Total:3)
 Control group       : 239.7.7.7
 Data group range(s) : 232.7.7.0/24
 Join interface(s)   : Eth1/25 (10.7.7.5)
 Site vlan           : 13 (up)
```


To verify that the overlay interface is operational and that the correct parameters have been configured for the overlay interface, you can use the **show otv overlay** command. If adjacencies are not properly established, the primary fields to check in the output of this command are the “VPN state” and “control group” fields. The VPN state should be “up” and the control group should match on the local and remote edge device.

Verifying MAC Address Learning

- Use the **show otv route** command to verify that MAC addresses are properly learned and announced across the overlay:

```
N7K-1# show otv route

OTV Unicast MAC Routing Table For Overlay1

VLAN MAC-Address     Metric  Uptime    Owner      Next-hop(s)
---- --------------  ------  --------  ---------  -----------
  10 0005.9b1f.7c7c  42      02:46:53  overlay    N7K-2
  10 0005.9b1f.89fc  1       02:46:54  site       Ethernet1/9
```

- If you are knowledgeable about IS-IS, you can verify the announcement of MAC addresses in the OTV IS-IS database:

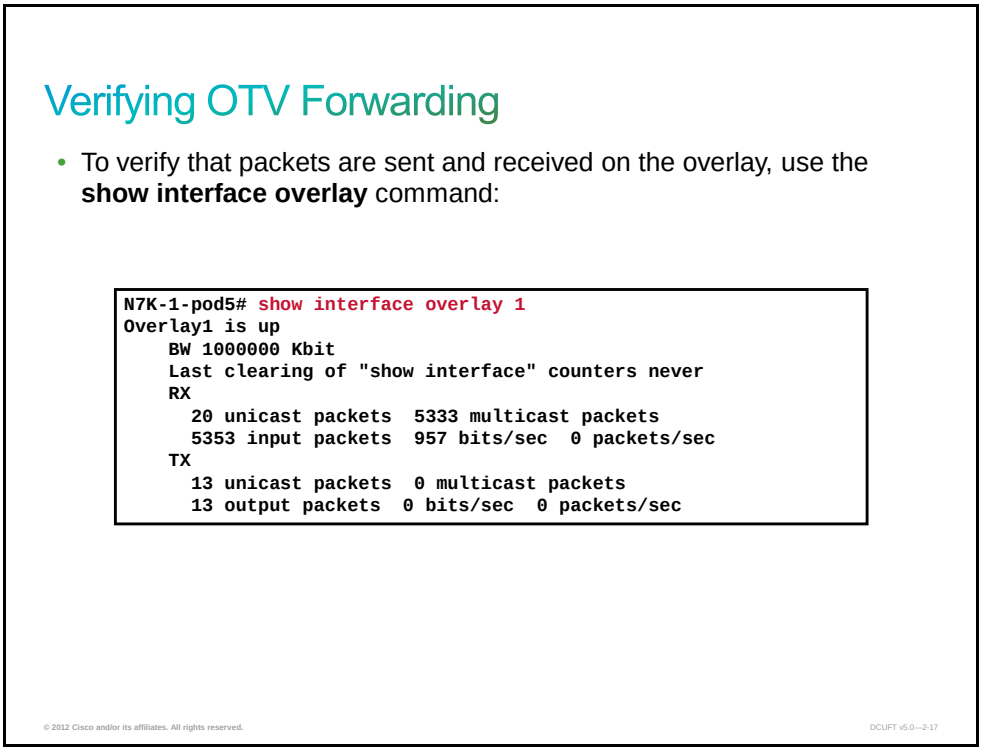
```
N7K-1# show otv isis database detail N7K-2.00-00 | include MAC
      MAC Address     : 0005.9b1f.7c7c
```


Once you have verified that the adjacencies were established, you should verify that the MAC addresses of the target hosts are properly advertised across the overlay. You can use the **show otv route** command to see all the MAC addresses that are learned for the extended VLANs. This command shows both local addresses that were learned on the internal interfaces and remote addresses that were learned through the overlay.

Tip	The **show otv route** command only shows MAC addresses for VLANs that are extended on the overlay. If a MAC address is displayed in the **show mac address-table** command for a VLAN on an internal interface, but the **show otv route** command does not show the address, then you should verify that the VLAN was added to the list of extended VLANs for the overlay.

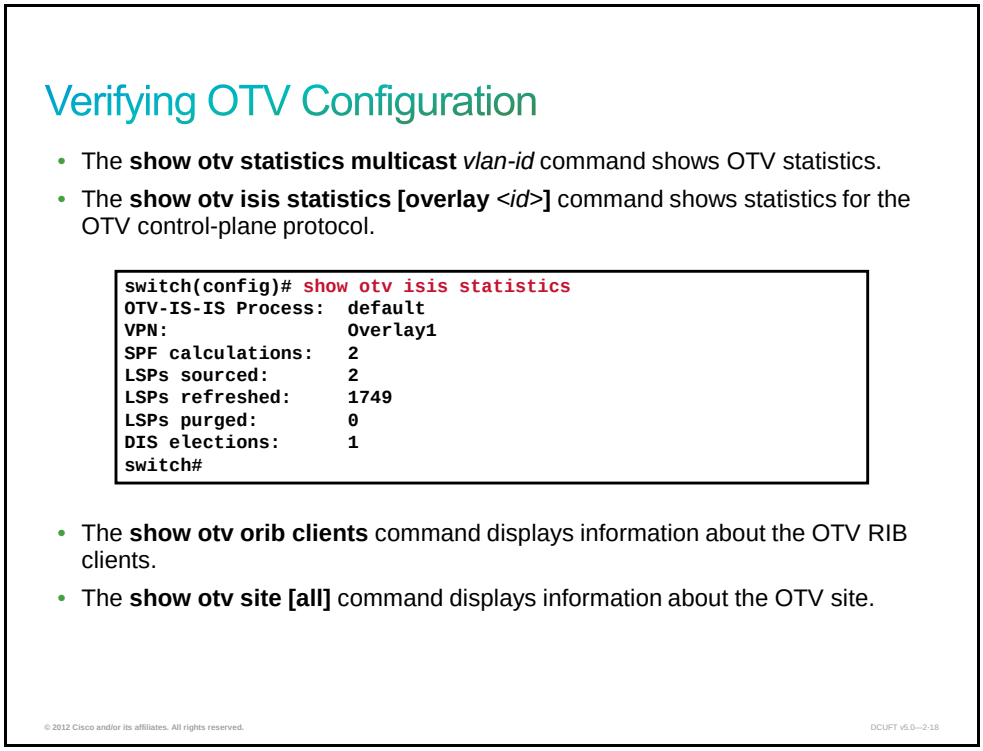
If you are proficient in troubleshooting IS-IS for IPv4 or IPv6, you can use that knowledge to troubleshoot the OTV IS-IS control protocol. For example, the OTV IS-IS database can be examined using the **show otv isis database** command.

For more specific troubleshooting, you could also use the **show otv overlay** *interface* **vlan** *vlan-id* command to see information about VLANs that are associated with an overlay interface.



If the OTV control plane looks correct, you need to verify data plane forwarding. In order to verify packet forwarding, generic techniques for troubleshooting IP unicast forwarding need to be used. For example, the OTV join interface could be monitored using the Switched Port Analyzer (SPAN) feature.

One of the OTV-specific commands that can be helpful in establishing whether packets are sent or received on the overlay is the **show interface overlay** command.



Finally, there are several other OTV verification commands:

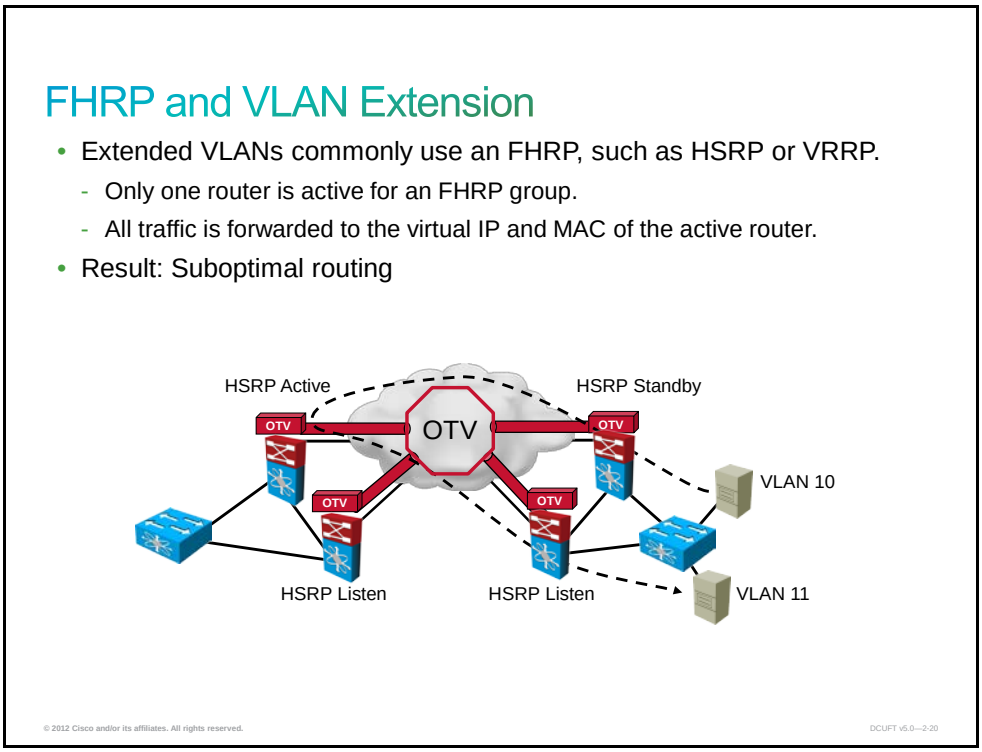
- The **show otv statistics multicast** *vlan-id* command shows OTV statistics.
- The **show otv isis statistics [overlay** *id***]** command shows statistics for the OTV control-plane protocol.
- The **show otv orib clients** command displays information about the OTV routing information base (RIB) clients.
- The **show otv site [all]** command displays information about the OTV site.

If there is any issue in establishing OTV IS-IS adjacency, you can look at the IS-IS adjacency log as follows:

```
SITE1-OED1# show otv isis internal event-history adjacency
ISIS default process
adjacency Events for ISIS process
2012 May 21 08:54:53.773678 isis_otv default [10371]: (Overlay1): LAN
adj L1
SITE2-OED1 over Overlay1 - UP
2012 May 21 08:54:53.773662 isis_otv default [10371]: [10376]: Sent
OTV add adjacency for overlay:Overlay1, addr: 10.10.17.6
2012 May 21 08:54:53.653906 isis_otv default [10371]: (Overlay1) : Set
adjacency
SITE2-OED1 over Overlay1 IPv4 address to 10.10.17.6
2012 May 21 08:54:53.653827 isis_otv default [10371]: (Overlay1) : LAN
adj L1
SITE2-OED1 over Overlay1 - INIT (New)
2012 May 21 08:54:53.653799 isis_otv default [10371]: [10376]:
Initialize adj for L1 MT-0 for iib Overlay1
```

HSRP Isolation Between Data Centers Using OTV

This topic describes the Hot Standby Router Protocol (HSRP) isolation between data centers using OTV.

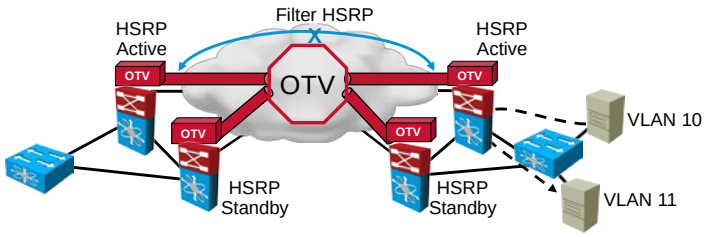


Data center networks commonly use a First Hop Redundancy Protocol (FHRP), such as HSRP, Virtual Router Redundancy Protocol (VRRP), or Gateway Load Balancing Protocol (GLBP). The combination of OTV and FHRPs could lead to suboptimal routing for VLANs that are extended across the overlay. When a VLAN is extended across multiple sites, only a single router will be elected as the active gateway and all traffic from that VLAN will go through that router.

As can be seen in the figure, this could cause traffic between two hosts that are in the same site, but on different VLANs, to be forwarded across the overlay to the other data center to be routed and then forwarded back across the overlay to the destination host. Clearly, this behavior is undesirable and should be prevented.

FHRP Filtering

- The FHRP suboptimal routing problem is solved by using FHRP traffic flitering.
- One active FHRP router per site
  - A VLAN access list filters the FHRP control packets
  - OTV MAC route filter stops the announcement of the FHRP virtual MAC address; a route map is applied to the OTV control protocol (IS-IS).

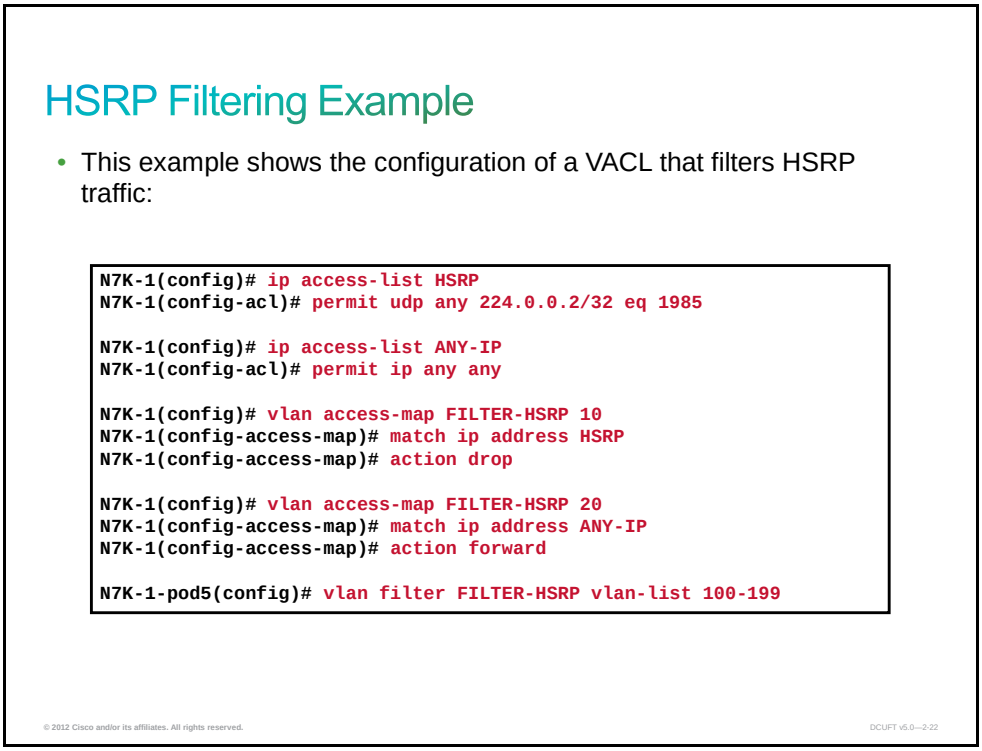


The desired behavior for the extended VLANs is to elect an FHRP active router per site, ensuring that traffic is routed locally when possible. In order to achieve the desired behavior, the FHRP traffic should be filtered from the overlay. If the routers in the different sites do not see hellos from each other, they will elect an active and standby router in each site. The active router in the site will be responsible for forwarding the traffic for the virtual IP and MAC address in that site.

To filter the FHRP traffic, two steps need to be taken. First, a VLAN access control list (VACL) must be implemented on the OTV edge device to block the FHRP hellos and prevent them from being forwarded on the overlay.

Second, the virtual FHRP MAC address should not be announced to other sites as it normally would. This can be achieved by implementing an OTV MAC route filter.

Even though HSRP traffic is filtered via the VACL defined in the first step, the virtual MAC used to source the HSRP packets is still learned by the OTV virtual device context (VDC). Therefore, OTV advertises this MAC address information to the other sites via an IS-IS update. While this in itself is not causing harm, it would cause the remote OTV edge devices to see constant MAC moves happening for the virtual MAC (from the internal interface to the overlay interface and vice versa). To prevent these MAC moves from being advertised and allow for a cleaner design, an OTV route map has to be configured.



The example in the figure shows how to filter HSRP packets on an OTV edge device. A VACL is implemented that matches HSRP traffic by its destination multicast address and UDP port number. The VACL drops these packets, but forwards any other IP traffic. The VACL is applied to VLANs 100–199, which represents the range of VLANs that are extended across the overlay.

HSRP Filtering Example (Cont.)

- This example shows the configuration of an OTV route filter that filters the announcement of the HSRP virtual MAC address:

```
N7K-1(config)# mac-list NOT-HSRP-VMAC deny 0000.0c07.ac00 0000.0000.00ff
N7K-1(config)# mac-list NOT-HSRP-VMAC permit 0000.0000.0000 ffff.ffff.ffff

N7K-1(config)# route-map NO-HSRP-ANNOUNCE permit 10
N7K-1(config-route-map)# match mac-list NOT-HSRP-VMAC

N7K-1(config)# otv-isis default
N7K-1(config-router)# vpn Overlay1
N7K-1(config-router-vrf)# redistribute filter route-map NO-HSRP-ANNOUNCE
```


The example in the figure shows how to configure an OTV route filter to filter the announcement of the HSRP virtual MAC addresses. First, a MAC address access list is created that permits all MAC addresses except the HSRP virtual MAC address block. Next, this access list is tied to a route map, because all route filtering in the Cisco NX-OS Software is performed through route maps. This route map filters out the HSRP MAC addresses defined in the access list. Finally, the route map is applied to the OTV IS-IS control protocol for the overlay.

The end result of the configuration is that HSRP remains strictly separated per site and each site has its own active HSRP router, which performs the inter-VLAN routing function.

Summary

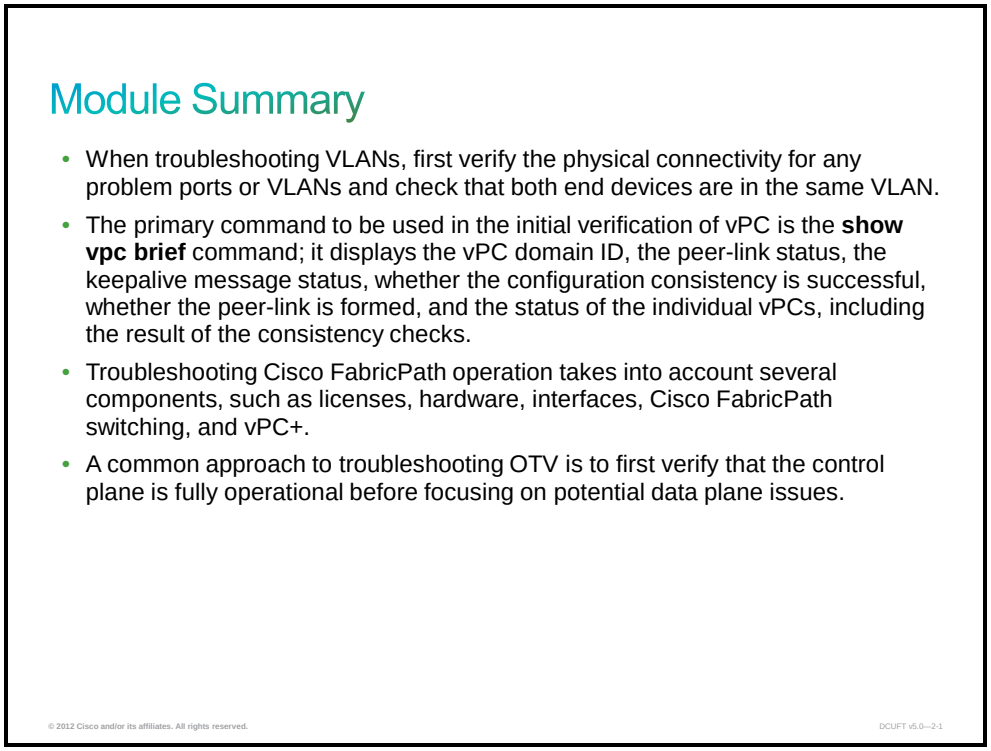
This topic summarizes the key points that were discussed in this lesson.

Summary

- OTV provides scalable Layer 2 extension across multiple data center sites.
- Troubleshooting OTV involves both control plane and data plane troubleshooting.
- To optimize the interaction between FHRP and OTV, filtering is required.

Module Summary

This topic summarizes the key points that were discussed in this module.



Usually, you would start troubleshooting the Layer 2 connectivity between devices because you have discovered that there is no Layer 3 connectivity between two adjacent Layer 2 hosts, such as two hosts in the same VLAN or a host and its default gateway.

Advanced Layer 2 switching features are available to build a solid Layer 2 foundation in the data center access and aggregation layers. Port channels and virtual port channels (vPCs) allow loop-free logical Layer 2 topologies to be created, which optimizes the use of bisectional bandwidth and increases the availability of the data center infrastructure.

Cisco FabricPath is a unique Cisco Nexus Operating System (NX-OS) Software feature that provides high availability and redundancy at a Layer 2 level in the form of multipathing capabilities. These capabilities bring the benefits of Layer 3 to a Layer 2 level, building highly resilient and scalable Layer 2 fabrics.

Overlay Transport Virtualization (OTV) provides Layer 2 connectivity between remote network sites by using MAC address-based routing and IP-encapsulated forwarding across a transport network. OTV provides support for applications that require Layer 2 adjacency, such as clusters and virtualization. OTV can be deployed on edge devices in each site, and requires no changes to the sites or transport network.

Module Self-Check

Use the questions here to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) In which three port modes can a Layer 2 interface on a Cisco Nexus switch be configured? (Choose three.) (Source: Troubleshooting VLANs and Private VLANs)

A) access
B) PVLAN
C) server
D) trunk
E) client

Q2) Which VLAN ID is allocated for internal use only on the Cisco Nexus switch? (Source: Troubleshooting VLANs and Private VLANs)

A) 1
B) 1000
C) 2000
D) 3000
E) 4000

Q3) Which feature must be enabled before configuring SVI? (Source: Troubleshooting VLANs and Private VLANs)

A) feature private-vlan
B) feature vlan
C) feature interface-vlan
D) feature fcoe

Q4) Promiscuous ports belong to the secondary VLAN. (Source: Troubleshooting VLANs and Private VLANs)

A) true
B) false

Q5) Which two options are possible LACP modes? (Choose two.) (Source: Troubleshooting Port Channels and Virtual Port Channels)

A) active
B) passive
C) auto
D) desirable

Q6) Which command displays the status of a port channel interface? (Source: Troubleshooting Port Channels and Virtual Port Channels)

A) **show interface port-channel** *channel-number*
B) **show port-channel traffic**
C) **show port-channel usage**
D) **show port-channel compatibility-parameters**

Q7) Which protocol is used for state synchronization and configuration validation between vPC peer devices? (Source: Troubleshooting Port Channels and Virtual Port Channels)

A) STP
B) CFS
C) NTP
D) HSRP

Q8) Which three options are the reasons for most vPC issues? (Choose three.) (Source: Troubleshooting Port Channels and Virtual Port Channels)

A) global configuration inconsistencies between the pair of switches that form the vPC domain
B) vPC specific configuration inconsistencies between the pair of switches that form the vPC domain
C) hardware failures
D) regular port channel inconsistencies

Q9) Cisco FabricPath is a Cisco NX-OS feature designed to bring the stability and performance of routing to Layer 2. (Source: Troubleshooting Cisco FabricPath)

A) true
B) false

Q10) Which protocol replaces STP as the control plane protocol in the Cisco FabricPath domain? (Source: Troubleshooting Cisco FabricPath)

A) BGP
B) IS-IS
C) HSRP
D) OSPF

Q11) What are the two tools that support the Cisco FabricPath troubleshooting process? (Choose two.) (Source: Troubleshooting Cisco FabricPath)

A) NX-OS CLI
B) FabricPath GUI
C) DCNM
D) ASDM
E) Cisco UCS Manager

Q12) Which command enables you to view the active Cisco FabricPath VLANs? (Source: Troubleshooting Cisco FabricPath)

A) **show fabricpath topology vlan active**
B) **show topology fabricpath vlan active**
C) **show fabricpath switch-id**
D) **show vlan**

Q13) Which two options are disadvantages of traditional data center interconnect technologies? (Choose two.) (Source: Troubleshooting OTV)

A) overhead associated with additional headers
B) complex operations
C) low bandwidth
D) transport dependent

Q14) Which protocol is used as the OTV control protocol? (Source: Troubleshooting OTV)

A) OSPF
B) IS-IS
C) BGP
D) MSDP
E) CFS

Q15) The **show otv adjacency** command can be used to verify whether the OTV control protocol adjacencies have been properly established. (Source: Troubleshooting OTV)

A) true
B) false

Q16) Which command enables you to see all of the MAC addresses that are learned for the extended VLANs? (Source: Troubleshooting OTV)

A) **show otv overlay**
B) **show otv adjacency**
C) **show interface overlay**
D) **show otv route**

Module Self-Check Answer Key

Q1) A, B, D

Q2) E

Q3) C

Q4) B

Q5) A, B

Q6) A

Q7) B

Q8) A, B, D

Q9) A

Q10) B

Q11) A, C

Q12) A

Q13) B, D

Q14) B

Q15) A

Q16) D

Module 3

SAN Switching Issue Troubleshooting

Overview

This module identifies common issues that relate to Cisco N-Port Virtualizer (NPV), zoning, SAN port channels, virtual storage area networks (VSANs), and Cisco Fabric Services. The module also presents methods for troubleshooting these issues. Understanding which tools to use when troubleshooting SAN switching issues is important. The ability to identify common configuration issues helps in the recovery of the fabric.

Module Objectives

Upon completing this module, you will be able to identify and resolve issues that relate to SAN switching in the Cisco data center architecture. This ability includes being able to meet these objectives:

- Identify and resolve issues that relate to Fibre Channel interface operation
- Identify and resolve issues that relate to Fibre Channel switching when the Cisco NX-OS switch is used in switched mode
- Identify and resolve issues that relate to Fibre Channel switching when the Cisco NX-OS switch is used in NPV mode

Lesson 1

Troubleshooting Fibre Channel Interfaces

Overview

This lesson is designed to provide you with some examples of common issues that relate to Fibre Channel connectivity and to show you how to identify and resolve these issues.

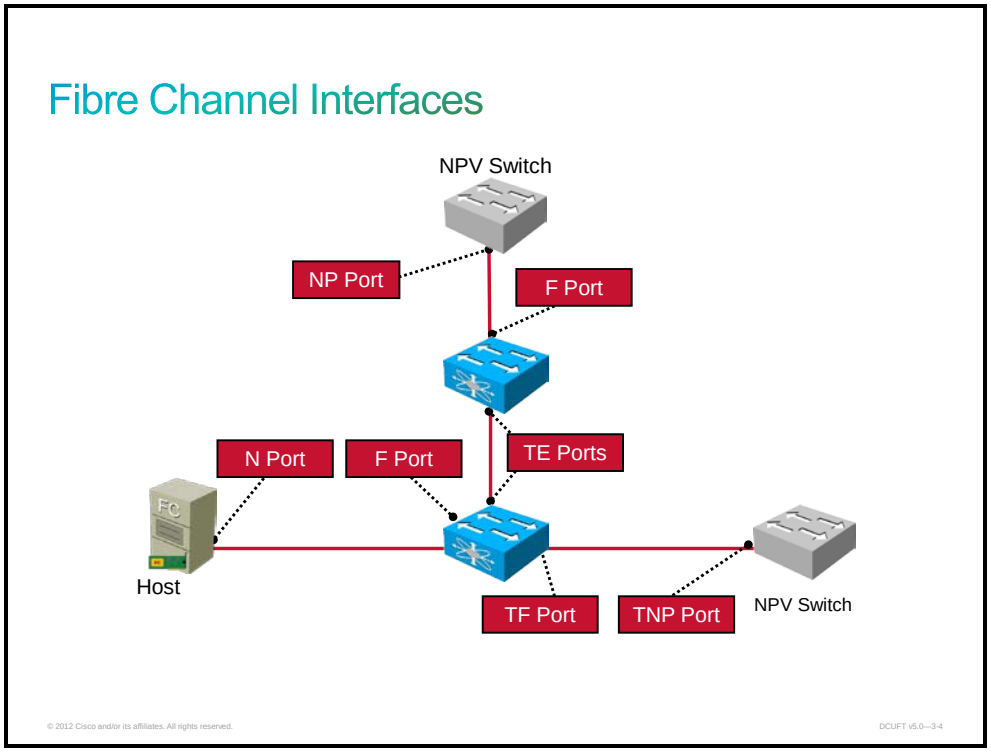
Objectives

Upon completing this lesson, you will be able to identify and resolve issues that relate to Fibre Channel interface operation. This ability includes being able to meet these objectives:

- Explain how to troubleshoot issues that relate to Fibre Channel ports on a Cisco Nexus or Cisco MDS Series switch
- Explain how to troubleshoot issues that relate to SAN port channel interfaces on a Cisco Nexus or Cisco MDS Series switch

Troubleshooting Fibre Channel Port Interfaces

This topic explains how to troubleshoot issues that relate to Fibre Channel port interfaces on a Cisco Nexus or Cisco MDS Series switch.



The most common problems that a SAN administrator might face can be categorized into three situations:

- **Switch-to-switch interconnectivity problems:** These types of problems can result in the isolation of a port or virtual storage area network (VSAN) because of incorrect parameters or settings on an Inter-Switch Link (ISL) or VSAN.
- **Node-to-switch connectivity problems:** These types of problems are identified by an Fx Port failure or caused by zone or VSAN configuration errors.
- **End-to-end connectivity problems:** These types of problems include misconfigured VSAN membership, VSAN trunking, or zoning.

A Fibre Channel interface supports multiple modes:

- **Expansion port (E Port):** In E Port mode, an interface functions as a fabric expansion port. This port is connected to another E Port to create an ISL between two switches. E Ports carry frames between switches for configuration and fabric management. These ports serve as a conduit between switches for frames that are destined to remote node ports (N Ports) and node loop ports (NL Ports).
- **Fabric port (F Port):** In F Port mode, an interface functions as an F Port. This port is connected to a peripheral device (host or disk) that operates as an N Port. An F Port can be attached to only one N Port.
- **Proxy N port (NP Port):** An NP Port operates on a device that is in N-Port virtualization (NPV) mode and that connects to the core switch via an F Port. NP Ports function like N Ports but also function as proxies for multiple N Ports.

- **Trunking expansion port (TE Port):** In TE Port mode, an interface functions as a TE Port. This port is connected to another TE Port to create an Enhanced Inter-Switch Link (EISL) between two switches. When an interface is in TE Port mode, all transmitted frames are in the EISL frame format, which contains VSAN information. Interconnected switches use the VSAN ID to multiplex traffic from one or more VSANs across the same physical link. TE Ports are specific to Cisco switches and expand the functionality of E Ports to support these functions:
  - VSAN trunking
  - Transport quality of service (QoS) parameters
  - Fibre Channel Traceroute (fctrace) feature
- **Trunking fabric port (TF Port):** In TF Port mode, an interface functions as a TE Port. This interface is connected to a trunking N Port (TN Port) or trunking NP Port (TNP Port) to create a link between a core switch and a Cisco NPV switch or a host bus adapter (HBA), to carry tagged frames. TF Ports are specific to Cisco switches and expand the functionality of F Ports to support VSAN trunking. In TF Port mode, all frames are transmitted in the EISL frame format, which contains VSAN information.
- **TNP Port:** In TNP Port mode, an interface functions as a TNP Port. This interface is connected to a TF Port to create a link to a core N-Port ID Virtualization (NPIV) switch from a Cisco NPV switch, to carry tagged frames.
- **Switched Port Analyzer (SPAN) destination port (SD Port):** In SD Port mode, an interface functions as a SPAN. An SD Port monitors network traffic that passes through a Fibre Channel interface. The port uses an attached, standard Fibre Channel analyzer (or similar switch probe). SD Ports cannot receive frames and only transmit a copy of the source traffic. This feature is nonintrusive and does not affect switching of network traffic for any SPAN source port.
- **Auto mode:** Interfaces that are configured in auto mode automatically detect the mode to which they should be configured during interface initialization. SD Ports must be configured administratively.

Fibre Channel Interface States

Each interface has an associated administrative and operational state.

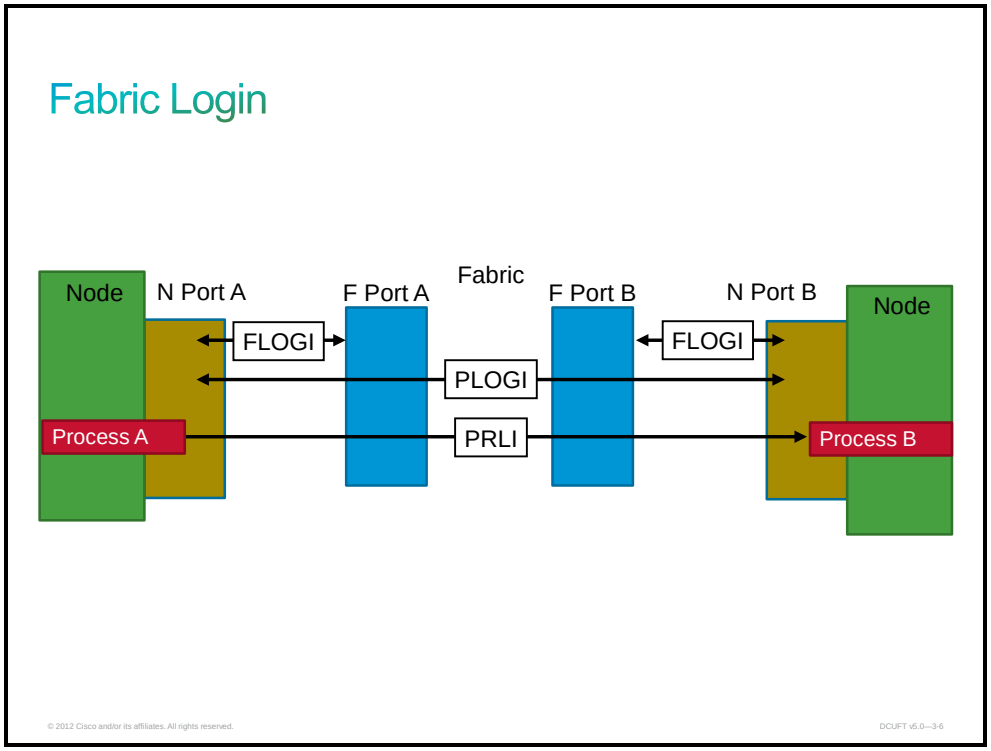
Administrative State	Description
Up	Use **no shutdown** command to enable the interface.
Down	Use **shutdown** command to disable the interface.

Operational State	Description
Up	Interface transmits and receives traffic as desired; administratively up, link layer state up, and interface initialization must be completed.
Down	Cannot transmit or receive data.
Trunking	Interface is operational in TE Port or TF Port mode.

Configuration allows for administrative enabling or disabling of an interface. Any interface that is administratively disabled does not function operationally.

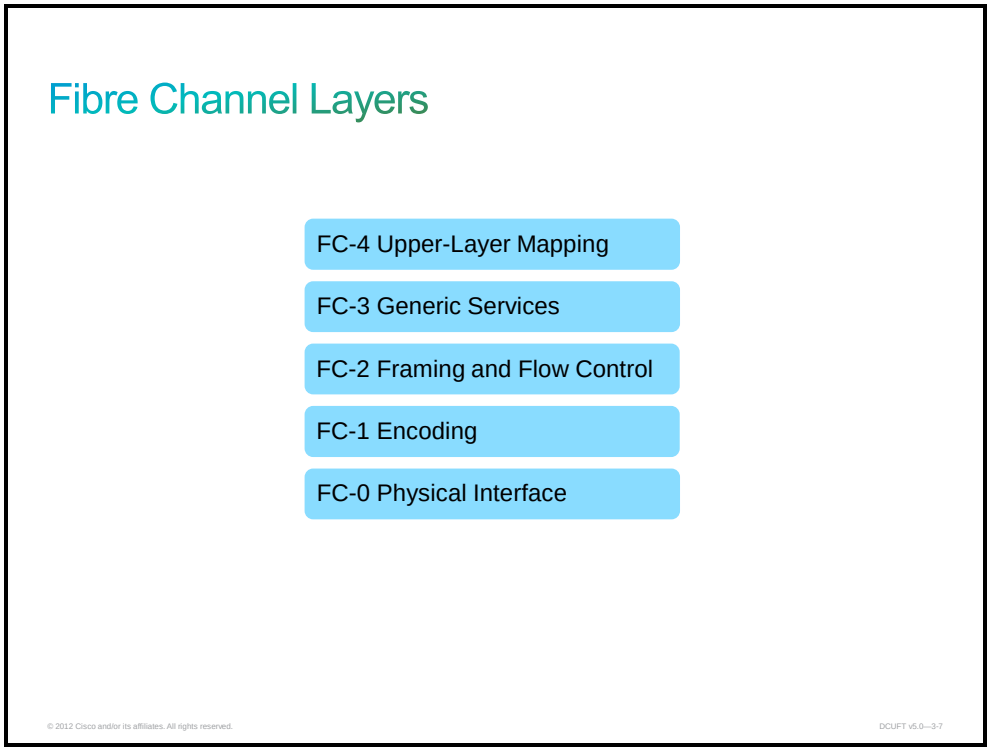
The **shutdown** command in the interface configuration submode disables a port. The **no shutdown** command enables the port and makes it administratively active.

The operational state can be down even when the data link layer is up because the operational state of an interface depends on the protocol that is running on the interface.



Before an N Port can begin exchanging data with other N Ports, three processes must occur:

- The N Port must log in to its attached F Port. This process is known as the fabric login (FLOGI).
- The N Port must log in to its target N Port. This process is known as the port login (PLOGI).
- The N Port must exchange information about upper-layer protocol (ULP) support with its target N Port, to ensure that the initiator and target process can communicate. This process is known as the process login (PRLI).



Fibre Channel is structured as a set of hierarchical functions.

The lowest level (FC-0) defines the physical link in the system, including the fibre, connectors, optical, and electrical parameters for various data rates.

The FC-1 level defines the transmission protocol, including serial encoding and decoding rules, special characters, and error control.

The signaling protocol (FC-2) level serves as the transport mechanism of Fibre Channel. FC-2 defines the framing rules of the data to be transferred between ports, the mechanisms for controlling the three service classes, and the means of managing the sequence of a data transfer.

The FC-3 level is intended to provide the common services that are required for advanced features such as striping, hunt groups, and multicast.

The highest level (FC-4) defines the application interfaces that can execute over Fibre Channel. FC-4 specifies the mapping rules of ULPs that use the lower Fibre Channel levels. Fibre Channel is equally adept at transporting both network and channel information and allows both protocol types to be concurrently transported over the same physical interface.

Troubleshooting Methodology

Identifying a SAN environment problem:

- Gather information.
- Verify physical connectivity and registration to the fabric.
- Verify storage subsystem and server configuration.
- Verify end-to-end connectivity and fabric configuration.

All commands for troubleshooting Fibre Channel can also be used for troubleshooting FCoE.

When a SAN administrator receives a call from network users inquiring as to why access to their storage has ceased, the challenge is to first determine whether the users are having a connectivity problem because of some anomaly within the SAN or whether some change on the user server created the problem. Assuming that a SAN infrastructure problem is the cause, the administrator must determine where the disconnect occurred.

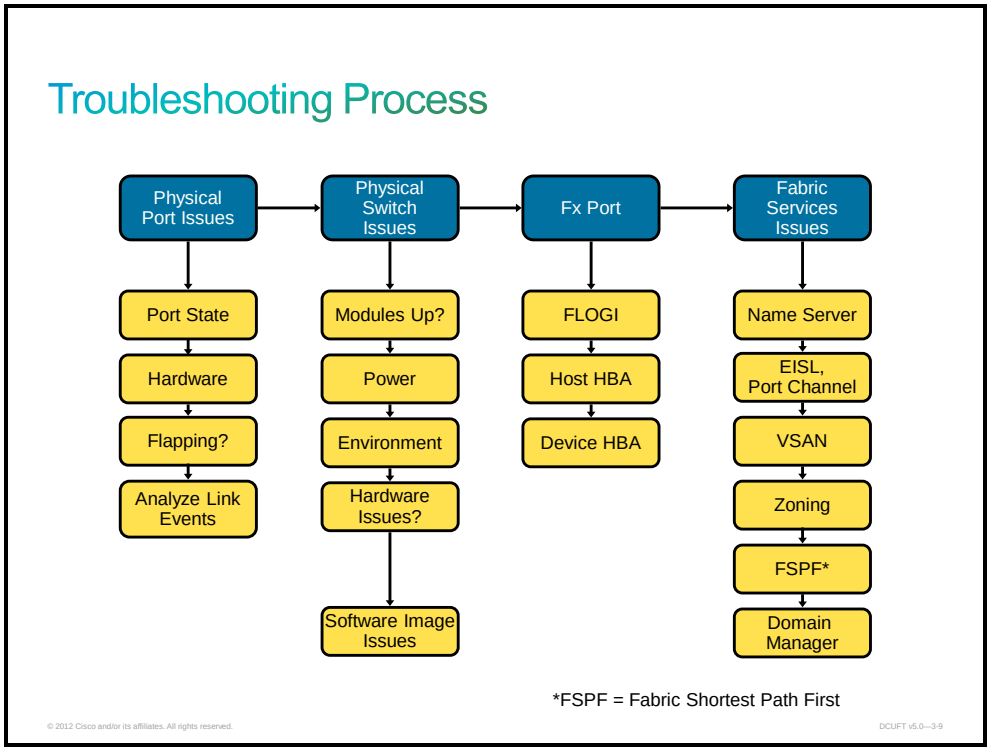
To identify a SAN environment problem, take these actions:

Step 1 **Gather information:** This effort can prove challenging. Users rarely understand that changes that they implement can cause the disconnect between the application server and the storage that is assigned to that server. First, verify that the server-storage connection had been working. Often, the resolution to the problem is on the server end. For example, the Fibre Channel HBA of the server might have been swapped without the administrator being properly notified. If the administrator satisfactorily determines that no changes have been implemented on the server end, the administrator must now investigate the location of the SAN problem.

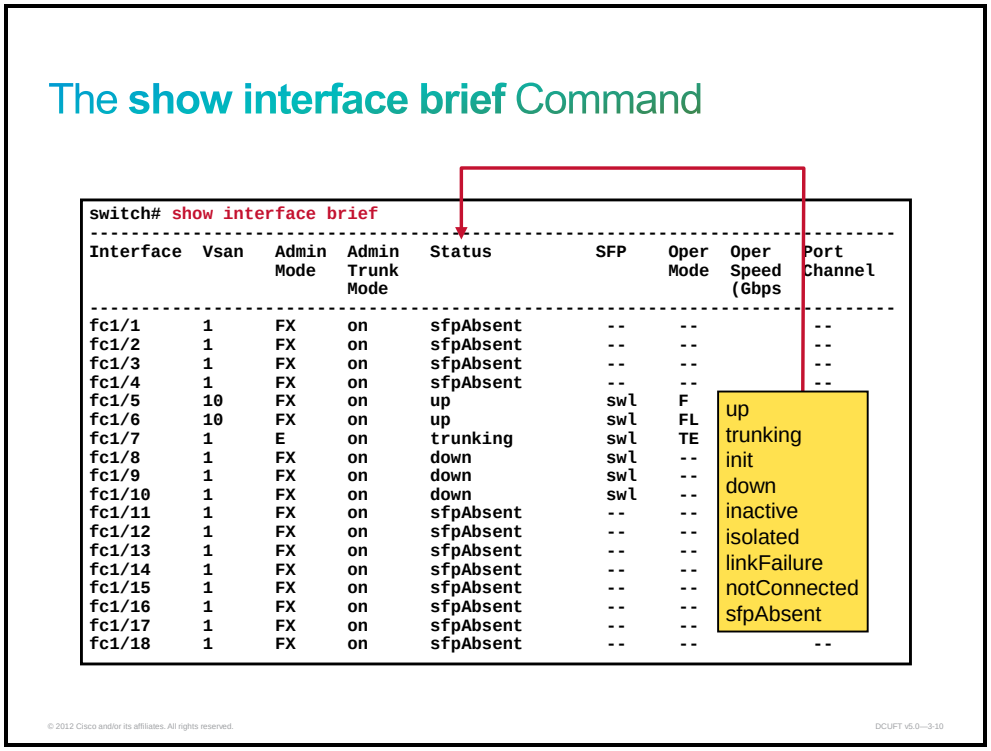
Step 2 **Verify physical connectivity and registration to the fabric:** Are all devices connected to their respective switches and registered to the name server?

Step 3 **Verify storage subsystem and server configuration:** If no changes have been implemented on the server, determine whether changes on a storage device, such as logical unit number (LUN) masking, have affected connectivity.

Step 4 **Verify end-to-end connectivity and fabric configuration:** Assuming that connectivity and name server registration are successful, the problem might be within the network infrastructure. For example, an ISL might have failed, a new switch might have joined the fabric and caused fabric disruption or domain isolation, the name server might not have synchronized, or a zone configuration might have changed.



This figure shows the overall troubleshooting process flow.



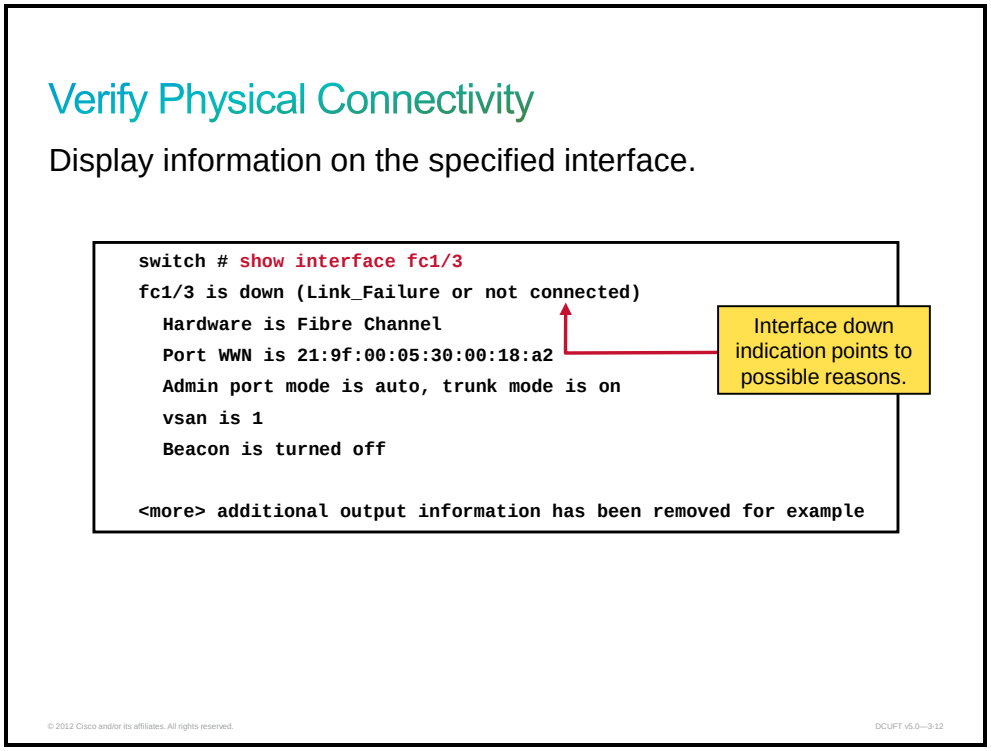
When you monitor interfaces, the **show interface brief** command provides broad information in an easily viewed, tabular format. Scan the Status column to view the operational state for each interface. On Cisco Nexus 5500 Platform switches, there are unified ports (LAN or Fibre Channel) and Fibre Channel ports (on the expansion modules).The operational state can be any of these entries:

- **up:** A Fibre Channel FC-2 layer link has been established. The attached node has completed a FLOGI.
- **trunking:** The port is up and VSAN trunking is enabled.
- **init:** The link is initializing. (Note that a port can get stuck in this state. This issue is explained later in the course.)
- **down:** An FC-1 layer link has been established, but the port is administratively disabled.
- **inactive:** A VSAN is suspended or deleted.
- **isolated:** Isolation has been caused by an Exchange Switch Capabilities (ESC) or exchange link parameter (ELP) failure, invalid domain ID, Fibre Channel domain overlap, disabled Cisco Domain Manager, failed zone or fabric merge, VSAN mismatch, or isolated peer E Port.
- **linkFailure:** The physical layer is not operational.
- **notConnected:** The physical layer is not operational, or there is no active device connection.
- **sfpAbsent:** A Fibre Channel small form-factor pluggable (SFP) transceiver is not present in the port.

Displaying the Interfaces in Cisco DCNM

- Choose **Dashboard > Switches** and click a specific switch.
- Choose the **Interfaces** tab.

When monitoring interfaces, you can also use Cisco Data Center Network Manager (DCNM). Choose **Dashboard > Switches,** choose a specific switch, and then choose the **Interfaces** tab.



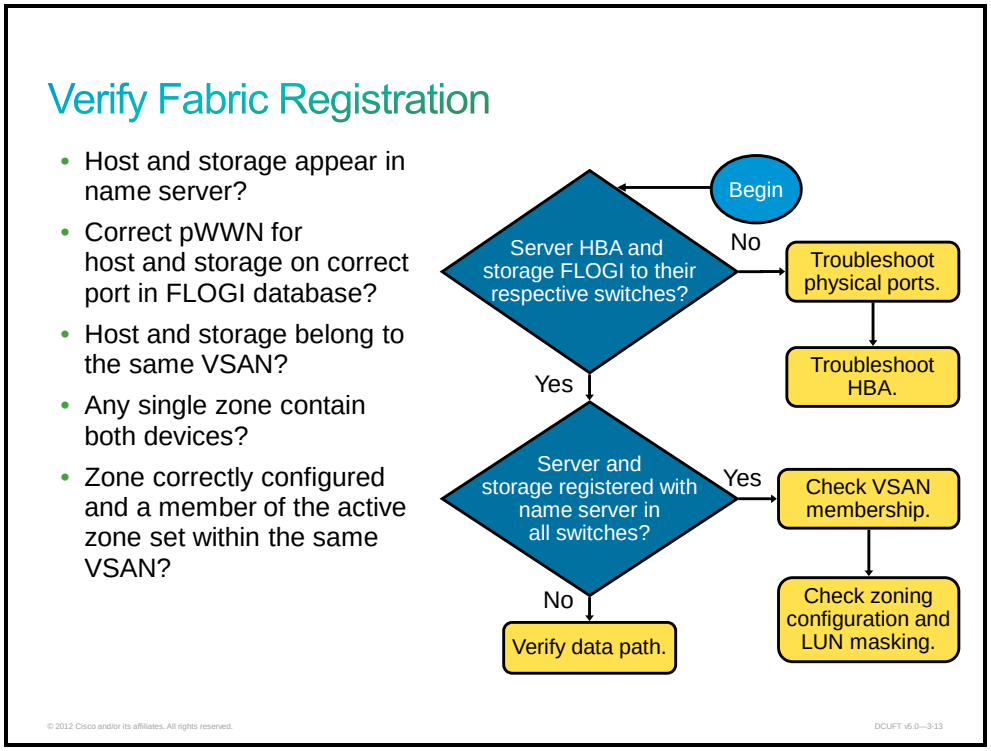
Before troubleshooting physical connectivity, ask some basic questions:

- Are you using the correct fiber type (single-mode fiber [SMF] or multimode fiber [MMF])?
- Has the fiber been checked for proper connection?
- Is the connection broken in any way?
- Is the LED on the connected module port green?
- Do the LEDs on any HBA or storage subsystem ports indicate normal functionality?

These basic questions can help save valuable time before you start more involved troubleshooting tasks. If additional troubleshooting is required, you can initiate various CLI commands to isolate the problem.

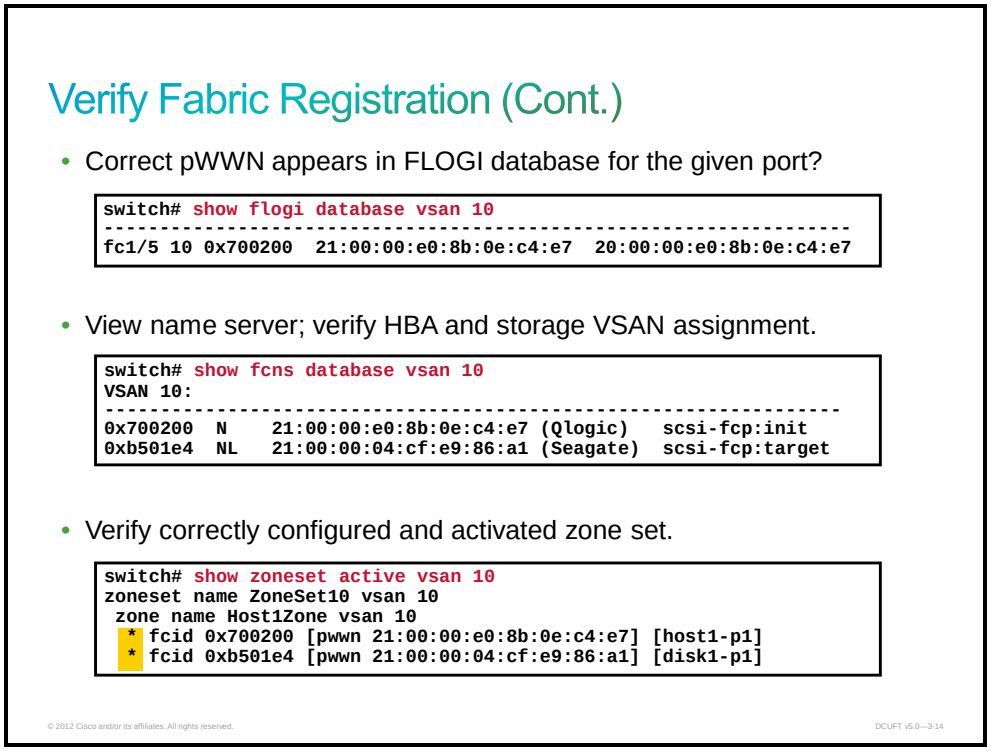
For example, the “SFP not present” and “Link failure or not connected” messages are two possible results of a **show interface** command when the interface is in the down condition. The example in the figure shows that interface 3 on the Fibre Channel switch module slot number 1 is down. The “Link failure or not connected” message can occur when nothing is connected to the specific interface, such as when a fiber is unplugged or broken, or when there is no bit synchronization between the switch interface and the Nx Port that connects directly to it. Bit synchronization is lost if the receive (Rx) path of the bit stream from the connected Nx Port is nonoperational. Check the fiber connections to both the Fx Port on the switch and the Nx Port on the storage device.

If the state of the **show interface** command indicates that SFP is not present, then the switch does not detect the presence of an SFP transceiver on the interface. In this case, you should verify that the SFP transceiver on the interface is seated properly. If reseating the SFP does not resolve the issue, then replace the SFP transceiver or try another port on the switch.



As the figure illustrates, the first step verifies that the end-node ports activate and that the devices successfully perform a FLOGI. If the devices fail to connect, you should troubleshoot physical port issues. If a host can see some, but not all, LUNs on an existing subsystem, then neither physical nor fabric-specific issues are the problem. If no LUNs are available, then investigate these questions:

- Do the host and storage belong to the same VSAN?
- Are both the HBA and the subsystem port successfully registered with the fabric name server?
- Does the correct port world wide name (pWWN) for the server HBA and the storage subsystem port show up on the correct port in the FLOGI database? In other words, is the device plugged in to the correct port?
- Does any single zone contain both devices? The zone members can be world wide names (WWNs) or Fibre Channel IDs (FCIDs).
- Is the zone configured correctly, and is it part of the active configuration or zone set within the same VSAN?



To troubleshoot port registration, follow these steps:

Step 1 Ensure that the Fibre Channel interface that is connected to the device is active and free of errors. If the interface is not working correctly, check the cabling and the host or storage device interface for faults. If the interface is working correctly, proceed to the next step.

Step 2 Verify that the device appears in the FLOGI database.

Step 3 Verify that the device appears in the Fibre Channel Name Server (FCNS) database.

Step 4 If the device does not appear in the FLOGI or FCNS database, verify that the device is configured properly. If the device is a host, check the device driver configuration and startup sequence.

Step 5 If the device appears to be configured correctly, reset the interface and view the FLOGI process to try to determine what happens when the device attempts to log in to the interface.

Step 6 If you still cannot resolve the problem, you can attempt to debug the registration process.

Viewing the FLOGI Database

For an F Port to come up, the switch port must first acquire bit and word synchronization with the connected N Port. The switch port must then receive the FLOGI that the N Port issued. If any one of these steps fails, the switch port will not come up as an F Port.

One of the first steps in troubleshooting this situation is to invoke the **show interface** command from the CLI. This command displays where the process of interface initialization has stopped, which determines the steps that are needed to solve the problem.

In a Fibre Channel fabric, each host and storage port requires an FCID. Use the **show flogi database** command to verify that the device is displayed in the FLOGI database in its respective switch.

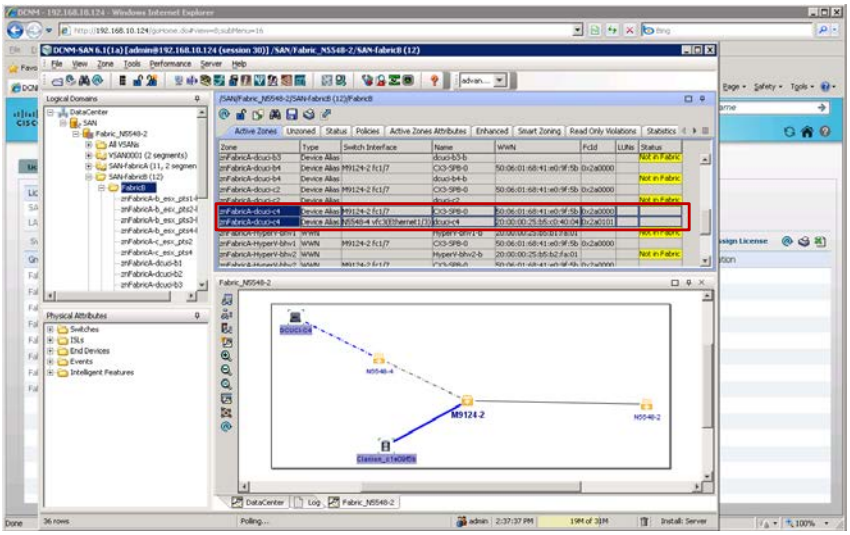
Viewing the FCNS Database

If both devices are registered in the FLOGI database on their respective switches, invoke the **show fcns database** command on each switch to verify that the name server has properly propagated. If only one of the end devices is visible, an ISL connection problem or name server issue might be at fault.

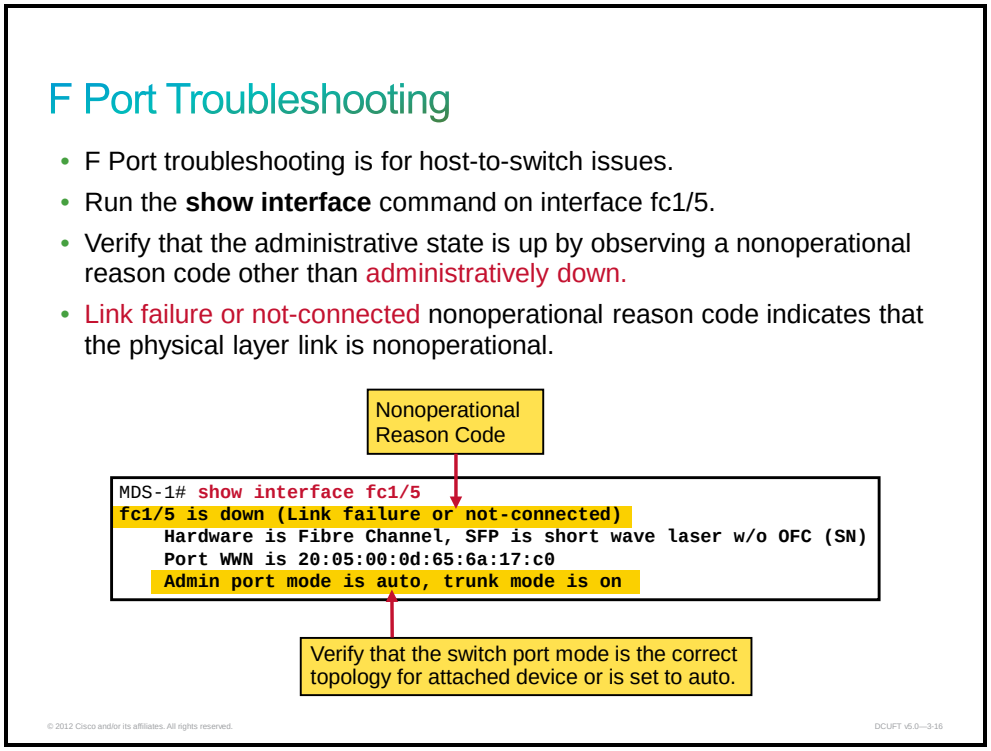
Verify Zoning Configuration

If both end devices are registered in the name server and cannot communicate, verify the zone configuration. Is there an active zone set, and are both devices visible? Look for asterisks (*) beside each zone entry from the **show zoneset active** command.

Verify Fabric Registration in Cisco DCNM for SAN



You can also verify fabric registration by using the Cisco DCNM. Open Cisco DCNM-SAN and choose the appropriate SAN ID. In the example in the figure, you can see that two devices are visible in the specified zone.



F Port troubleshooting is for host-to-switch issues. In this figure, a **show interface** command is executed on fc1/5. You can see that fc1/5 is down and is displaying a nonoperational code "Link failure or not connected". This code indicates that the physical layer connectivity between the host and fc1/5 is nonoperational but that the port is administratively up because it does not reflect the nonoperational, administratively down status.

Other nonoperational reason codes are possible:

- **SFP not present:** The SFP hardware is not plugged in.
- **Initializing:** The physical layer link is operational and the protocol initialization is in progress.
- **Reconfigure Fabric in progress:** The fabric is being reconfigured. The offline Cisco Nexus Operating System (NX-OS) waits for the time specified by the resource allocation timeout value (R_A_TOV) before retrying initialization.
- **Inactive:** The interface VSAN is deleted or in a suspended state. To make the interface operational, assign that port to a configured and active VSAN.
- **Hardware failure:** A hardware failure is detected.
- **Error disabled:** Error conditions require administrative attention. Interfaces can be error-disabled for various reasons:
  - Configuration failure
  - Incompatible buffer-to-buffer credit configuration

To make the interface operational, you must first fix the error conditions that are causing this state. Then, you must administratively shut down or enable the interface.

F Port Troubleshooting (Cont.)

- Use the **clear counters** command to clear the interface counters for fc1/5.
- Use **shutdown** and **no shutdown** commands to reinitialize the interface.

```
MDS-1# clear counters interface fc1/5
MDS-1# conf
MDS-1(config)# interface fc1/5
MDS-1(config-if)# shutdown
MDS-1(config-if)# no shutdown
```

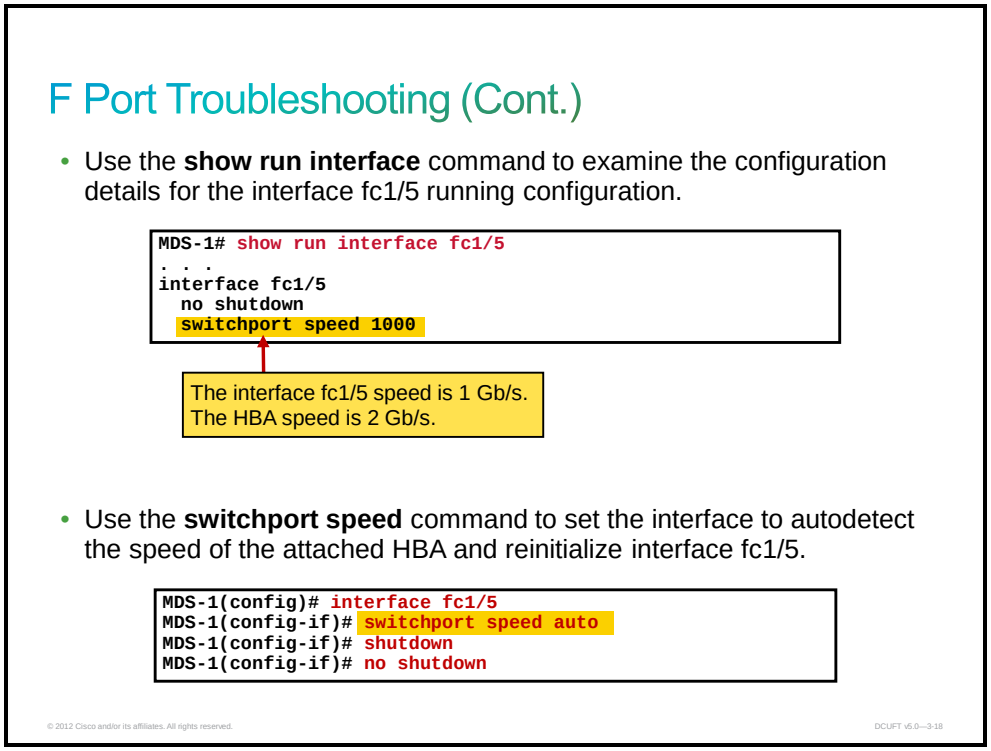
- Run the **show interface** command again on fc1/5 to verify the nonoperational reason code and observe link initialization ordered set counter values.

```
MDS-1# show interface fc1/5
fc1/5 is down (Link failure or not-connected)
. . .
      0 input OLS, 0 LRR, 0 NOS, 0 loop inits
      0 output OLS, 0 LRR, 0 NOS, 0 loop inits
```

Absence of FC-1 counters indicates that the interface is not achieving bit or word synchronization.

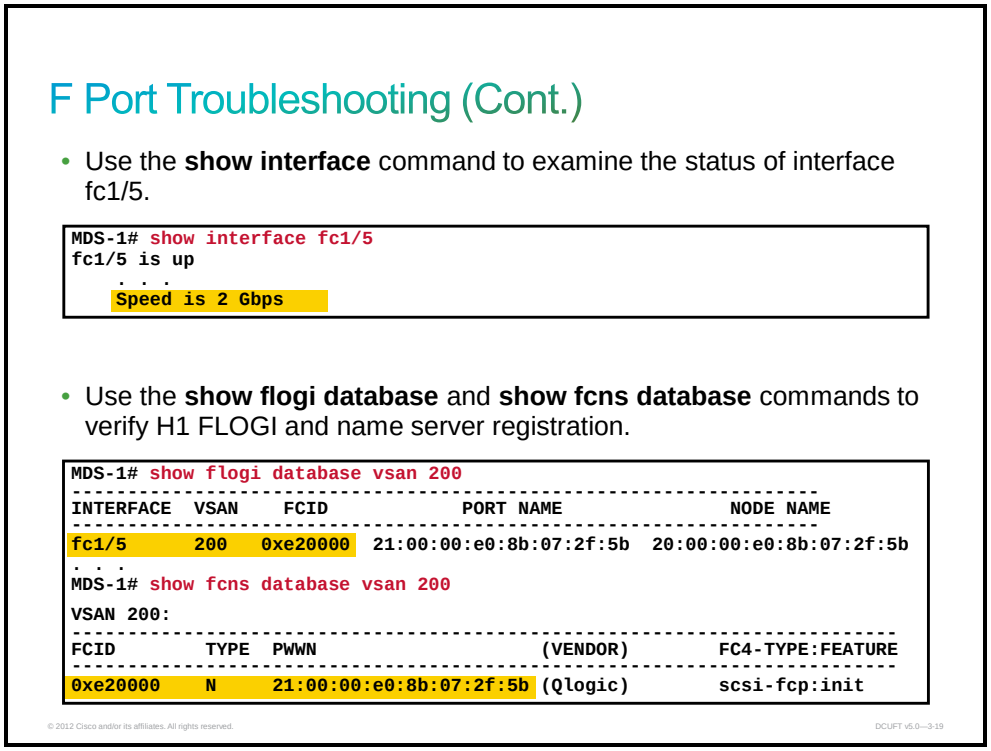
Clearing the interface counters on fc1/5 will help the administrator to determine which type of failure has occurred. After the interface counters have been baselined through use of the **clear counters interface fc1/5** command, the interface is initialized through the use of the **shutdown** and **no shutdown** sequence.

The absence of counter values in the link initialization ordered-set counter values indicates that the HBA in the host and interface fc1/5 on the Cisco MDS Series switch cannot achieve FC-0 bit synchronization or FC-1 word synchronization. One reason might be a configuration mismatch between the speed of interface fc1/5 and the HBA installed in the host.

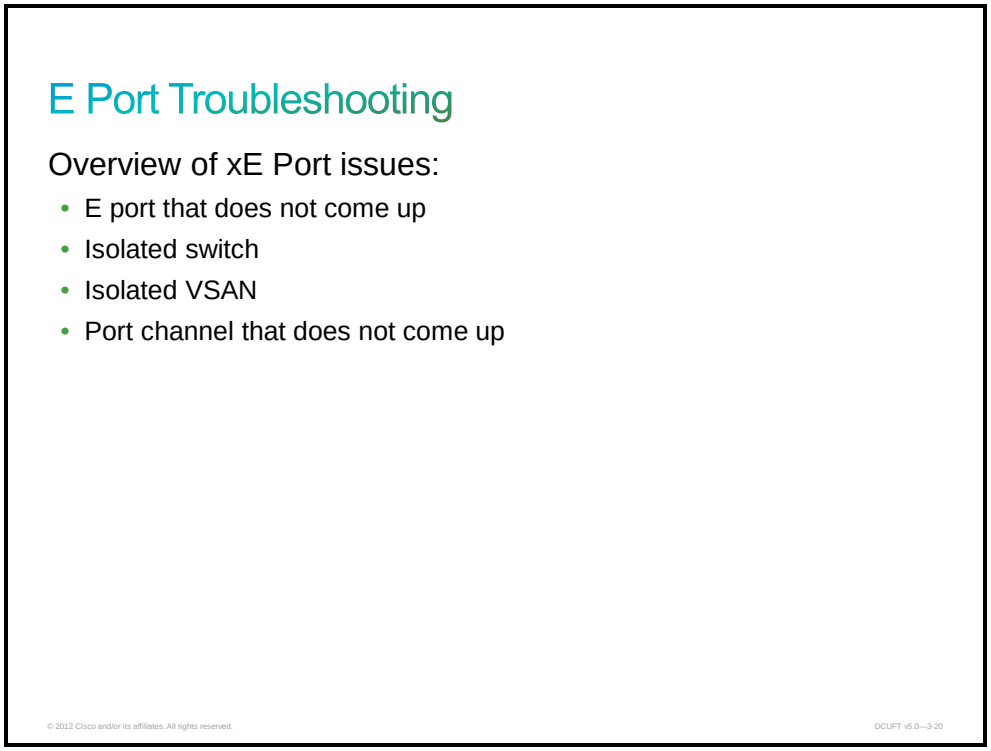


In this figure, a **show run interface** command is executed on fc1/5 to display the running configuration of the interface. The command output identifies a configuration mismatch between the speed of interface fc1/5 and the HBA that is installed in the host. This configuration error is preventing the HBA from initializing the link with the interface. The administrator has chosen to remedy the configuration error by using the **switchport speed auto** command to set interface fc1/5 to autodetect the speed of attached devices. Another valid alternative would be to use the **switchport speed 2000** command to hardcode the speed of interface fc1/5 to match the configured speed (2 Gb/s) of the HBA in the host.

After the speed has been changed to autodetect, the interface is reinitialized by using the **shutdown** and **no shutdown** sequence.

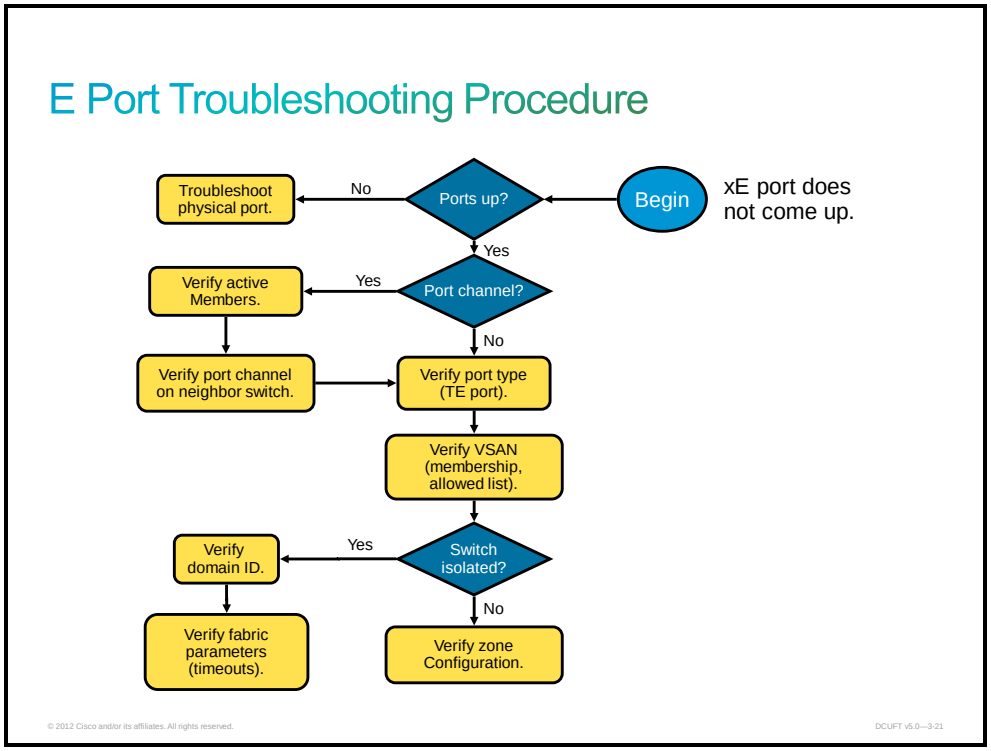


In this figure, a **show interface** command is executed on fc 1/5. The first few lines of the output provide all the information that is needed to verify that the host has initialized with interface fc1/5 and that the link is both administratively and operationally up, running at 2 Gb/s. The host login to the fabric and registration with the name server are then verified by using the **show flogi database vsan 200** and **show fcns database vsan 200** commands. Before the fix, both sets of output were empty.



E Port troubleshooting is for switch-to-switch issues. These are typical symptoms of such problems:

- An isolated fabric
- An isolated switch
- An E Port that does not come up
- A port channel that does not come up

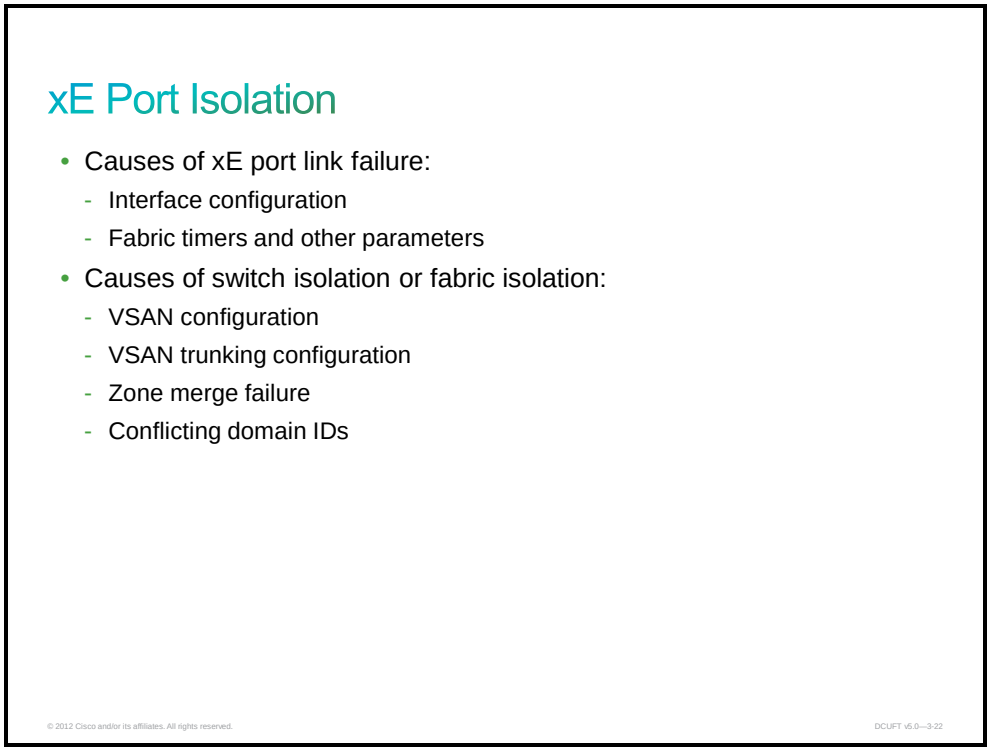


The E Port troubleshooting procedure is much like F Port troubleshooting. The difference involves the focus of your troubleshooting; that is, for E Ports, you must investigate switch-to-switch issues instead of device-to-switch issues.

On an E Port, only one VSAN can be passed and possibly isolated. However, in a TE Port, multiple VSANs can become isolated while others are passing traffic. The same troubleshooting approach applies in both cases, except that on a TE Port the troubleshooting might need to be done on a per-VSAN basis or on multiple VSANs.

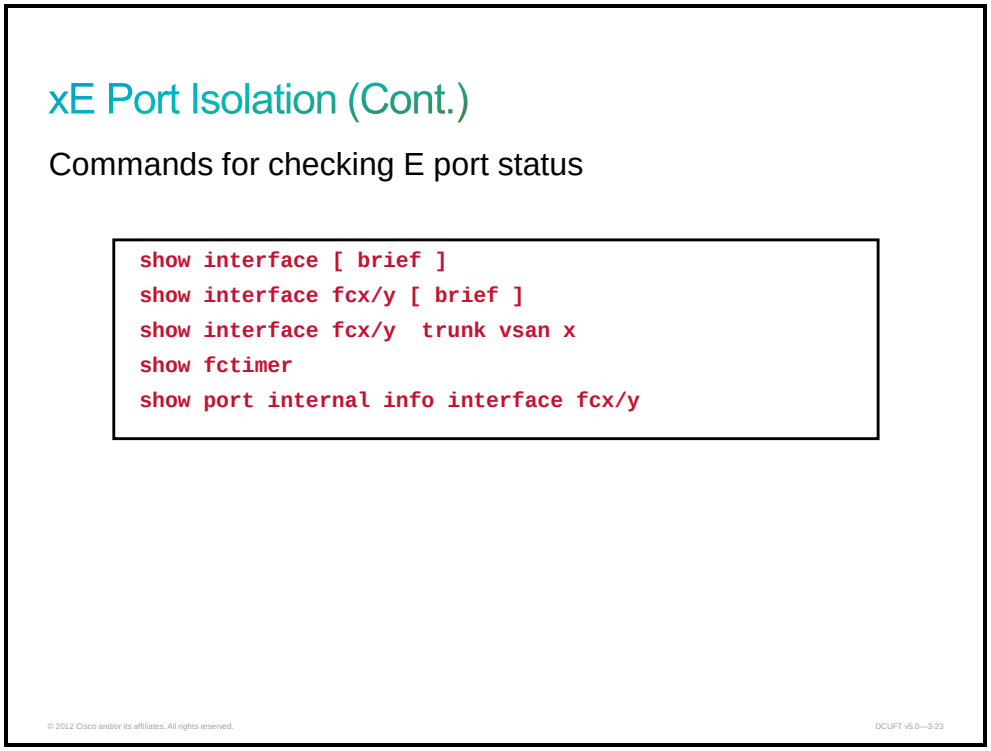
Misconfigured fabric or zone parameters can result in two types of isolation:

- ISL isolation can occur when you try to merge two separate fabrics or add a new domain to an existing fabric and there is a domain ID conflict or a mismatch in fabric timeout values.
- E Port isolation can result from a failed zone merge.

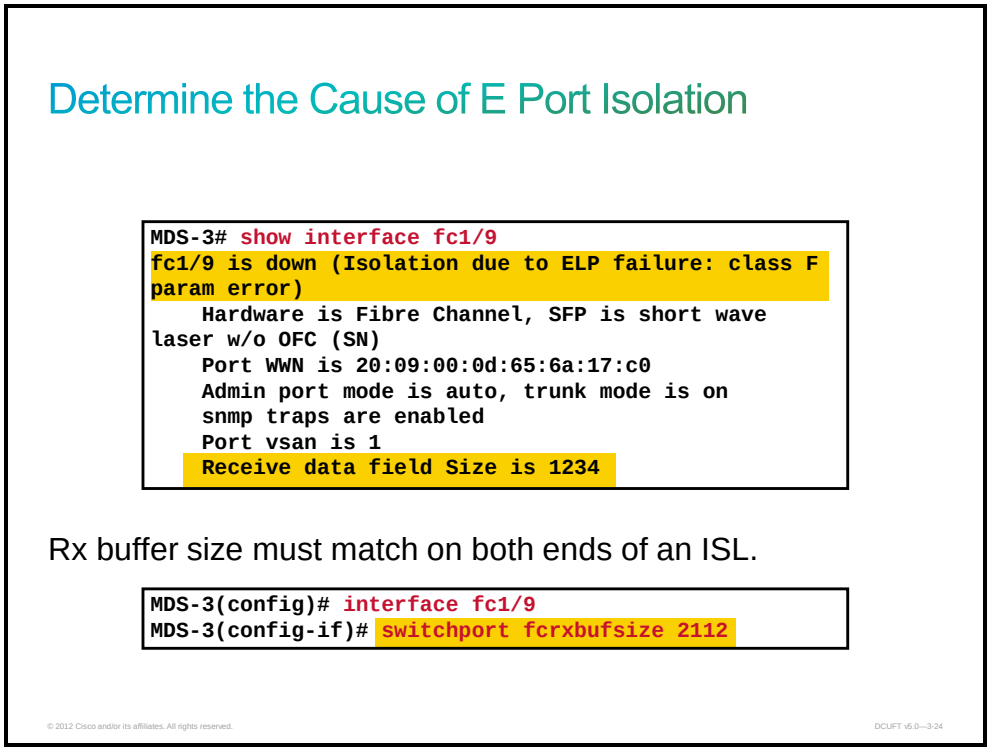


E Port isolation causes can range from mismatched port configuration settings to incorrect zoning and fabric configuration settings:

- **ELP failure and mismatch in the switch or port parameters:** Port configuration settings, such as speed or port mode, can prevent an ISL from achieving initialization.
- **Zone merge failure:** A port becomes isolated if the zone configuration settings do not match.
- **Port VSAN mismatch:** A mismatch in VSAN configuration can result in xE port isolation.
- **Domain overlap:** As a switch attempts to join a fabric, it becomes isolated if its configured domain ID is already in use.
- **Invalid fabric reconfiguration:** If timeout values do not match, for example, a switch isolates itself from the neighboring switch.



The first step in the process is to determine the nature of the problem by checking the status of the E Port. You can use the **show interface** command to display the port status. If an error has occurred, this command displays the protocol error or configuration mismatch that caused the problem.



In the example in the figure, the **show interface** command indicates that the E Port did not come up because of an ELP failure. The receive buffer size must match on both ends of an ISL. The default size is 2112.

To configure the receive buffer size, choose the interface and use the **switchport fcrxbufsize** command:

```
MDS-3(config)# interface fc1/9
MDS-3(config-if)# switchport fcrxbufsize 2112
```


show fctimer Command

Verify that Fibre Channel timeout values (TOVs) match on all switches.

```
MDS-3# show fctimer ?
 <CR>
  >             Redirect it to a file
  >>            Redirect it to a file in append mode
  D_S_TOV       D_S_TOV in milliseconds
  E_D_TOV       E_D_TOV in milliseconds
  F_S_TOV       F_S_TOV in milliseconds
  R_A_TOV       R_A_TOV in milliseconds
  last          Show the status of the last cfs commit/abort operation
  pending       Show the status of pending fctimer commands
  pending-diff  Show the difference between pending database and running
config
  session       Show the state of fctimer cfs session
  status        Cfs distribution is enabled or disabled
  vsan          Specify VSAN id
  |             Pipe command output to filter

MDS-3# show fctimer vsan 100

vsan no.  F_S_TOV   D_S_TOV   E_D_TOV   R_A_TOV
-------------------------------------------------
100       5000 ms   5000 ms   2000 ms   10000 ms
```


A mismatch in the configured switch or port parameters affects the link-initialization process and eventually the initial ELP exchange. (The ELP frame is sent between two switches to negotiate fabric parameters.) If this failure occurs, verify that the fabric parameters are the same for both switches. Although VSAN parameters are configured on a per-switch basis, they must be the same for all switches within a fabric.

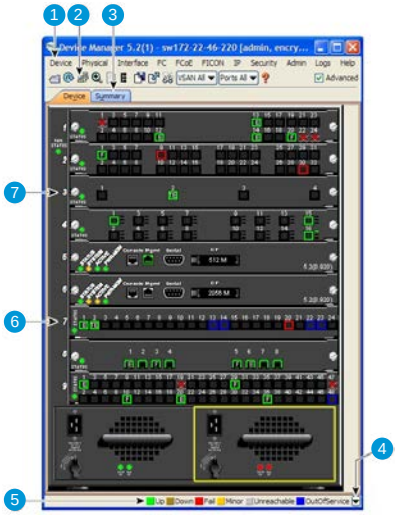
If the interface indicates an ELP failure, use the **show fctimer** command to verify that these parameters match:

- **Error detect timeout value (E_D_TOV):** The E_D_TOV determines the life of an individual Fibre Channel frame in any particular fabric element. The effects of the E_D_TOV on the fabric as a whole are typically cumulative because each fabric element contains its own E_D_TOV timers for any received frame.
- **R_A_TOV:** The R_A_TOV determines the life of an individual Fibre Channel frame in the fabric as a whole. For a fabric, the R_A_TOV implies that no particular frame remains in (and can thus be emitted from) the fabric after the timer expires.
- **Fabric stability timeout value (F_S_TOV):** The F_S_TOV is used to ensure that fabric stability has been achieved during fabric configuration. The F_S_TOV determines how long the fabric waits for the principal switch selection to complete.
- **Distributed services timeout value (D_S_TOV):** The D_S_TOV determines how long a switch service waits for a reply from another instance of that service on another switch.

Note The D_S_TOV must be greater than the E_D_TOV.

Cisco Device Manager

Cisco Device Manager: Device View



Number	Description
1	Menu bar
2	Toolbar
3	Tabs
4	Legend
5	Status
6	Supervisor modules
7	Switching or services modules

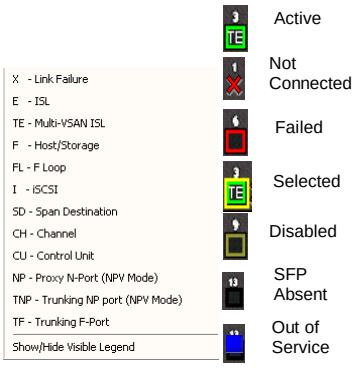
From Cisco DCNM, you can click DM to open Cisco Device Manager. Cisco Device Manager provides a graphical representation of a Cisco MDS 9000 Series switch chassis or Cisco Nexus 5500 Platform switch chassis, including the installed switching modules, the supervisor modules, the status of each port within each module, the power supplies, and the fan assemblies.

Cisco Device Manager provides two views: Device View and Summary View. Use Summary View to monitor interfaces on the switch. Use Device View to perform switch-level configurations:

- Configure virtual Fibre Channel interfaces.
- Configure Fibre Channel over Ethernet (FCoE).
- Configure zones for multiple VSANs.
- Manage ports, port channels, and trunking.
- Manage Simple Network Management Protocol version 3 (SNMPv3) security access to switches.
- Manage CLI security access to the switch.
- Manage alarms, events, and notifications.
- Save and copy configuration files and software image.
- View hardware configuration.
- View chassis, module, port status, and statistics.

Cisco Device Manager (Cont.)

Cisco Device Manager: Port display descriptions



Basic port monitoring, using Cisco Device Manager, begins with the visual display in Device View. The port display includes these items:

- **Green box:** A successful fabric login has occurred, and the connection is active.
- **Red X:** An SFP is present but there is no connection. This could indicate a disconnected or faulty cable or no active device connection.
- **Red box:** An SFP is present but FLOGI has failed. The issue is typically a mismatch in port or fabric parameters with the neighboring device. For example, a port parameter mismatch occurs if a node device is connected to a port that is configured as an E Port. An example of a fabric parameter mismatch is differing timeout values.
- **Yellow box:** In Cisco Device Manager, a port has been selected.
- **Gray box:** The port is administratively disabled.
- **Black box:** An SFP is not present.

In Cisco Device Manager, selecting the Summary View expands the information available for port monitoring. The display includes these items:

- Speed
- Frames that are transmitted and received
- Percent utilization for the CPU, dynamic memory, and flash memory

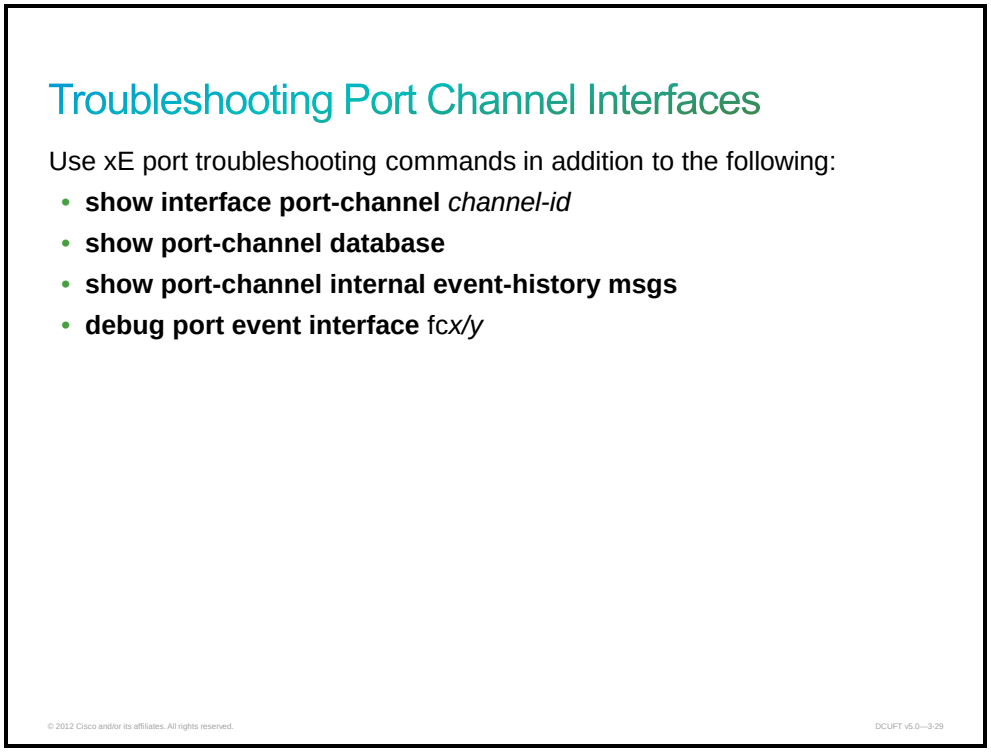
To find more detailed or additional port information with either the Device View or Summary View, choose and double-click any port. The initial display shows administrative settings for mode, speed, and status, plus current operational status, failure cause, and date of the last configuration change.

Additional tabs are available:

- **Rx BB Credit:** Configure and view buffer-to-buffer credits (BB_Credits).
- **Other:** View port channel ID, WWN, and maximum transmission unit (MTU) and configure maximum Rx buffer size.
- **FLOGI:** View FCID, pWWN, node world wide name (nWWN), BB_Credits, and class of service (CoS) for N-Port connections.
- **ELP:** View pWWN, nWWN, BB_Credits, and supported classes of service for ISLs.
- **Trunk Config:** View and configure trunk mode and allowed VSANs.
- **Trunk Failure:** View failure cause for ISLs.
- **Physical:** Configure beaconing and view SFP information.
- **Capability:** View current port capability for hold-down timers, BB_Credits, and maximum receive buffer size.
- **Diagnostics:** View diagnostics for the port.

Troubleshooting SAN Port Channel Interfaces

This topic explains how to troubleshoot issues related to SAN port channel interfaces on a Cisco Nexus or Cisco MDS Series switch.



Commands for troubleshooting E Ports can be used for troubleshooting port channels, with the addition of the following **show port-channel** commands:

- **show interface port-channel**
- **show port-channel database**
- **show port-channel internal event-history msgs**

show interface port-channel Command

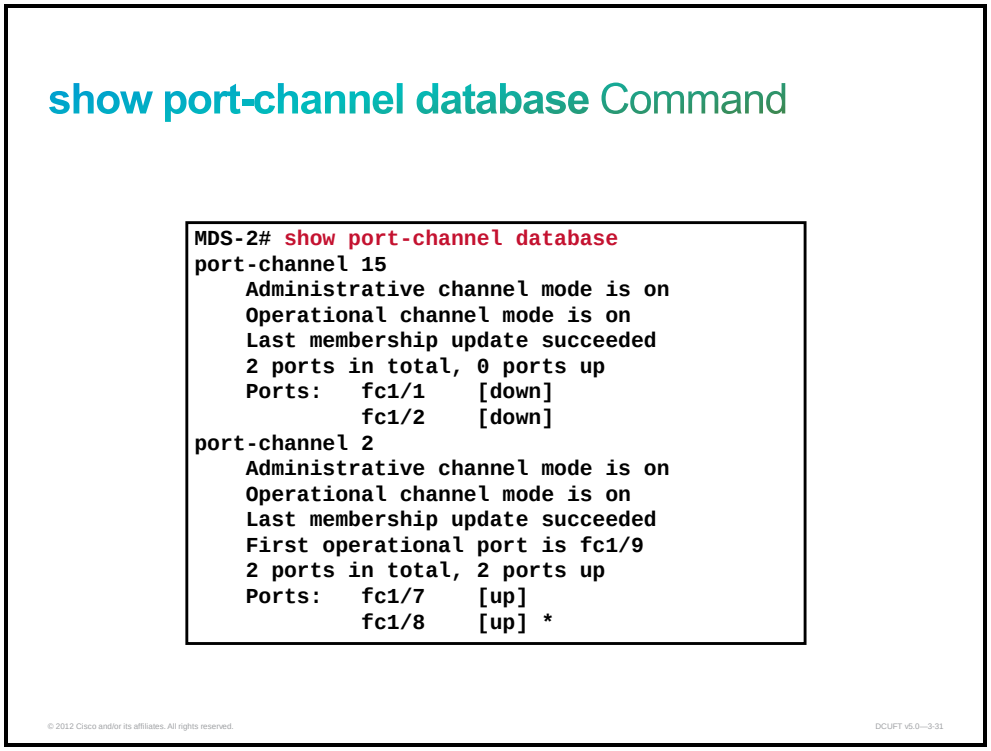
Check port channel configuration.

```
MDS-2# show interface port-channel 2
port-channel 2 is trunking
    Port description is To 10.0.15.3
    Hardware is Fibre Channel
    Port WWN is 24:73:00:0b:be:77:6f:40
    Admin port mode is E, trunk mode is on
    Port mode is TE
    Port vsan is 1
    Speed is 4 Gbps
    Trunk vsans (admin allowed and active) (1,100-200)
    Trunk vsans (up)                       (1,100-200)
    Trunk vsans (isolated)                 ()
    Trunk vsans (initializing)             ()
    5 minutes input rate 200 bits/sec, 25 bytes/sec, 0 frames/sec
    5 minutes output rate 136 bits/sec, 17 bytes/sec, 0 frames/sec
       109943 frames input, 7068672 bytes
```


To confirm port channel properties, use the **show interface port-channel** command. Use the display to verify trunk mode, VSAN status (allowed, isolated), and members.

This text is an example of the command output:

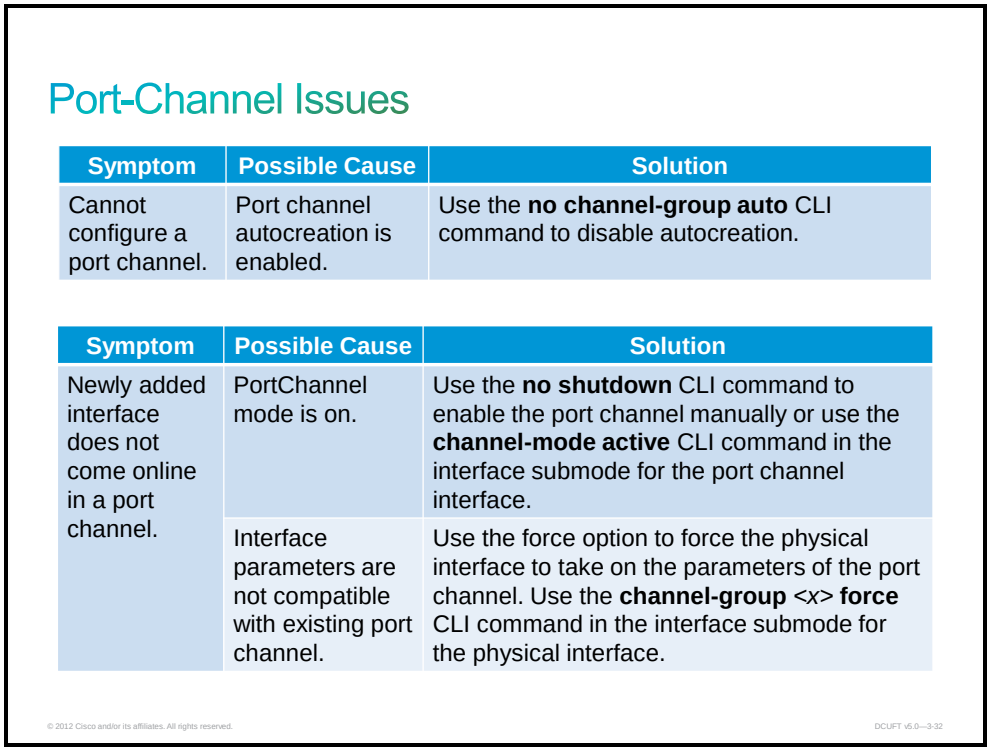
```
mds# show interface port-channel 115
port-channel 115 is trunking
Port description is To 10.0.15.3
Hardware is Fibre Channel
Port WWN is 24:73:00:0b:be:77:6f:40
Admin port mode is E, trunk mode is on
Port mode is TE
Port vsan is 1
Speed is 4 Gbps
Trunk vsans (admin allowed and active) (1,151-153)
Trunk vsans (up)                       (1,151-153)
Trunk vsans (isolated)                 ()
Trunk vsans (initializing)             ()
5 minutes input rate 200 bits/sec, 25 bytes/sec, 0 frames/sec
5 minutes output rate 136 bits/sec, 17 bytes/sec, 0 frames/sec
109943 frames input, 7068672 bytes
0 discards, 0 errors
0 CRC,  0 unknown class
0 too long, 0 too short
109945 frames output, 4884876 bytes
0 discards, 0 errors
6 input OLS, 6 LRR, 8 NOS, 0 loop inits
16 output OLS, 4 LRR, 5 NOS, 0 loop inits
Member[1] : fc1/8
Member[2] : fc1/9
```



Another command that displays port-channel configuration is the **show port-channel database** command. This command displays configuration information for all port channels:

```
mds# show port-channel database
port-channel 15
Administrative channel mode is on
Operational channel mode is on
Last membership update succeeded
2 ports in total, 0 ports up
Ports:   fc1/1    [down]
fc1/2    [down]
port-channel 115
Administrative channel mode is on
Operational channel mode is on
Last membership update succeeded
First operational port is fc1/9
2 ports in total, 2 ports up
Ports:   fc1/8    [up]
fc1/9    [up] *    <= first active port
```

Note	The asterisk (*) indicates a port that carries control plane traffic that is not load-balanced. If this port goes down, another port begins carrying control plane traffic.



The figure shows typical port-channel issues. If you cannot configure a port channel, use the **no channel-group auto** CLI command. If the newly added interface does not come online in a port channel, use the **no shutdown** CLI command to enable the port channel manually or use the **channel-mode active** CLI command in the interface submode for the port channel interface. Or, use the **channel-group** *x* **force** CLI command in the interface submode for the physical interface.

Port-Channel Issues (Cont.)

- Fibre Channel port is down when trying to connect switches via SAN port channel.
  - Possible cause
    - A SAN port-channel compatibility parameter is misconfigured in the configuration:
      - Type of interface, Gigabit Ethernet at both ends, or Fibre Channel at both ends
      - Speed, mode, rate mode, port VSAN, allowed VSAN list, and port security
      - Remote switch WWN (sWWN) and trunking mode
  - Solution
    - Use **show san-port-channel compatibility-parameters** command to verify which parameters need to be checked in the configuration.
- You cannot configure trunking.
  - Use the **trunk protocol enable** CLI command to enable trunking.

If the issue during connectivity persists with a different error message, debug further by running one or more of these commands:

- **show port internal info interface fc** *slot/port*
- **show port internal event-history interface fc** *slot/port*
- **show san-port-channel internal event-history errors**
- **show logging log | grep fc** *slot/port*
- **show san-port-channel internal event-history all**
- **show tech-support detail > bootflash:showtechdet**

Port-Channel Issues (Cont.)

- VSAN traffic does not traverse trunk.
  - Possible cause
    - VSAN is not listed in the allowed-active VSAN list.
  - Solution
    - Use the **switchport trunk allowed vsan** command to add VSAN to the allowed-active list.
- xE port is isolated in a specific VSAN under interface of SAN port channel.
  - Possible cause
    - Fabric timers or port parameters might be misconfigured or a zoning mismatch might be present.
  - Solution
    - Use the **show interface fc** <*slot/port*> command on the TE port to determine the VSAN number. The isolated VSAN number must match the VSAN number to which the host and the storage are connected.
    - Use the **show port internal info interface san-port-channel** <*number*> command to determine the cause of the VSAN isolation.

A host cannot gain access to a target that is on the same VSAN and connected to two switches using TE Ports. The VSAN traffic is unable to traverse the trunk. Depending on the path from host to target, you might observe performance degradation or you might be unable to access any disks.

One possible cause is that the VSAN is not listed in the allowed-active VSAN list. Use the **switchport trunk allowed vsan** command to add the VSAN to the allowed-active list.

The xE Port is isolated in a specific VSAN that is under a SAN port channel interface. The following error message might appear in the logging log:

```
"%$VSAN <VSAN#>%$ Interface port-channel <channel #>, vsan
<vsan #> is down (isolation due to [cause])".
```

The xE Port can be isolated in a specific VSAN for many reasons:

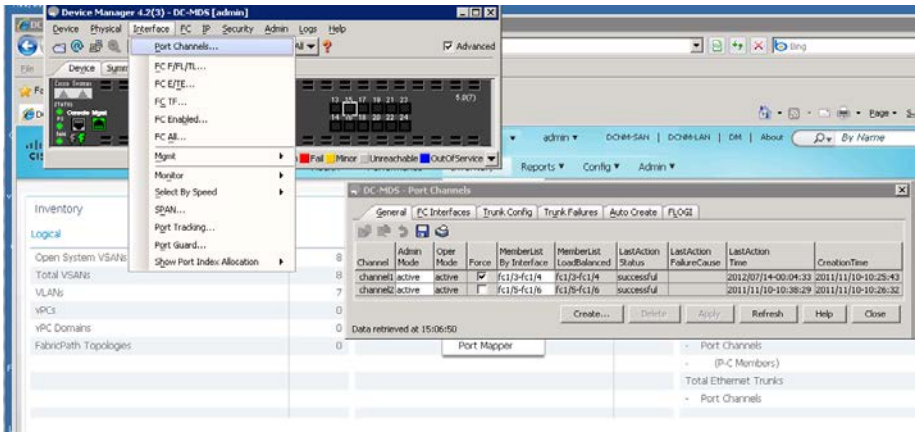
- Fabric timers might be misconfigured.
- Port parameters might be misconfigured.
- There might be a zoning mismatch.

To resolve the VSAN isolation on the TE Port, use the **show interface fc** *slot/port* command on the port, to determine the VSAN number. The isolated VSAN number must match the VSAN number to which the host and the storage are connected.

In the output of the command, look for information such as Trunk vsans (isolated) (*vsan id*).

Use the **show port internal info interface san-port-channel** *number* command to determine the cause of the VSAN isolation.

Review SAN Port Channels on Cisco DCNM



You can also review SAN port channels by using Cisco DCNM. From the Cisco DCNM general window, click **DM** and open Cisco Device Manager on a specific device. From there, click **Interface** and choose **Port Channels** to open a new window. Use this window to review the configuration and possible issues with SAN port channels.

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- The E Port and F Port troubleshooting procedures are similar. The difference involves the focus of the troubleshooting: For E Ports, you need to investigate switch-to-switch issues and for F Ports, device-to-switch issues.
- Commands for troubleshooting E Ports can be used to troubleshoot port channels, with the addition of some **show port-channel** commands.

Lesson 2

Troubleshooting Fibre Channel Fabric Services

Overview

This lesson is designed to provide some examples of common issues that relate to SAN switching when the switch is in Fibre Channel switch mode. The lesson also shows you how to identify and resolve the issues.

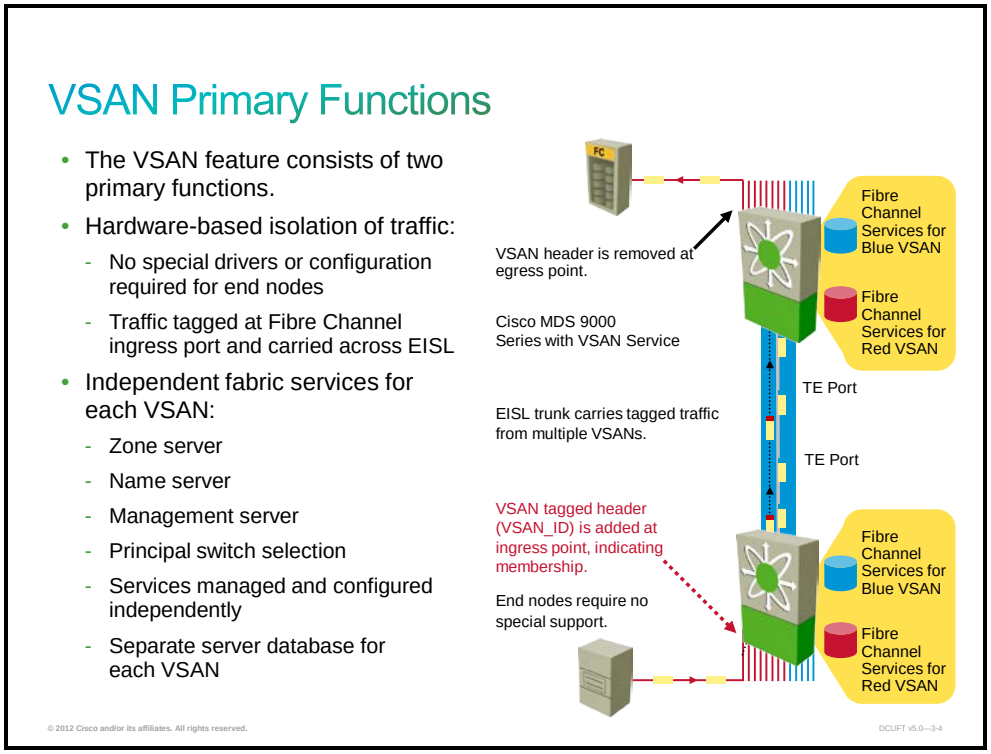
Objectives

Upon completing this lesson, you will be able to identify and resolve issues that relate to Fibre Channel switching when the Cisco Nexus Operating System (NX-OS) switch is used in switched mode (versus node port [N-Port] Virtualization [NPV] mode, which is covered in the lesson “Troubleshooting NPV Mode”).This ability includes being able to meet these objectives:

- Explain how to troubleshoot issues that relate to VSANs on a Cisco Nexus or Cisco MDS Series switch
- Explain how to troubleshoot issues that relate to the Fibre Channel Domain on a Cisco Nexus or Cisco MDS Series switch
- Explain how to troubleshoot issues that relate to the Fibre Channel Name Services on a Cisco Nexus or Cisco MDS Series switch
- Explain how to troubleshoot issues that relate to the Fibre Channel zoning on a Cisco Nexus or Cisco MDS Series switch
- Explain how to troubleshoot issues that relate to Cisco Fabric Services on a Cisco Nexus or Cisco MDS Series switch

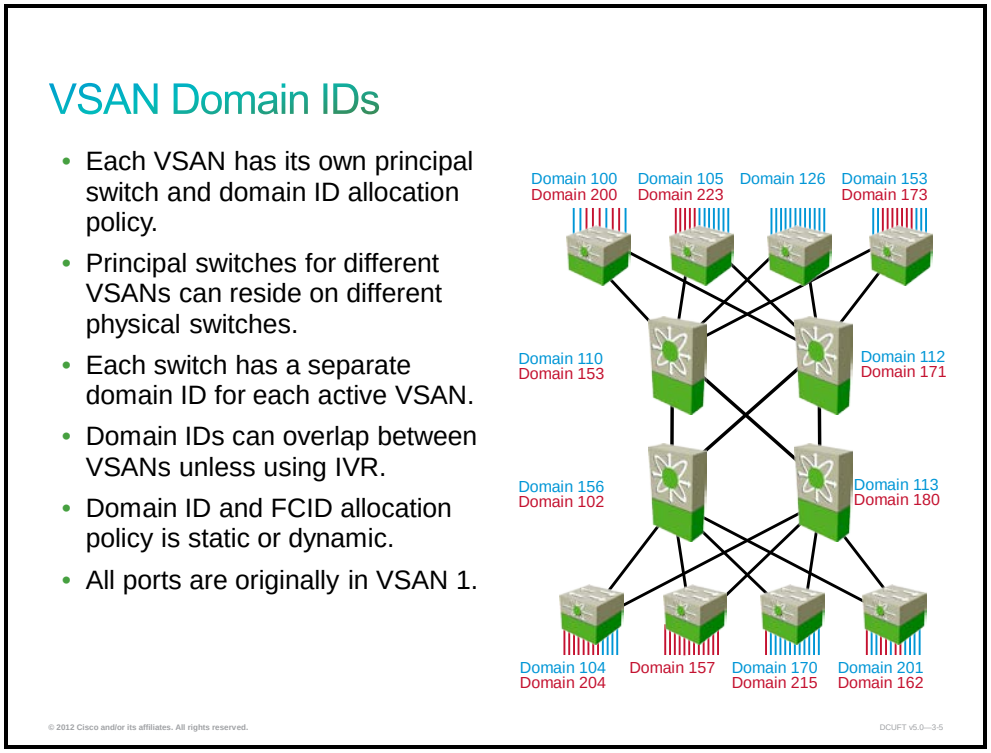
Troubleshooting VSANs

This topic explains how to troubleshoot issues that relate to virtual storage area networks (VSANs) on a Cisco Nexus or Cisco MDS Series switch.



The VSAN feature provides two primary functions: hardware-based isolation of traffic and independent Cisco Fabric Services for each VSAN:

- Hardware-based isolation of tagged traffic that belongs to different VSANs requires no special drivers or configuration at the end nodes, such as hosts, disks, and so on.
- Traffic is tagged at the Fibre Channel ingress port (Fx Port) and carried across Enhanced Inter-Switch Links (EISLs) between Cisco MDS 9000 Series Multilayer or Cisco Nexus 5500 Platform switches. Because VSANs use explicit frame tagging, they can be extended over the metropolitan-area network (MAN) or WAN.
- Fibre Channel, and therefore VSANs, can easily be carried across dark fiber. However, VSANs add 8 bytes of header, which might be a concern for channel extenders. The channel extenders might consider the header an invalid frame and drop it.
- Dense wavelength-division multiplexing (DWDM) switches might also count frames as invalid but pass the frames anyway. Qualification is ongoing within Cisco to validate various extension methods.
- Each Cisco Fibre Channel Fabric Service maintains a separate database for each newly created VSAN. These services include zone server, name server, management server, and principal switch selection. Each service runs independently on each VSAN and is independently managed and configured.



VSAN Domain IDs

Each VSAN has its own principal switch and domain ID allocation policy, either static or dynamic. Principal switches for different VSANs do not need to reside on the same physical switch. Each switch has a separate domain ID for each active VSAN.

- Each VSAN can also have a separate Fibre Channel ID (FCID) allocation policy, either static or dynamic.
- All ports are originally configured in default VSAN 1.

As shown in the figure, each switch that has end ports in a particular VSAN has a domain ID that is assigned to that VSAN. Core switches that trunk these VSANs also have assigned domain IDs in the VSANs.

VSAN Numbering

Default VSAN: The factory settings for the Cisco MDS 9000 Series and Cisco Nexus 5500 Platform switches have only the default VSAN 1 enabled. VSAN 1 should not be used in a production environment. If no VSANs are configured, then all devices in the fabric are considered part of the default VSAN. By default, all ports are assigned to the default VSAN. VSAN 1 cannot be deleted but can be suspended.

Exchange Virtual Fabric Protocol (EVFP) isolated VSAN: VSAN 4079 is used on trunking fabric port-to-trunking node port (TF-TN Port) links, to carry EVFP. The TF-TN port link allows a server to reach multiple VSANs through a TF Port without Inter-VSAN Routing (IVR). For TF port-to-trunking NP Port (TF-TNP Port) links, the Picture Transfer Protocol (PTP) is used to carry tagged frames and supports trunking port channels. The TF-TNP Port link between a third-party NPV core and a Cisco NPV switch is established by using EVFP.

Isolated VSAN: VSAN 4094 is known as the isolated VSAN. All nontrunking ports are transferred to this VSAN when the VSAN to which they belong is deleted. All ports in the deleted VSAN are isolated (disabled).

Note VSAN 0 and VSAN 4095 are reserved and not used.

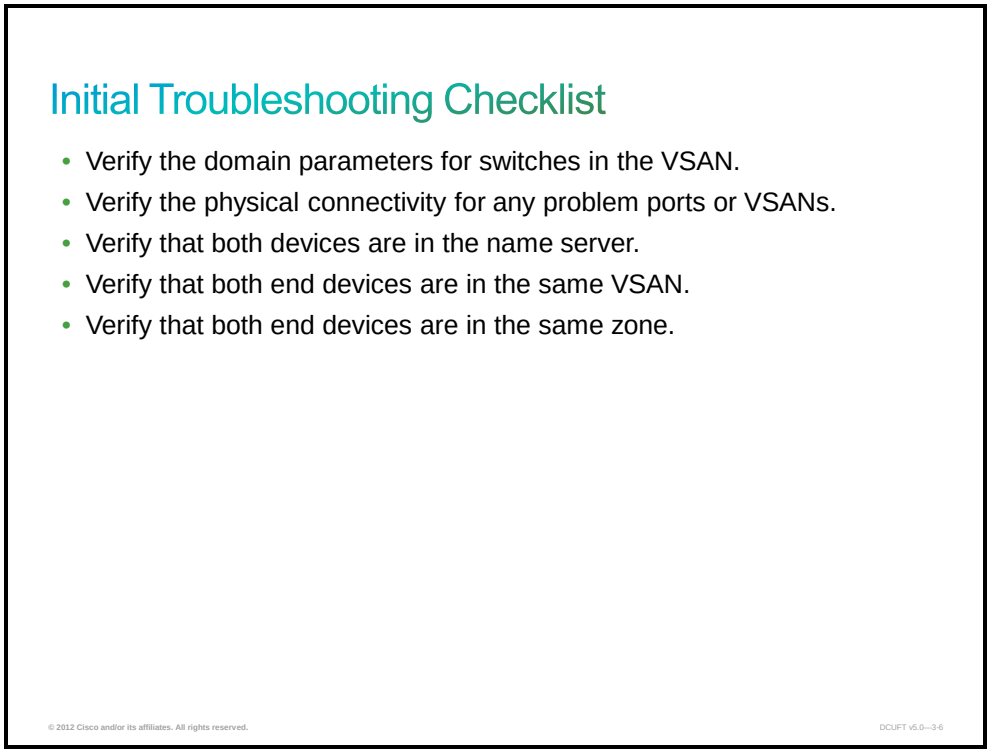
Basic VSAN Recommended Practices

VSANs provide a method of isolating devices that are physically connected to the same storage network but logically considered to be part of different SAN fabrics that do not need to be aware of one another.

VSANs provide a way to provide practical isolation of devices that are physically connected to the same fabric. VSANs reduce the size of a Fibre Channel distributed database and enable more scalable and secure fabrics.

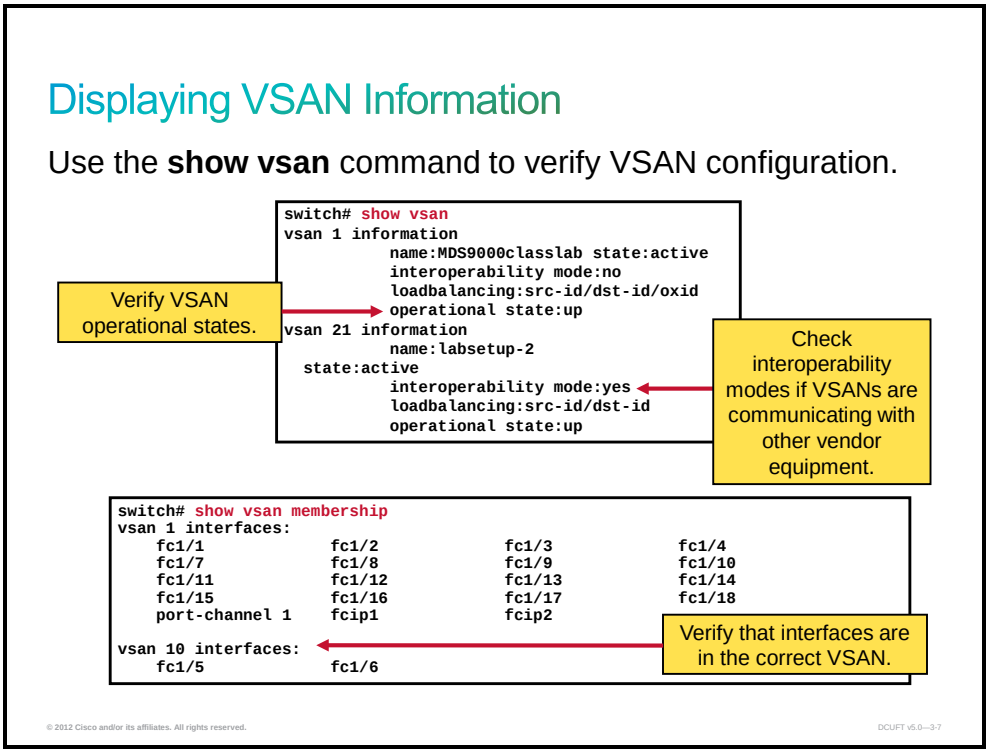
Follow these guidelines when implementing VSANs:

- Avoid using VSAN 1 (the default VSAN) for production network traffic.
- Create at least one VSAN to carry your network traffic.
- Isolate devices in VSANs whenever practical. Isolation by department and by application are two common practices. You should also isolate test environments from production environments.
- Continue to use zones inside each VSAN.



Use these CLI commands to display VSAN, Fibre Channel domain, and Fabric Shortest Path First (FSPF) information:

- **show vsan**
- **show vsan** *vsan-id*
- **show vsan membership**
- **show interface fc** *slot/port* **trunk** *vsan-id*
- **show** *vsan-id* **membership**
- **show vsan membership interface fc** *slot/port*



To verify the current VSAN configuration and status, use the **show vsan** command from executive mode. Use this command to display this information:

- Created VSANs
- VSAN name
- Administrative state (active and suspended)
- Interoperability setting (default, 1, 2, 3)
- Load-balancing scheme (source ID [SID]/destination ID [DID]/originator exchange ID [OXID]—default, SID/DID)
- Operational state (up and down)

To report the status for a specific VSAN, add the VSAN number to the **show vsan** command; for example, **switch# show vsan 20**.

Displaying VSAN information (Cont.)

Use the **show vsan membership** command to verify VSAN membership.

```
switch# show vsan membership
vsan 1 interfaces:
    fc1/1           fc1/2           fc1/3           fc1/4
    fc1/7           fc1/8           fc1/9           fc1/10
    fc1/11          fc1/12          fc1/13          fc1/14
    fc1/15          fc1/16          fc1/17          fc1/18

vsan 10 interfaces:
    fc1/5           fc1/6

vsan 20 interfaces:


vsan 4079(evfp_isolated_vsan) interfaces:


vsan 4094(isolated_vsan) interfaces:
```


To verify port VSAN membership, use the **show vsan membership** command from executive mode. The report displays port VSAN assignments, including the isolated VSAN (4094).

To report the membership for a specific VSAN, add the VSAN number to the **show vsan membership** command:

```
switch# show vsan 10 membership
vsan 10 interfaces:
    fc1/5           fc1/6
```

Allowed VSAN List Errors

Isolation because VSAN is not configured on peer:

- Use the **show interface trunk** command on both switches to determine which switch has VSAN allow-list configuration errors.

```
MDS-2# show interface fc1/9 trunk vsan
fc1/9 is trunking
    Vsan 1 is up (None), FCID is 0x1b0300
    Vsan 100 is down (Isolation due to vsan not configured on peer)
    Vsan 200 is up (None), FCID is 0x920000
```

```
MDS-3# show interface fc1/9 trunk vsan
fc1/9 is trunking
    Vsan 1 is up (None), FCID is 0xb00300
    Vsan 200 is up (None), FCID is 0xe20100
```

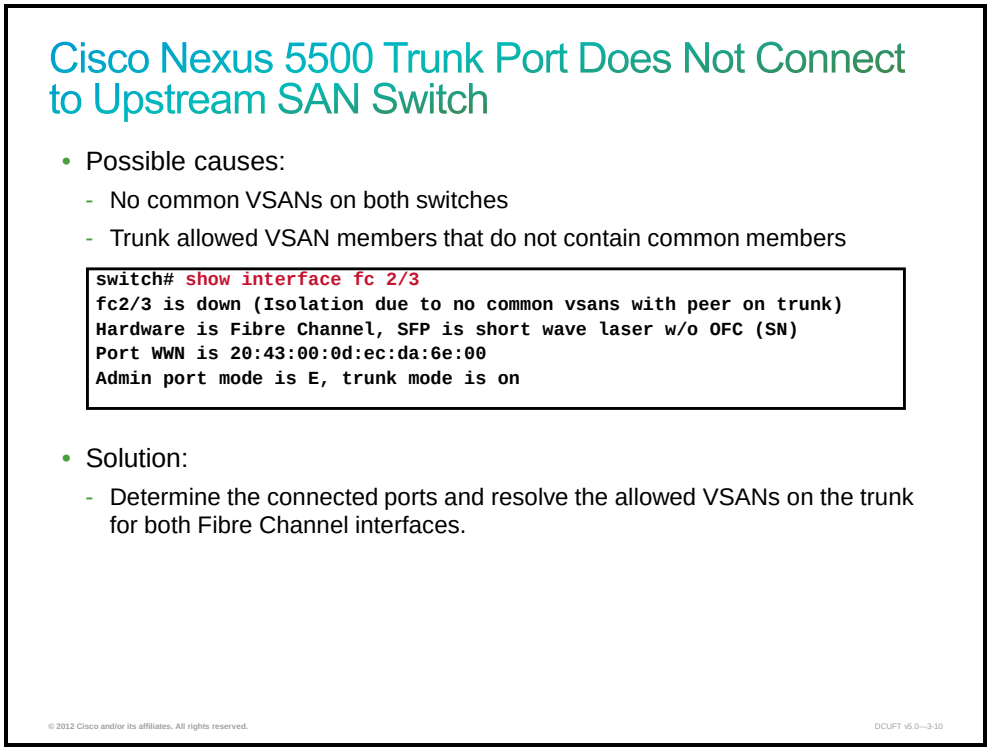
- Use the **switchport trunk allowed** command to add the VSAN to the VSAN allow list on the interface of the switch that does not display the isolation.

```
MDS-3(config)# interface fc1/9
MDS-3(config-if)# switchport trunk allowed vsan add 100
```


If you are confident that these two issues are not causing the problem, then the cause of the isolation is probably a misconfigured **switchport trunk allowed** entry on an Inter-Switch Link (ISL) interfaces:

- The VSAN that is experiencing isolation is configured on both switches in adjacency.
- The nonoperational status of isolation because a VSAN is not configured on the peer is displayed in the output of the **show interface trunk vsan** command on one interface and not on other trunk ports.

This can be verified by running the **show interface fc** *x/y* **trunk vsan** command on both switches. The switch interface that does not display the isolation should be updated to include the VSAN by using the **switchport trunk allowed** command.



The Cisco Nexus 5500 Platform trunk port does not connect to the upstream SAN switch for these reasons:

- The status of the trunk port connected to the upstream switch is isolated.
- The switch port trunk mode is enabled on both sides.
- Physical cabling has been checked and verified.
- Ports are up on both switches.

The VSAN allow list is not the same for both interfaces. Specifically, no common VSAN is allowed on both interfaces. This situation might have several causes:

- There are no common VSANs on both switches.
- The trunk allowed VSAN members that do not contain common members.

In the example in the figure, the trunk VSAN allow list on the Cisco Nexus 5500 Platform and Cisco MDS Series Fibre Channel interfaces do not match.

To solve the issue, determine the connected ports and resolve the allowed VSANs on the trunk for both Fibre Channel interfaces.

Cisco Nexus 5500 E Port (Nontrunking) Does Not Connect to Upstream SAN Switch

- Possible causes:
  - The error is displayed by the **show interface brief** and **show vsan membership** commands, which show that the E Port on one switch is in the wrong VSAN.
- Solution:
  - Move the nontrunking E Port into the proper VSAN.

```
Switch1 # show interface brief
--------------------------------------------------------------------------------
Interface         Vsan     Admin Admin Status SFP Oper        Oper     Port
                           Mode  Trunk               Mode     Speed    Channel
                                 Mode                        (Gbps)
--------------------------------------------------------------------------------
fc2/4             50       E       off   up   swl   E       2        --

Switch2 # show interface brief
--------------------------------------------------------------------------------
Interface         Vsan     Admin Admin Status SFP Oper        Oper     Port
                           Mode  Trunk               Mode     Speed    Channel
                                 Mode                        (Gbps)
--------------------------------------------------------------------------------
fc1/2             100      E       off   up   swl   E       2        --
```


The Cisco Nexus 5500 Platform expansion port (E Port) does not connect to the upstream SAN switch because one of these issues has occurred:

- The status of the interconnected nontrunking E Ports shows that the status is up. However, no Fibre Channel services are working between the switches.
- Devices in the same VSAN do not appear in the Fibre Channel Name Server (FCNS) database for both switches.
- The show topology command does not list peer switch information.
- Zones show that members are not logged in.

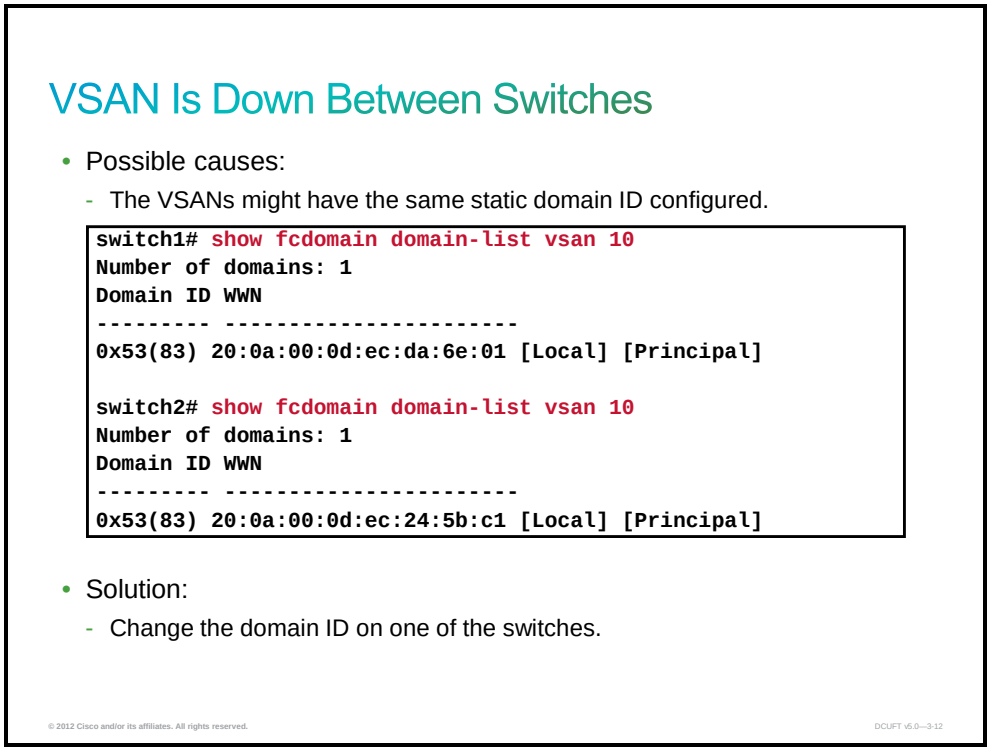
The error is displayed by the **show interface brief** and the **show vsan membership** commands. These commands show that the E Port on one switch is in the wrong VSAN.

To solve the issue, move the nontrunking E Port into the correct VSAN; for example:

```
switch(config-vsan-db)# vsan 100 interface fc 2/4
Traffic on fc2/4 may be impacted. Do you want to continue? (y/n) [n] y
```

The zone set is now active and the Fibre Channel topology is correct. Example:

```
switch(config-if)# show zoneset active vsan 100
zoneset name ZoneSet_Host_Storage vsan 100
zone name Zone_Host_Storage vsan 100
* fcid 0x640114 [pwwn 20:00:00:25:b5:00:20:0e] [Host]
* fcid 0x5a0000 [pwwn 50:0a:09:81:86:78:39:66] [Storage]
switch(config-if)# show topology
FC Topology for VSAN 100 :
--------------------------------------------------------------------------
Interface     Peer Domain    Peer Interface       Peer IP Address
--------------------------------------------------------------------------
fc1/2         0x5a(90)         fc2/4              172.25.183.124
```



The VSAN is down between switches because one of these issues has occurred:

- The VSAN is configured on both switches.
- The trunk allow list allows the VSAN.
- The VSAN is reported to be down (initializing state).
- The zones are active.
- Both the host and storage are logged into the SAN.

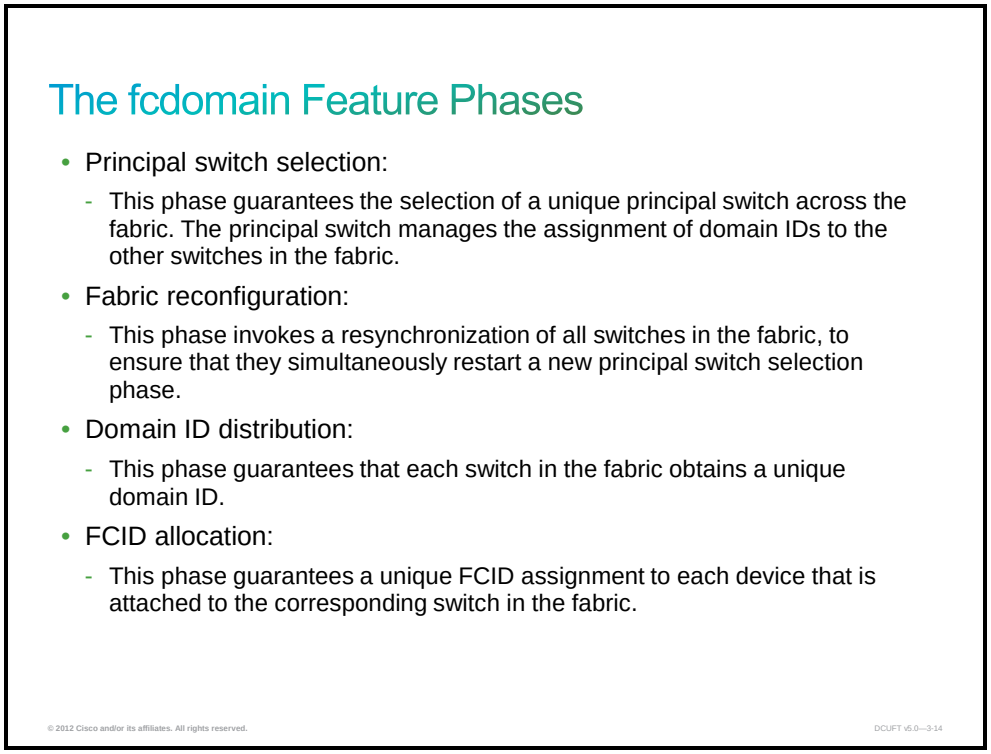
In this failure, the storage port is not logged into the active zone set. After examining the interface, the error can be seen as in this example:

```
switch# show interface fc 2/4 trunk vsan 10
fc2/4 is trunking
Vsan 10 is down (Isolation due to domain id assignment failure)
switch# show port internal info interface fc 2/4 | grep Isolation
fc2/4, Vsan 10 - state(down), state reason(Isolation due to domain id
assignment failure), fcid(0x000000)
fc2/4, Vsan 50 - state(down), state reason(Isolation due to vsan not
configured on peer), fcid(0x000000)
```

The VSANs might have the same static domain ID configured. To solve the issue, change the domain ID on one of the VSANs.

Troubleshooting Fibre Channel Domain

This topic explains how to troubleshoot issues that relate to the Fibre Channel domain on a Cisco Nexus or Cisco MDS Series switch.



The Fibre Channel domain (fcdomain) feature performs principal switch selection, domain ID distribution, FCID allocation, and fabric reconfiguration as described in the Fibre Channel Switch Fabric 2 (FC-SW-2) standards. The domains are configured on a per-VSAN basis, and if you do not configure a domain ID, the local switches use a random ID.

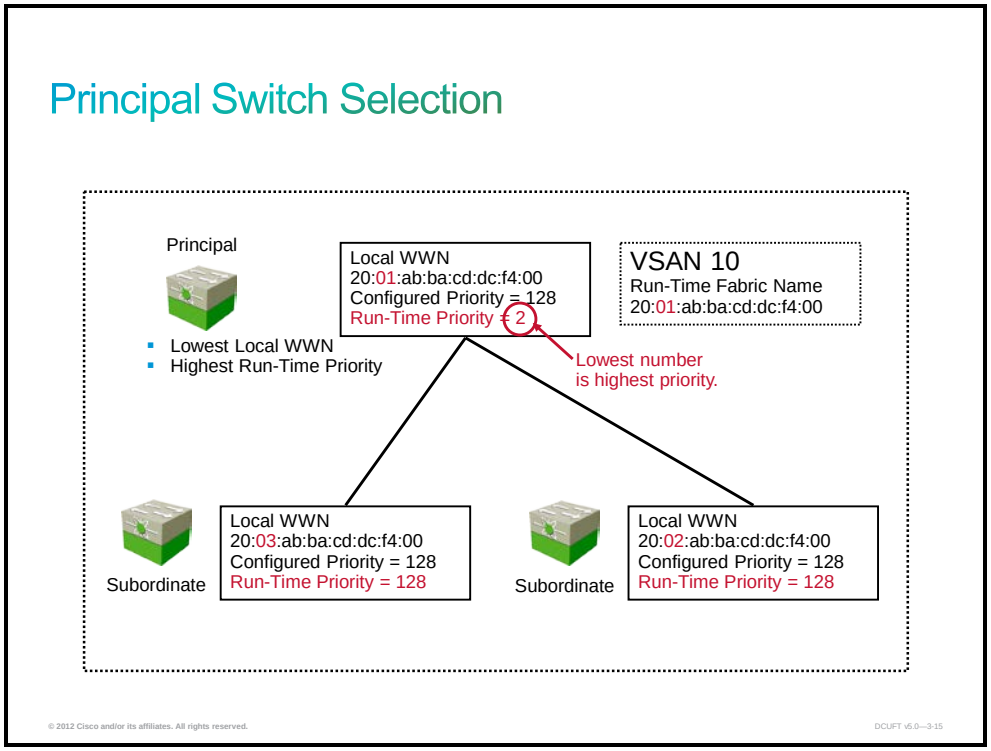
To successfully configure domain parameters and prevent fabric segmentation, you must understand the anticipated behavior of the four fcdomain feature phases:

- **Principal switch selection:** This phase guarantees the selection of a unique principal switch across the fabric.
- **Domain ID distribution:** This phase guarantees that each switch in the fabric obtains a unique domain ID.
- **FCID allocation:** This phase guarantees a unique FCID assignment to each device that is attached to the corresponding switch in the fabric.
- **Fabric reconfiguration:** This phase guarantees a resynchronization of all switches in the fabric, to ensure that they simultaneously restart a new principal switch selection phase.

Domain ID Guidelines

Domain IDs must be unique across interconnected VSANs. To ensure unique domain IDs across interconnected VSANs, consider these guidelines:

- Minimize the number of switches that require a domain ID assignment, to ensure minimum traffic disruption.
- Minimize the coordination between interconnected VSANs when configuring the SAN for the first time, as well as when you add each new switch.

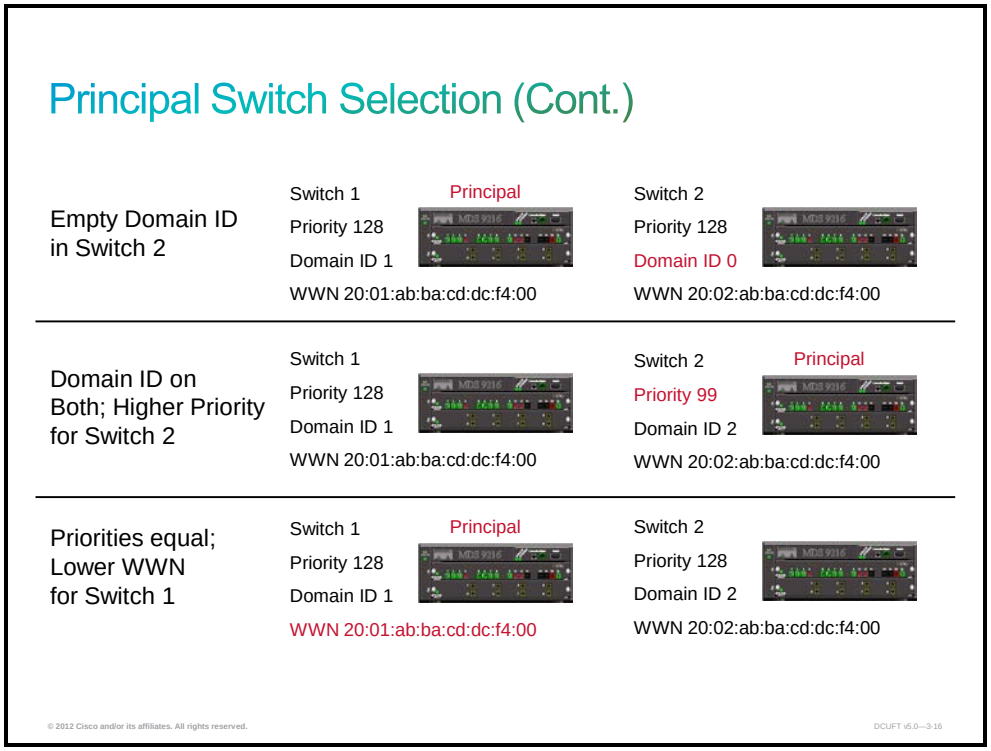


The principal switch selection phase guarantees the selection of a unique principal switch across the fabric. The principal switch allocates domain IDs to subordinate switches.

When the fcdomain feature is disabled, the run-time fabric name is the same as the configured fabric name.

When the fcdomain feature is enabled, the run-time fabric name is the same as the world wide name (WWN) of the principal switch.

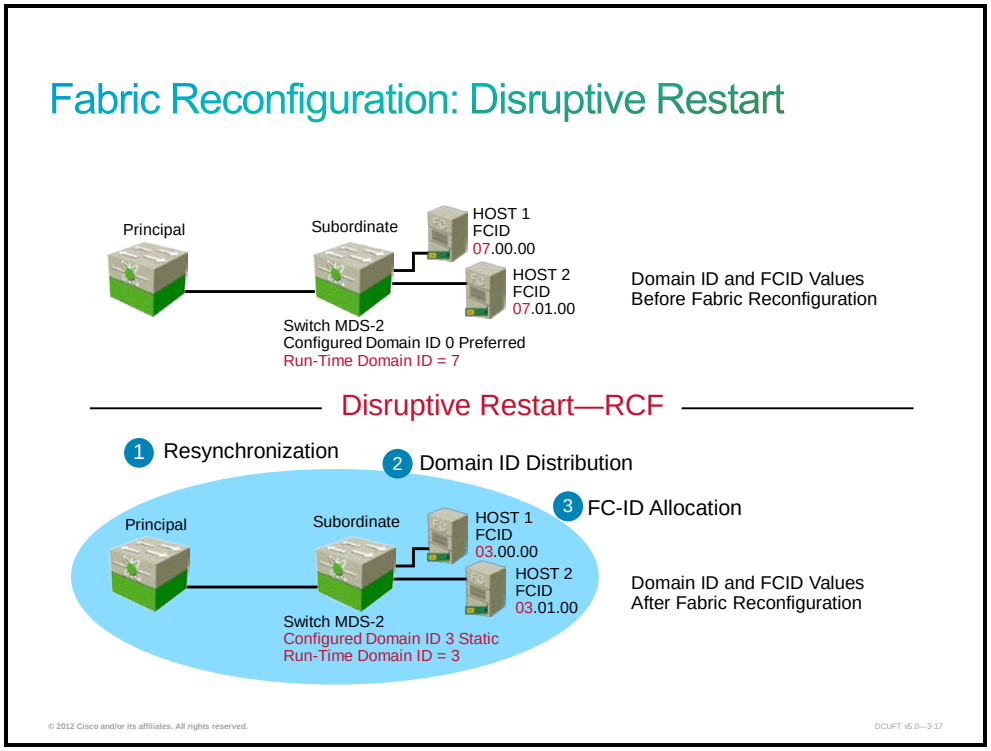
In this example, the configured fabric name is 20:01:ab:ba:cd:dc:f4:00.



The principal switch is used to issue domain IDs when a new switch is added to an existing fabric. When two fabrics merge, the principal switch selection process determines which one of the existing switches becomes that principal switch.

The election of the new principal switch is characterized by these rules:

- A switch with a nonempty domain ID list has priority over a switch with an empty domain ID list, and the principal switch is the principal switch of the first fabric. For a single-switch fabric, the switch does not contain a domain ID list.
- If both fabrics have a domain ID list, the priority between the two principal switches is determined by configured switch priority. The user can set this parameter; the lower the value, the higher the priority. However, when you are connecting a single-switch fabric to a multiswitch fabric, the multiswitch fabric always retains its principal switch, regardless of the principal switch priority setting on the single-switch fabric.
- If the principal switch cannot be determined by either of the two previous criteria, then the WWNs of the two switches determine the principal switch: The lower value has the higher priority.



Domain Restart

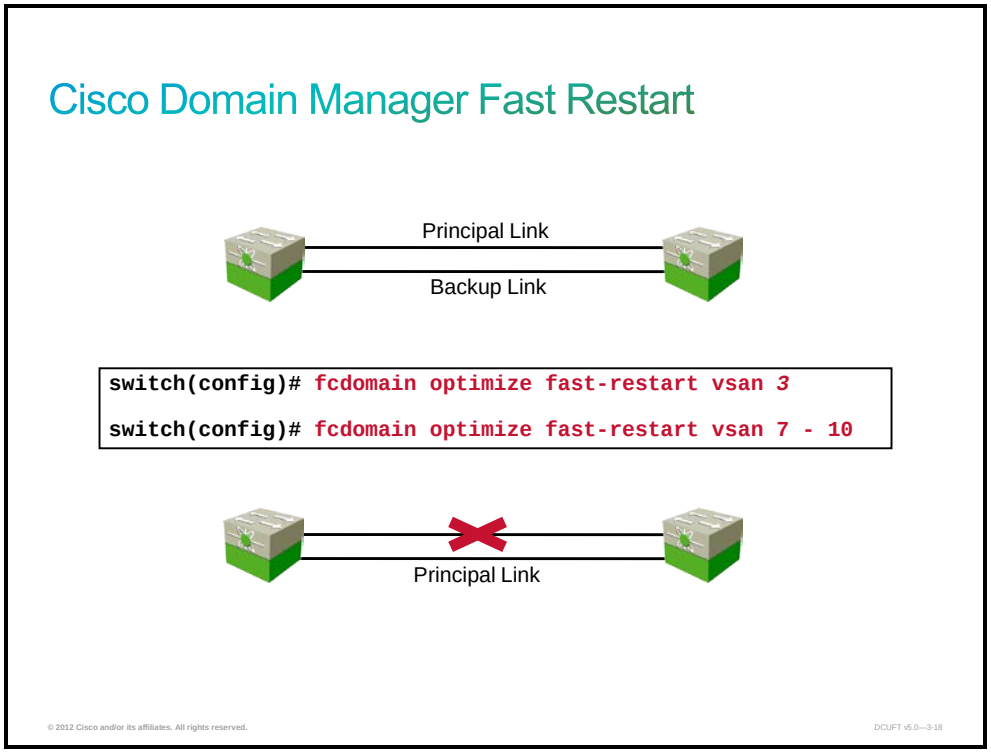
Fibre Channel domains can be started disruptively or nondisruptively.

- If you perform a disruptive restart, Reconfigure Fabric (RCF) frames are sent to other switches in the fabric.
- If you perform a nondisruptive restart, Build Fabric (BF) frames are sent to other switches in the fabric.

Note A static domain is specifically configured by the user and can be different from the run-time domain. If the domain IDs are different, the run-time domain ID changes to take on the static domain ID after the next restart.
If a VSAN is in interoperability mode, you cannot disruptively restart the fcdomain feature for that VSAN.

You can apply most of the configurations to their corresponding run-time values. Each of the following sections provides further details on how the fcdomain parameters are applied to the run-time values.

The **fcdomain restart** command applies your changes to the run-time settings. Use the restart disruptive option to apply most of the configurations to their corresponding run-time values.



When a principal link fails, Cisco Domain Manager must select a new principal link. By default, Cisco Domain Manager starts a BF phase, followed by a principal switch selection phase. Both phases involve all switches in the VSAN and together take at least 15 seconds to complete. To reduce the time that is required for Cisco Domain Manager to select a new principal link, you can enable the Cisco Domain Manager fast restart feature.

When fast restart is enabled and a backup link is available, Cisco Domain Manager needs only a few milliseconds to select a new principal link to replace the one that failed. Also, the reconfiguration that is required to select a new principal link affects only the two switches that are directly attached to the failed link, not the entire VSAN. When a backup link is not available, Cisco Domain Manager reverts to the default behavior and starts a BF phase, followed by a principal switch selection phase. The fast restart feature can be used in any interoperability mode.

show fcdomain

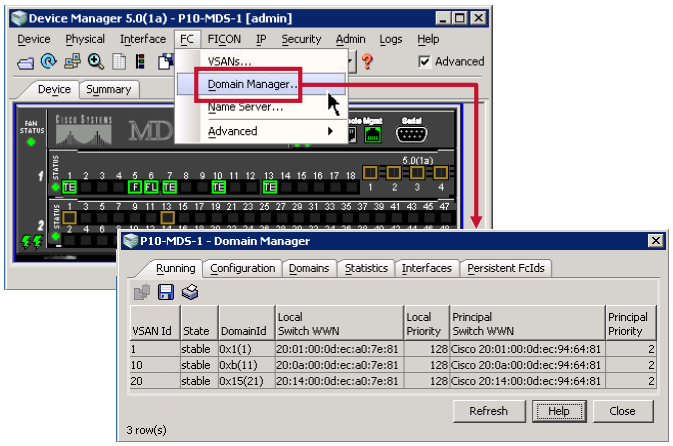
```
switch# show fcdomain
VSAN 1
The local switch is a Subordinated Switch.
Local switch run time information:
        State: Stable
        Local switch WWN:    20:01:00:05:30:00:13:9f
        Running fabric name: 20:01:00:05:30:00:13:9e
        Running priority: 128
        Current domain ID: 0x4a(74)
Local switch configuration information:
        State: Enabled
        FCID persistence: Enabled
        Auto-reconfiguration: Disabled
        Contiguous-allocation: Disabled
        Configured fabric name: 20:01:00:05:30:00:28:df
        Configured priority: 128
        Configured domain ID: 0x00(0) (preferred)
Principal switch run time information:
        Running priority: 2
```

```
switch# show fcdomain vsan 23
Error: VSAN 23 not active or WWN not available.
```


Issue the **show fcdomain** command with no arguments to display all VSANs. The VSANs should be active or an error is generated; for example:

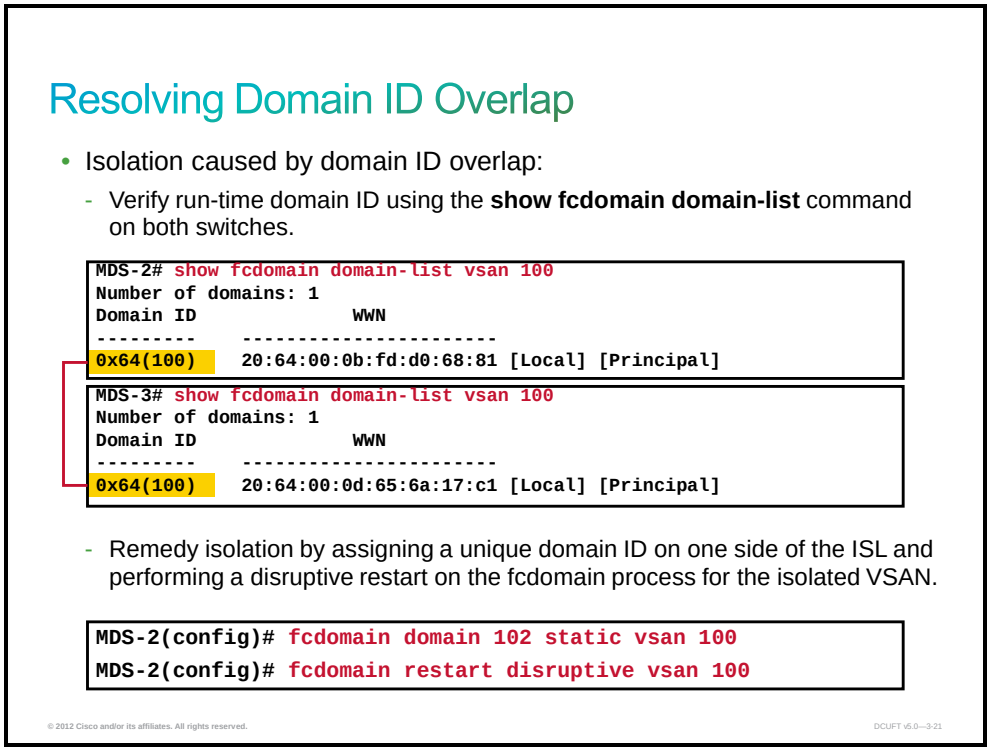
```
DC-MDS# show fcdomain vsan 23
Error: VSAN 23 not active or WWN not available.
```

Cisco Device Manager Domain Database



Displaying the Cisco Domain Database with Cisco Device Manager

To view fcdomain information, in Cisco Device Manager choose FC > Domain Manager.



If two switch fabrics with two or more switches are connected and both fabrics have switches with the domain ID already assigned, then the E Ports that are used to connect the two fabrics are isolated in that VSAN.

The auto-reconfigure option prevents isolation by forcing a disruptive restart of the VSAN. However, a disruptive restart is still required if autoreconfiguration is enabled after isolation occurs. The auto-reconfigure option must be enabled on all switches and is disabled by default.

To verify that each switch can see the other switches, use the **show fcdomain domain-list vsan** *vsan* command. If the command does not include a specific VSAN number, then the display lists the output for all VSANs.

The command output lists the set of domain IDs and associated WWNs for each switch within a VSAN. This list provides the WWN of the switches that own each domain ID and information about whether a switch is the principal switch of the switches in the fabric or VSAN to which it belongs.

Example: Two Switches in VSAN 1

This example shows two switches in VSAN 1. This output indicates that the switch on which the command was issued has built its adjacency in VSAN 1, with the other switch in the same VSAN:

```
switch1# sh fcdomain domain-list vsan 1
Number of domains: 2
Domain ID WWN
--------- --------------------------------
0x4a(74) 20:01:00:05:30:00:13:9f [Local]
0x4b(75) 20:01:00:05:30:00:13:9e [Principal]
------- ------------------------------------
```


Example: One Switch in VSAN 1

In this example, only one switch is recognized. This output indicates that the switch on which the command was issued has not established adjacency with the neighboring switch in VSAN 1:

```
switch2# sh fcdomain domain-list vsan 1
Number of domains: 1
Domain ID WWN
--------- -----------------------
0x4a(74) 20:01:00:05:30:00:13:9f [Local] [Principal]
--------- -----------------------
```

Resolving Domain ID Overlaps

To resolve a domain ID overlap, you can manually assign the domain ID by using one of these commands in configuration mode:

- **switch(config)# fcdomain domain** *domain-id* **static vsan** *x*
- **switch(config)# fcdomain domain** *domain-id* **preferred vsan** *x*

The static option tells the switch to request that particular domain ID. If the switch does not get that particular address, it isolates itself from the fabric. With the preferred option, the switch requests the specified domain ID. If that domain ID is unavailable, the switch accepts another domain ID.

After configuring the domain ID, you must restart Cisco Domain Manager.

Note	Although the **static** option can be applied to run time after a disruptive or nondisruptive restart, the **preferred** option is applied to run time only after a disruptive restart.

Verify Automatic Reconfiguration Status

When joining two stable fabrics with overlapping domain assignments, these cases apply:

- If the auto-reconfigure option is disabled (default) on either or both switches, the links between the two switches become isolated.
- Beware: If the auto-reconfigure option is enabled on both switches, a disruptive reconfiguration phase is started.

```
switch# show fcdomain
VSAN 1
 . . .
Local switch configuration information:
    State: Enabled
    FCID persistence: Enabled
    Auto-reconfiguration: Disabled
```


By default, the **auto-reconfigure** option is disabled. When you join two switches that belong to two stable fabrics with overlapping domains, these cases apply:

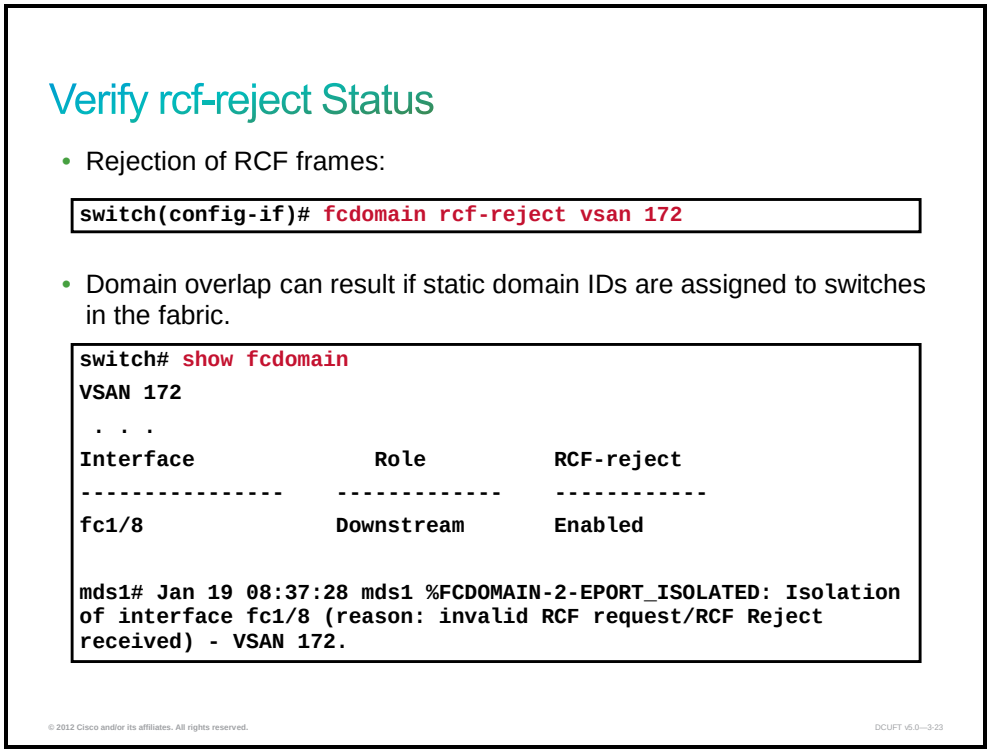
- **Case 1:** If the **auto-reconfigure** option is enabled on both switches, a disruptive reconfiguration phase is started.
- **Case 2:** If the **auto-reconfigure** option is disabled on either or both switches, the links between the two switches become isolated.

To display whether the **auto-reconfigure** option is enabled, use the **show fcdomain** command. The display shows the setting on the local switch for each configured VSAN.

To enable the **auto-reconfigure** option on a particular VSAN, use the **fcdomain auto-reconfigure vsan** *vsan* command in configuration mode.

The **auto-reconfigure** option takes immediate effect at run time. You do not need to reissue the **fcdomain** command.

If a domain is isolated because of domain overlap, and you later enable the **auto-reconfigure** option on both switches, the fabric continues to be isolated. However, if you enable the option on both switches before connecting the fabric, a disruptive RCF occurs. A disruptive reconfiguration can affect data traffic. You can nondisruptively reconfigure the Fibre Channel domain by manually changing the configured domains on the overlapping links and eliminating the domain overlap.



The administration of domain IDs typically involves restarting the fabric or VSAN, which guarantees a resynchronization of all switches in the fabric to ensure that they simultaneously restart a new principal switch selection phase. The principal switch guarantees that each switch in the fabric obtains a unique domain ID. Fibre Channel domains can be started disruptively or nondisruptively. With each disruptive restart, RCF frames are sent to other switches in the fabric.

Cisco MDS 9000 Series and Cisco Nexus 5500 Platform switches can be configured to reject RCF frames. To determine whether your switch is configured to reject or accept RCFs, use the **show fcdomain** command. The output shows the RCF rejection status for each ISL interface. This figure shows that interface fc1/8 has RCF rejection that is enabled for VSAN 172. If an RCF for VSAN 172 is received from the fabric, then this logging message is echoed:

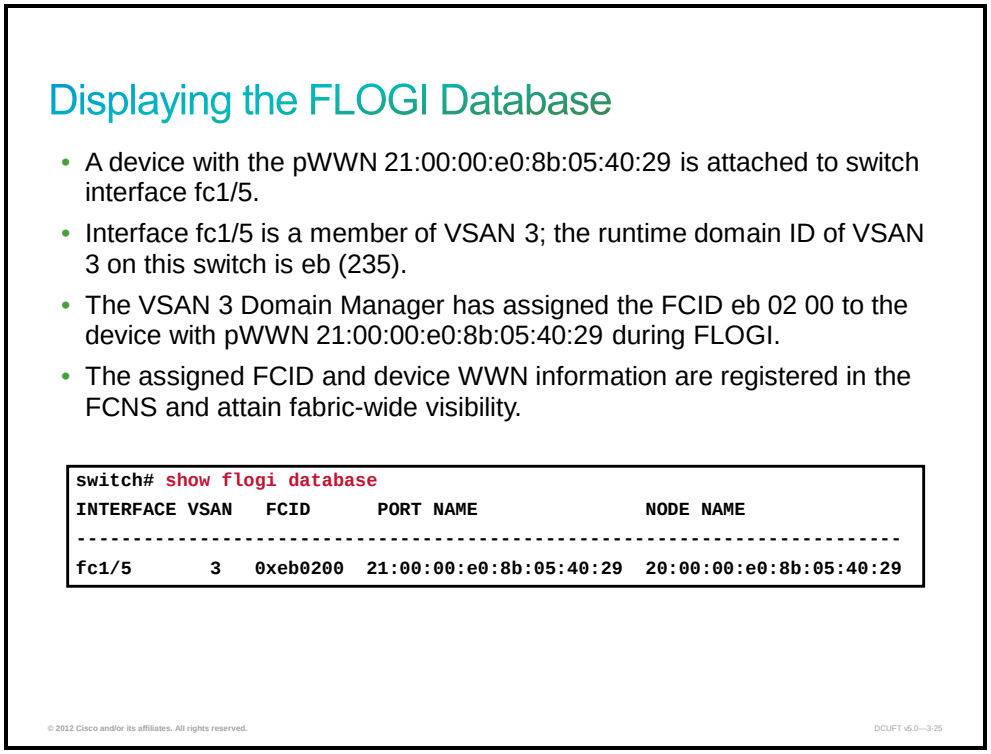
Jan 19 08:37:28 mds1 %FCDOMAIN-2-EPORT_ISOLATED: Isolation of interface fc1/8 (reason: invalid RCF request/RCF Reject received) - VSAN 172.

To verify the trunking status of the interface, use the show interface fc *x*/*y* trunk vsan command:

```
switch# show interface fc 1/8 trunk vsan
fc1/8 is trunking
Vsan 1 is up, FCID is 0x640200
Vsan 171 is up, FCID is 0x640000
Vsan 172 is down (Isolation due to invalid fabric
reconfiguration)
Vsan 173 is up, FCID is 0x620100
```

Troubleshooting Fibre Channel Name Services

This topic explains how to troubleshoot issues that relate to the Fibre Channel Name Services on a Cisco Nexus or Cisco MDS Series switch.



In a Fibre Channel fabric, each host or disk requires an FCID. Use the **show flogi** command to verify that a storage device is displayed in the FLOGI table, as in the figure.

If the required device is displayed in the FLOGI table, the fabric login is successful. Examine the FLOGI database on a switch that is directly connected to the host bus adapter (HBA) of the host and to connected ports.

Displaying FCNS Entries

- Use the **show fcns database** command to display the FCNS and statistical information for a specific VSAN or for all VSANs.
- Use the **show fcns detail** command to display extended information.

```
switch# show fcns database vsan 3
VSAN 3:
--------------------------------------------------------------------------------
FCID      TYPE     PWWN                    (VENDOR)  FC4-TYPE:FEATURE
--------------------------------------------------------------------------------
0xeb00e2    NL    21:00:00:04:cf:d6:f3:ac       (Seagate)       scsi-fcp
0xeb00e4    NL    21:00:00:0c:50:9e:8b:d8       (Seagate)       scsi-fcp
0xeb00e8    NL    21:00:00:0c:50:9e:8b:78       (Seagate)       scsi-fcp
0xeb00ef    NL    21:00:00:0c:50:9e:8b:76       (Seagate)       scsi-fcp
0xeb0100    N     21:01:00:e0:8b:25:45:29       (Qlogic)        scsi-fcp:init
0xeb0200    N     21:00:00:e0:8b:05:40:29       (Qlogic)        scsi-fcp:init
0xec00e2    NL    22:00:00:04:cf:d6:f3:ac       (Seagate)       scsi-fcp
0xec00e4    NL    22:00:00:0c:50:9e:8b:d8       (Seagate)       scsi-fcp
<…>
```

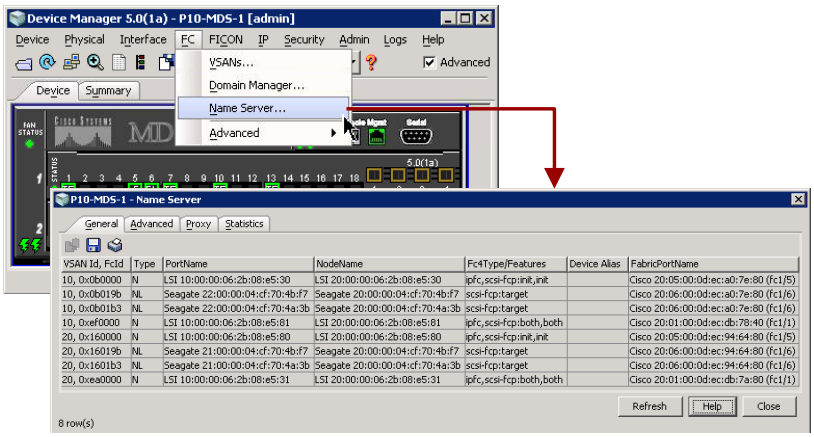

The FCNS stores name entries for all hosts in the FCNS database. The name server permits an Nx Port to register attributes during a FLOGI (to the name server) to obtain attributes of other hosts. These attributes are deregistered when the Nx Port logs out either explicitly or implicitly.

In a multiswitch fabric configuration, the name server instances that run on each switch share information in a distributed database. One instance of the name server process runs on each switch, with separate databases for each VSAN.

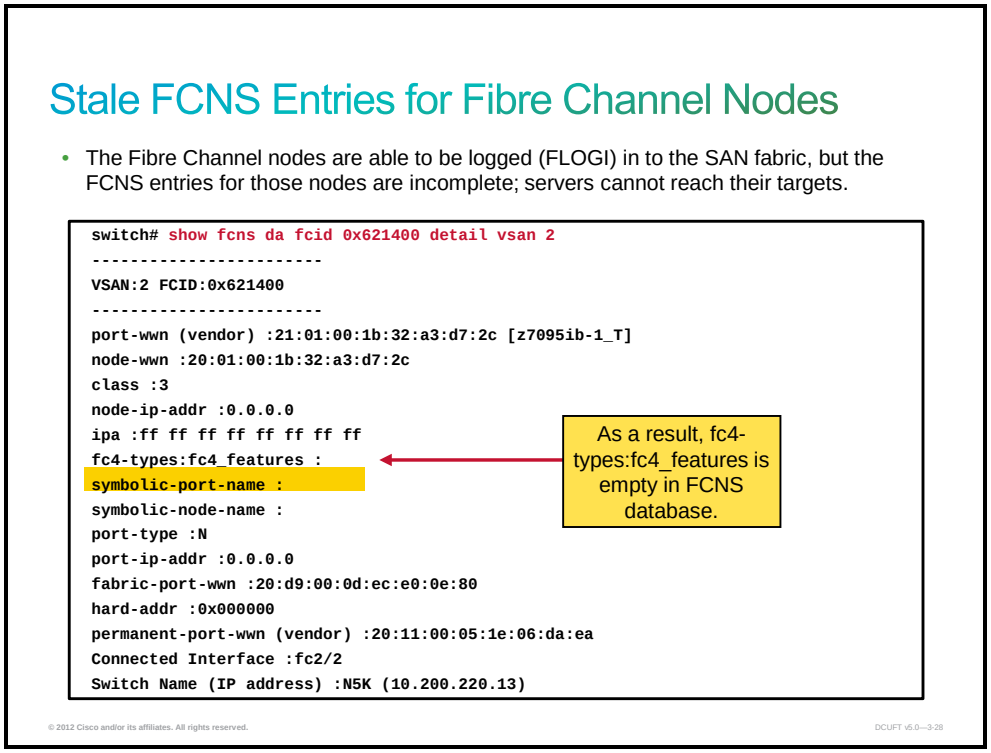
Use the **show fcns** command to display the name server database and statistical information for a specific VSAN or for all VSANs.

Displaying FCNS Entries in Cisco Device Manager

Choose **FC > Name Server** to view the FCNS database.



The menu bar at the top of the Cisco Device Manager main window provides options for managing and troubleshooting a single switch.



The Fibre Channel nodes are able to be logged (FLOGI) in to the SAN fabric, but the FCNS entries for those nodes are incomplete. Servers cannot reach their targets. As a result, fc4-types:fc4_features will be empty in the FCNS database.

Possible Cause

The Fibre Channel nodes might not be registering their FC4Types and FC4Features in the FCNS database in a topology in which Cisco Nexus 5000 Series Switches are configured as NPV core (feature N-Port ID Virtualization [NPIV]) and connected to existing gateway switches. The fc4-types:fc4_features can be verified by using the **show fcns database detail** command, as shown in this example:

```
switch# show fcns da fcid 0x621400 detail vsan 2
------------------------
VSAN:2 FCID:0x621400
------------------------
port-wwn (vendor) :21:01:00:1b:32:a3:d7:2c [z7095ib-1_T]
node-wwn :20:01:00:1b:32:a3:d7:2c
class :3
node-ip-addr :0.0.0.0
ipa :ff ff ff ff ff ff ff ff
fc4-types:fc4_features :
symbolic-port-name :
symbolic-node-name :
port-type :N
port-ip-addr :0.0.0.0
fabric-port-wwn :20:d9:00:0d:ec:e0:0e:80
hard-addr :0x000000
permanent-port-wwn (vendor) :20:11:00:05:1e:06:da:ea
Connected Interface :fc2/2
Switch Name (IP address) :N5K (10.200.220.13)
```

Some existing gateway switches might require the area part of the FCID to be the same for the switch and for all the blades that are logged in through that port.

However, because of an old issue with QLogic HBAs, the Cisco Nexus 5000 Series domain server assigns a separate area for each QLogic HBA that matches a certain Organizationally Unique Identifier (OUI) by default. Therefore, a conflict between existing gateway requirements and the Cisco domain allocation scheme exists. Cisco still implements this setup to support old existing QLogic HBAs in the field.

Solution

Us the **no fcid-allocation area company** *oui* command for all used QLogic OUIs (ensuring flat FCID allocation in the future), force all affected blades to log out of the fabric, delete the already created persistent FCID entry from the Cisco Nexus 5000 Series switch configuration, and allow the blade to log in again.

In the following **show flogi database** command output, all devices obtain a unique area ID (x01, x08, x0c):

```
Fc2/1 2 0x620104 20:10:00:05:1e:5e:6a:85 10:00:00:05:1e:5e:6a:85
Fc2/1 2 0x620800 21:01:00:1b:32:a3:c0:2e 20:01:00:1b:32:a3:c0:2e
Fc2/1 2 0x620c00 21:01:00:1b:32:33:8b:8e 20:01:00:1b:32:33:8b:8e
```

Because of the specific area ID requirement of the existing switch, the last two blades must also have area x01. To force the QLogic adapters to log in again and obtain FCIDs in 0x6201xx range, follow these steps:

Step 1 Configure (force) the future FCID allocation scheme to be flat for all WWNs that match the OUIs in this situation:

```
switch(configure)# no fcid-allocation area company 0x001B32
```

Step 2 Force the FCID under reconfiguration to log out of the fabric.

Note If you shut down the Cisco Nexus 5000 Series interface that serves as the primary uplink for that server, the server will log in through another interface. The appropriate method is to shut down the affected blade and ensure that the FLOGI for the WWN is gone.

Step 3 Delete the automatically created configuration entry for persistent FCID allocation, as shown in this example:

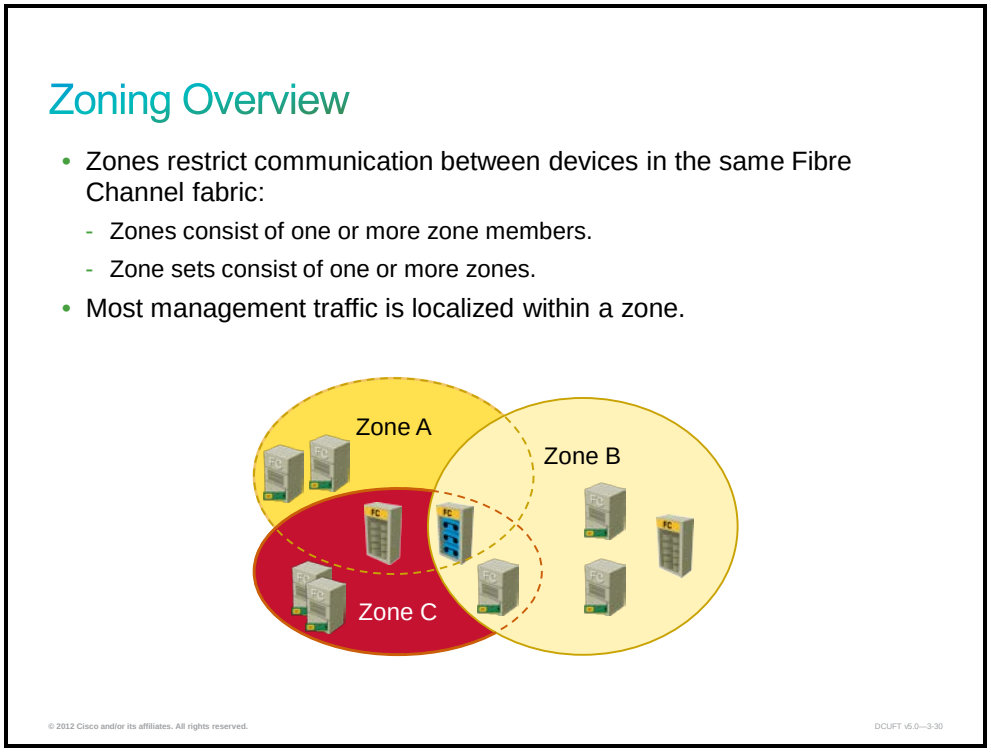
```
switch(config)# fcdomain fcid database
switch(config-fcid-db)# no vsan 2 wwn 21:01:00:1b:32:a3:c0:2e fcid
0x620800 area dynamic
```

Step 4 Bring up the blade and ensure that it receives a proper FCID.

```
Fc2/1 2 0x620104 20:10:00:05:1e:5e:6a:85 10:00:00:05:1e:5e:6a:85
Fc2/1 2 0x620123 21:01:00:1b:32:a3:c0:2e 20:01:00:1b:32:a3:c0:2e
```

Troubleshooting Fibre Channel Zoning

This topic explains how to troubleshoot issues that relate to Fibre Channel zoning on a Cisco Nexus or Cisco MDS Series switch.



With many types of servers and storage devices on the network, the need for security is critical. For example, if a host gains access to a disk that another host (potentially one with a different operating system) is using, the data on this disk could become corrupted. To avoid any compromise of critical data within the SAN, zoning allows the user to overlay a security map that dictates which devices (namely hosts) can see which targets, thus reducing the risk of data loss.

As the figure shows, a zone set consists of one or more zones:

- A zone set can be activated or deactivated as a single entity across all switches in the fabric.
- Only one zone set can be activated at any time.
- A zone can be a member of more than one zone set.
- A zone consists of multiple zone members. Members in a zone can access one another; members in different zones cannot access one another.

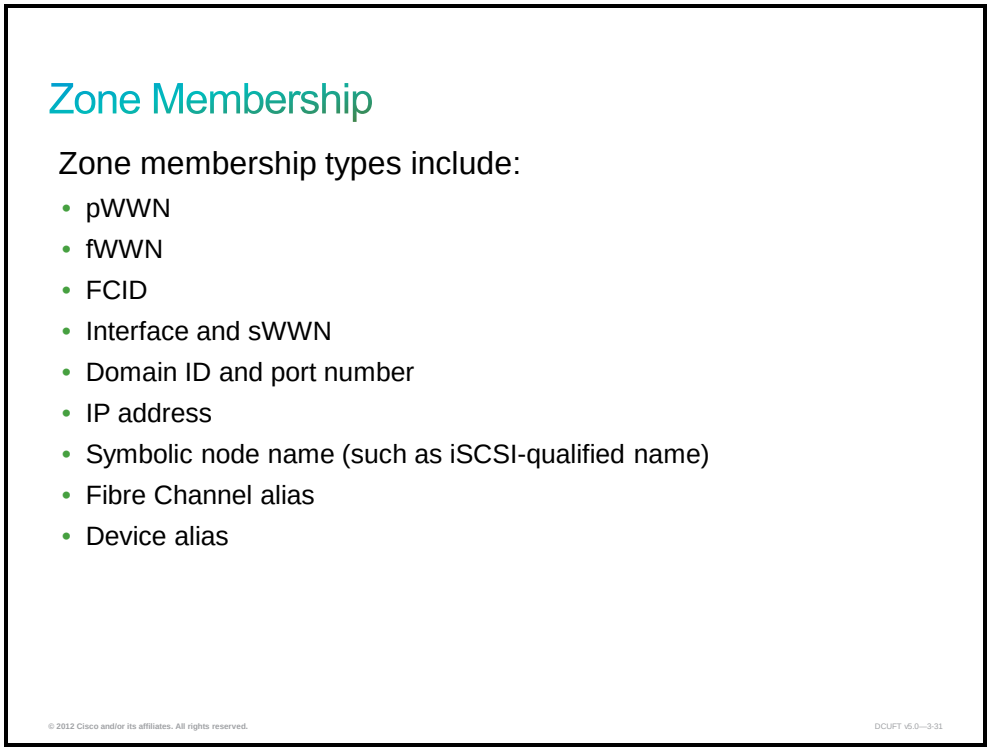
Uses for Zoning

Zoning typically has several uses:

- **Separating initiators from their targets:** Frequently, each initiator port belongs in a separate zone with its targets.
- **Separating devices:**
  - You can use zoning to separate devices that use different operating systems. This practice is useful to protect some operating systems from treating disks that are formatted by other operating systems as blank disks and potentially taking over and overwriting their storage.

- You can use zoning to separate devices that have no need to communicate with other devices in the fabric or that have classified data.
- You can use zoning to separate devices into departmental, administrative, or other functional groupings.

- **Localizing management traffic:** Most management traffic does not cross zone boundaries. Zones help to reduce the impact of management traffic on the fabric.



A zone member is used to uniquely identify a device or devices that are to be included in a zone. Zone members can be identified by using any of nine possible methods:

- **Device port world wide name (pWWN):** When zoning is based on device pWWN, zone membership is determined by using the device pWWN of a node port (N Port) that is attached to the Fibre Channel switch.
- **Fabric world wide name (fWWN):** When zoning is based on fWWN, zone membership is determined by using the pWWN of the fabric port (F Port) of the Fibre Channel switch to which the device is attached. In this case, zone membership is associated with a given port on a given line card in a given Fibre Channel switch.
- **FCID:** When zoning is based on the FCID of an N Port that is attached to the switch, zone membership is determined by the FCID that the fabric domain controller assigns to a device.

Note	For FCID zoning to be useful, you should use static FCID assignment.

- **Interface and switch world wide name (sWWN):** Zoning that is based on the switch port to which the device is attached is typically referred to as interface-based zoning. This form of zoning allows zone membership to be globally specified, based on a given sWWN and interface on that switch; for example, member interface fc3/10 swwn 20:00:00:05:30:00:91:9e.

Note	Interface- and sWWN-based zoning is not yet part of any ANSI Fibre Channel standard, so this form of zoning cannot be used in interoperability VSANs.

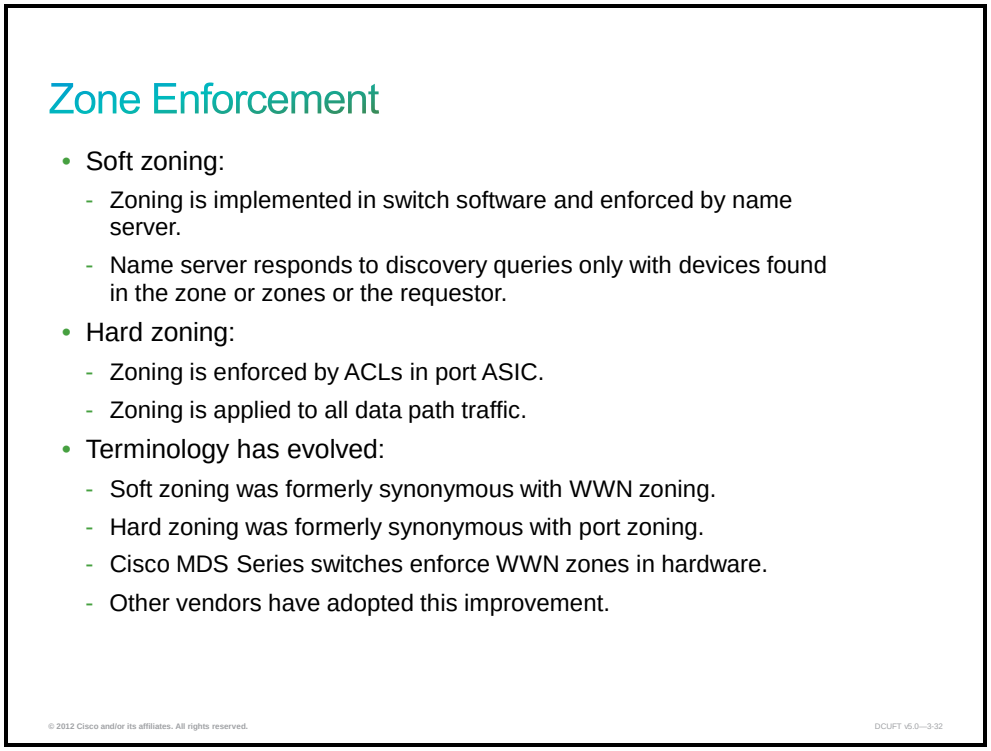
- **Domain ID and port number:** Zoning membership can be based on the domain ID of the switch and the port number on the switch to which the device is attached. Because domain IDs can be allocated dynamically, the use of static domain ID allocation is recommended. The port number is specified as a port index between 0 and 255. Associating a port number to a given module slot and port combination can be difficult, so you should use zoning

based on interface and sWWN rather than domain ID and port number. Domain ID and port number-based zoning is specified in the ANSI Fibre Channel standards, but Fibre Channel vendors can number ports in different ways. Therefore, this form of zoning is not recommended for multivendor fabrics and is unavailable for standards-based interoperability mode.

- **IP address:** Zoning membership can be based on device IP address. This form of zoning can be used for Internet Small Computer Systems Interface (iSCSI) devices.
- **Symbolic node name:** Symbolic node name zoning allows zone members to be defined by their unique symbolic node name, such as the iSCSI Qualified Name (IQN) or IP address that is associated with iSCSI devices. In this manner, iSCSI devices can have dynamic pWWN and node world wide names (nWWNs) associated with them but can continue to make use of zoning membership by using their globally unique IQN.

Note	Symbolic node name-based zoning is not yet part of any ANSI Fibre Channel standard, so this form of zoning is unavailable for multivendor fabrics.

- **Fibre Channel alias**: Zoning membership is based on a previously defined Fibre Channel alias (**fcalias** command).



There are two basic ways to enforce zoning: software-based (soft) zoning and hardware-based (hard) zoning:

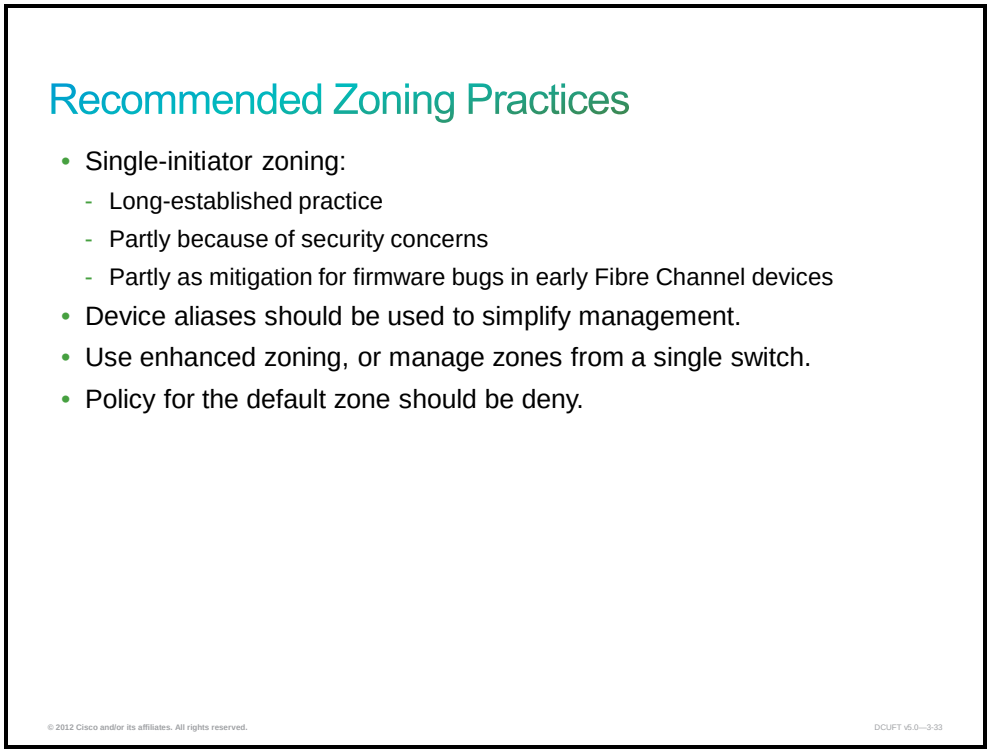
- **Soft zoning:** Soft zoning is enforced by the FCNS service. When devices connect to a Fibre Channel fabric, they use the FCNS to correlate WWNs to FCIDs and to discover devices on the fabric. With soft zoning, a Fibre Channel switch responding to a name server query from a device responds with a list of only those devices that are registered in the same zone or zones as the querying device.
- **Hard zoning:** Hard zoning is enforced through access control lists (ACLs) that the switch port ASIC applies to every Fibre Channel frame that is switched.

Because soft zoning does not enforce ACLs on a per-frame basis, soft zoning is not as secure as hard zoning. Soft zoning does not prevent a rogue device from attaching to Cisco Data Center Network Manager (DCNM) for SAN and obtaining a list of all FCIDs.

Until recently, most Fibre Channel switches enforced interface-based zone membership in hardware, so the term hard zoning has often been used synonymously with port zoning. Similarly, WWN-based zone members were enforced only by the name server, so the term soft zoning has been used synonymously with WWN zoning. However, this is no longer the case.

Cisco MDS 9000 Series switches support hard zoning for 8000 zones and 20,000 zone members, as well as supporting soft zoning. Soft-zoning information is downloaded to a hardware ternary content addressable memory (TCAM) module that enforces soft zoning in the port ASIC. Therefore, all zone memberships on Cisco MDS 9000 Series switches are hardware-enforced. Cisco Nexus 5500 Platform switches support hard zoning for 640 zones per switch and 1280 zone members. Other switch vendors are also beginning to enforce WWN zone membership at the hardware level. There is no longer a one-to-one relationship between zone membership type (interface or WWN) and enforcement method (soft or hard).

Note	The hard zoning from some vendors reverts to soft zoning when a threshold number of zone members is exceeded.

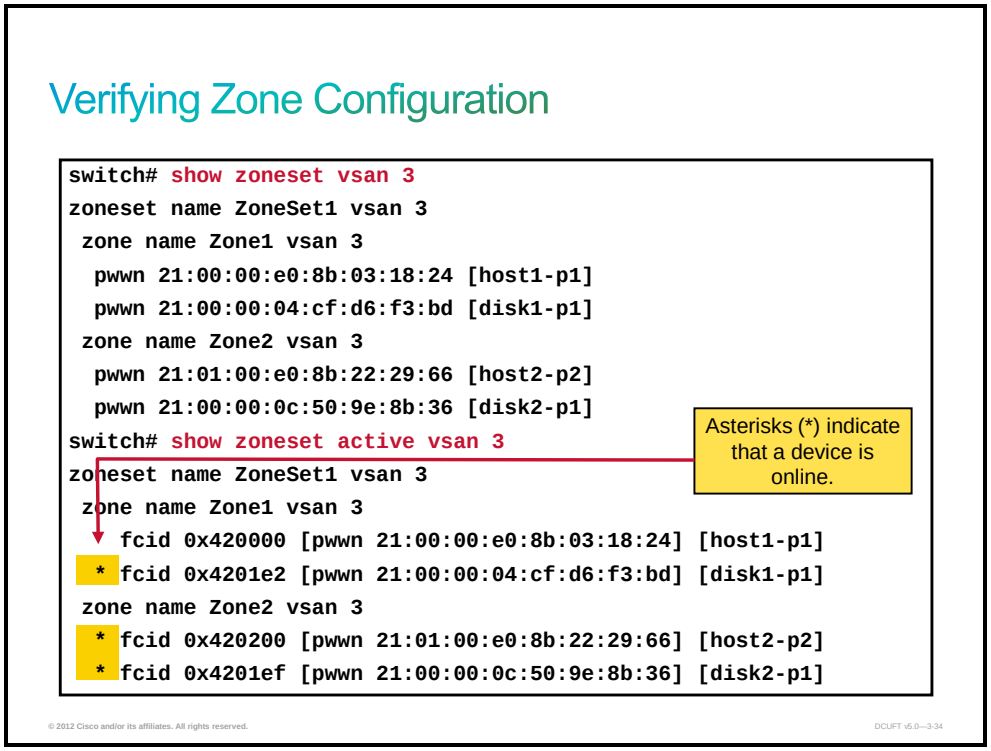


Follow these guidelines for implementing zones:

- Zoning should always be implemented in a Fibre Channel fabric, if not from a security perspective, then from the perspective of minimizing loss of data. In general, you should use as many zones as there are HBAs that communicate with storage. For example, if you have two hosts each, with two HBAs that are communicating with three storage devices, you should use four zones. This type of zoning is sometimes referred to as single-initiator zoning.

Note	Single-initiator zoning is a long-established practice that was popularized partly because of the additional level of security that is provided but also as mitigation for firmware bugs in early Fibre Channel devices. Although single-initiator zoning is still considered a best practice, this practice is not as important as it was in the early years of Fibre Channel and might not be appropriate for all SAN environments.

- To simplify management, use Fibre Channel aliases wherever possible. It is easier to identify devices with aliases than with WWNs. In general, you should assign aliases to WWNs.
- Zone administration should generally be confined to one Fibre Channel switch within a given fabric, to ensure that there is no possibility of activating an incomplete zone set (which might happen if the full zone set database is not consistent across Fibre Channel switches).
- Leave the default zone policy as deny so that devices cannot inadvertently access each other when placed in the default zone.



To display the current zoning configuration on the local switch, in EXEC mode use the command **show zoneset**.

To verify the current active zone set, use **show zoneset active** in EXEC mode. The asterisks (*) indicate that a device is visible (online). A missing asterisk might indicate an offline device or an incorrectly configured zone, possibly a mistyped pWWN.

show zone analysis Command

To display zone and zone set information, use the **show zone analysis** command:

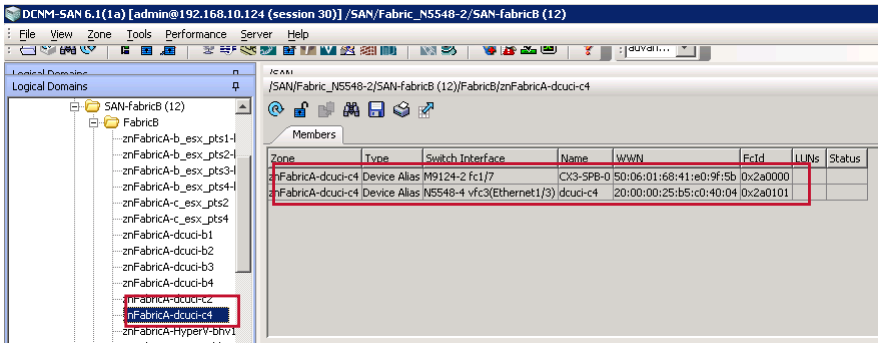
- **show zone analysis vsan 10**
- **show zone analysis active vsan 10**
- **show zone analysis zoneset zs1 vsan 10**

```
switch# show zone analysis active vsan 10
Zoning database analysis vsan 10
  Active zoneset: zoneset1
    Activated at: 16:57:22 UTC May 04 2012
    Activated by: Local [ CLI ]
    Default zone policy: Deny
    Number of devices zoned in vsan: 2/4 (Unzoned: 2)
    Number of zone members resolved: 2/2 (Unresolved: 0)
    Num zones: 1
    Number of IVR zones: 0
    Number of IPS zones: 0
    Formatted size: 60 bytes / 2048 Kb
```


To better manage the zones and zone sets on your switch and to display zone and zone set information, use the show zone analysis command:

```
switch# show zone analysis active vsan 10
Zoning database analysis vsan 10
  Active zoneset: zoneset1
    Activated at: 16:57:22 UTC May 04 2012
    Activated by: Local [ CLI ]
    Default zone policy: Deny
    Number of devices zoned in vsan: 2/4 (Unzoned: 2)
    Number of zone members resolved: 2/2 (Unresolved: 0)
    Num zones: 1
    Number of IVR zones: 0
    Number of IPS zones: 0
    Formatted size: 60 bytes / 2048 Kb
switch# show zone analysis zoneset zoneset1 vsan 10
Zoning database analysis vsan 10
  Zoneset analysis: zoneset1
    Num zonesets: 1
    Num zones: 1
    Num aliases: 0
    Num attribute groups: 0
    Formatted size: 112 bytes / 2048 Kb
```

Displaying Zone Information in Cisco DCNM for SAN



Zone sets, zones, and zone member information can be displayed by expanding the zone set folder for the VSAN in question. Selecting a zone or zone set highlights member devices in the zone in the Cisco DCNM for SAN topology map. Zone configuration changes can be made from the Zone > Edit Local Full Zone Database menu.

Troubleshooting Fabric Merges

Fabric merge failures are usually caused by one of these factors:

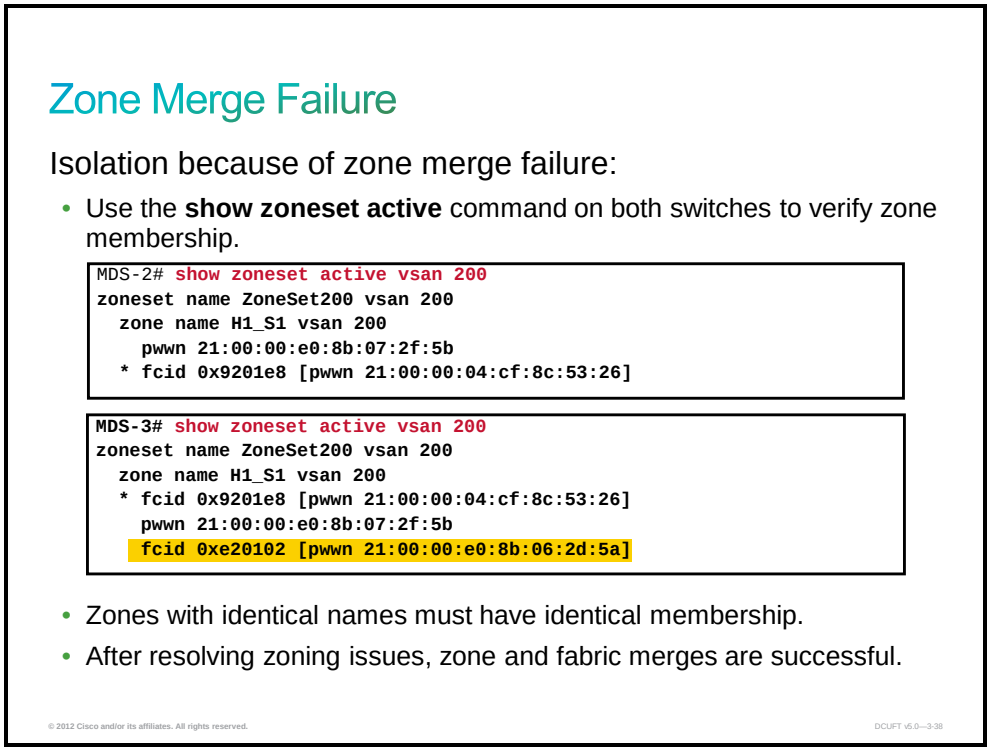
- Domain ID conflicts
  - All switches in a fabric must have a unique domain ID.
- Zone merge failure
  - Zone merge failure is usually caused by incompatibilities in the active zone databases.
- Incorrect VSAN allow list on ISL
  - VSAN allow list does not allow traffic to flow between switches.
- Missing VSAN
  - VSAN is missing from one switch, so there is no fabric to merge with.

Domain ID conflicts can cause merge failures. Therefore, all switches in a fabric must have a unique domain ID.

Zone merge failure is usually the result of incompatibilities in the active zone databases. If two fabrics have active zones with the same name, then the members of those zones must be identical.

If the VSAN allow list on an ISL is not allowing traffic to flow between switches, check the interfaces on both ends of the isolated ISL. If only one interface shows the isolation, then the VSAN that shows the isolation is probably missing from the VSAN allow list over the trunking expansion port (TE Port).

If the VSAN is missing from one of the switches, a fabric merge does not occur because only VSANs with identical IDs can merge. For example, VSAN 100 can merge with VSAN 100 but not with VSAN 200.



Use the **show logging** command to find out whether the isolation is because of a zone merge:

```
%ZONE-2-ZS_MERGE_FULL_DATABASE_MISMATCH: %$VSAN 10%$ Zone merge full
database mismatch on interface port-channel 2
%ZONE-2-ZS_MERGE_FAILED: %$VSAN 10%$ Zone merge failure, isolating
interface port-channel 2
%PORT-5-IF_TRUNK_DOWN: %$VSAN 10%$ Interface port-channel 2, vsan 10
is down (Isolation due to zone merge failure)
```

Use the **show zoneset active** command against the VSAN that is experiencing isolation because of zone merge failure on both switches. Zones with identical names must have identical membership. In this example, an additional member FCID 0xe20102 *[pwwn 21:00:00:e0:8b:06:2d:5a]* in zone H1_S1 on MDS-3 is not present in the zone database of VSAN 200 on MDS-2.

There are several ways to fix this conflict:

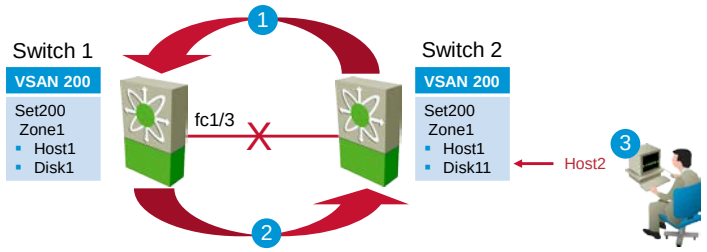
- The Cisco MDS 9000 Series switches and Cisco Nexus 5500 Platform switches support **zone merge import** and **zone merge export** commands, which allow you to push or pull the active database from one switch to another, thus eliminating one database and using only the other.
- Rename one of the zones to a name that is not duplicated on the other switch.
- In this example, remove the device FCID 0xe20102 [pwwn 21:00:00:e0:8b:06:2d:5a] from MDS-3, or add the same device to MDS-2.

All these methods create identical zone databases and allow the fabric merge to succeed.

In terms of methodology, the output that the **show zoneset active** command generates is good for highlighting which individual zone has the conflict. However, before you go looking for that information, you need to know that the isolation is caused by a zone merge. The **show logging** command output can help you to do so.

Zone Set Import and Export

When merging fabrics, TE and E Ports might become isolated if the active zone set databases differ between the two switches or fabrics.



You can recover from isolation by using one of three options:

1. Import the active zone set database of the neighboring switch and replace the current active zone set.
2. Export the current database to the neighboring switch.
3. Manually resolve the conflict by editing the full zone set, activating the corrected zone set, and then bringing up the link.

The Cisco MDS 9000 Series and Cisco Nexus 5500 Platform switches provide a facility to correct a merge failure, either by importing the database of an adjacent switch or by exporting its database to the adjacent switch. This capability avoids the need to manually edit and fix the configuration at either of the switches. Use this capability with caution, because it affects all devices within the configured zones.

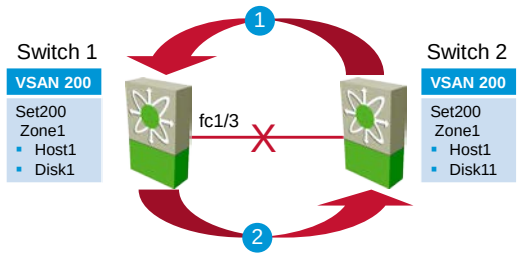
If a zone merge fails, you can use one of three options to recover from isolation:

- Import the active zone set database of the neighboring switch and replace the current active zone set.
- Export the current database to the neighboring switch.
- Manually resolve the conflict by editing the full zone set, activating the corrected zone set, and then bringing up the link.

Use these commands to resolve a failed zone merge by importing or exporting an active zone set:

- **zoneset import interface** *interface-id* **vsan** *vsan-id*
- **zoneset export vsan** *vsan-id*

Zone Set Import and Export (Cont.)



Import the zone set from the adjacent switch that is connected through the fc1/3 interface for VSAN 200.

1 `switch# zoneset import interface fc1/3 vsan 200`

Export the zone set to the adjacent switch that is connected through VSAN 200.

2 `switch# zoneset export vsan 200`

An E Port is segmented (isolation because of zone merge failure) if these conditions are true:

- The active zone sets on the two switches differ from each other in terms of zone membership (provided there are zones at either side with identical names).
- The active zone set on both switches contains a zone with the same name but with different zone members.

To use the CLI to resolve the link isolation that a failed zone merge caused, follow these steps:

Step 1 Use the **show interface** command to confirm that the port is isolated because of a zone merge failure.

```
switch# show interface fc 1/3
Fc1/3 is down (Isolation due to zone merge failure)
Hardware is Fibre Channel, WWN is 20:4e:00:05:30:00:63:9e
vsan is 200
Beacon is turned off
40 frames input, 1056 bytes, 0 discards
0 runts, 0 jabber, 0 too long, 0 too short
0 input errors, 0 CRC, 3 invalid transmission words
0 address id, 0 delimiter
0 EOF abort, 0 fragmented, 0 unknown class
79 frames output, 1234 bytes, 16777216 discards
Received 23 OLS, 14 LRR, 13 NOS, 39 loop inits
Transmitted 50 OLS, 16 LRR, 21 NOS, 25 loop inits
```

Step 2 Verify the zoning information by using the following commands on each switch:

- **show zone vsan** *vsan-id*
- **show zoneset vsan** *vsan-id*

Step 3 You can use two approaches to resolve a zone merge failure by overwriting the zoning configuration of one switch with the configuration of the other switch. To do so, use either of these commands:

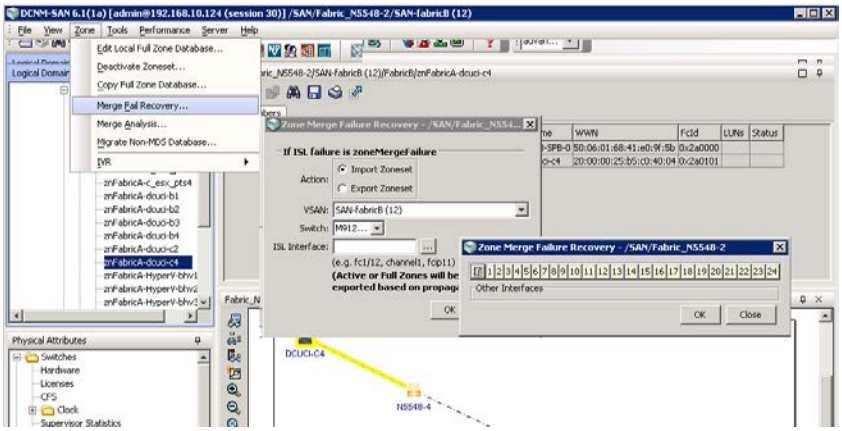
- **zoneset import interface** *interface-id* **vsan** *vsan-id*
- **zoneset export vsan** *vsan-id*

The **import** option in the first command overwrites the active zone set of the local switch with that of the remote switch. The **export** option in the second command overwrites the active zone set of the remote switch with that of the local switch.

Note If the zoning databases between the two switches are overwritten, you cannot use the **import** option. To work around this, you can manually change the content of the zone database on either switch, and then issue a **shutdown/no shutdown** command sequence on the isolated port.
If the isolation is specific to one VSAN and not on an E Port, the correct way to issue the cycle up and down is to remove the VSAN from the list of allowed VSANs on that trunk port and reinsert it.

Zone Set Import and Export in Cisco DCNM for SAN

To initiate import or export of an active zone set, choose **Zone > Merge Fail Recovery** from Cisco DCNM for SAN.



Zone Set Import and Export in Cisco DCNM for SAN

Note	Importing from one switch and exporting from another switch can lead to isolation again.

You can import active zone sets (perform a merge fail recovery) if the cause of an ISL failure is a zone merge failure.

To import an active zone set, follow these steps:

Step 1 From DCNM for SAN, choose **Zone** > **Merge Fail Recovery**. You will see the Zone Merge Failure Recovery dialog box.

Step 2 Choose the **Import Zoneset** radio button.

Step 3 Choose the switch from which to import the zone set information from the drop-down list.

Step 4 Choose the VSAN from which to import the zone set information from the drop-down list.

Step 5 Choose the interface to use for the import process.

Step 6 Click the **OK** button to import the active zone set, or click the **Close** button to close the dialog box without importing the active zone set.

Exporting Active Zone Sets

You can export active zone sets (perform a merge fail recovery) if the cause of an ISL failure is a zone merge fail.

To export an active zone set, follow these steps:

Step 1 From DCNM for SAN, choose **Zone** > **Merge Fail Recovery**. You will see the Zone Merge Failure Recovery dialog box.

Step 2 Choose the **Export Zoneset** radio button.

Step 3 Choose the switch to which to export the zone set information from the drop-down list.

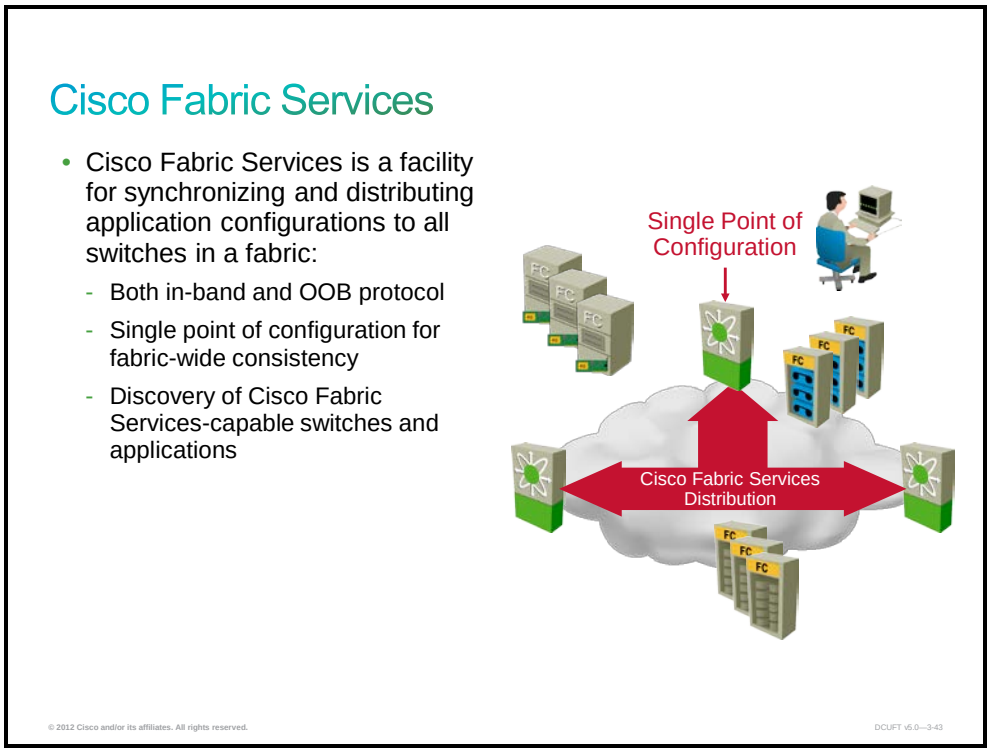
Step 4 Choose the VSAN to which to export the zone set information from the drop-down list.

Step 5 Choose the interface to use for the export process.

Step 6 Click the **OK** button to export the active zone set, or click the **Close** button to close the dialog box without exporting the active zone set.

Troubleshooting Cisco Fabric Services on Cisco MDS Series and Cisco Nexus Switches

This topic explains how to troubleshoot issues that relate to Cisco Fabric Services on a Cisco MDS Series or Cisco Nexus switch.



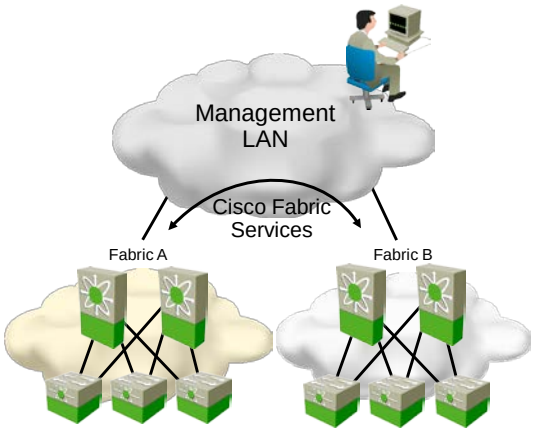
Many features in Cisco MDS 9000 Series Multilayer Switches require configuration synchronization in all switches in the fabric. Maintaining configuration synchronization across a fabric is important to maintain fabric consistency. In the absence of a common replication infrastructure, this synchronization is achieved through manual configuration of each switch within the fabric. This manual process is tedious and error prone.

The Cisco NX-OS Software uses the Cisco Fabric Services infrastructure to enable efficient database distribution and to foster device flexibility. Cisco Fabric Services simplifies SAN provisioning by automatically distributing configuration information to all switches within a fabric. Several Cisco NX-OS applications use the Cisco Fabric Services infrastructure to maintain and distribute the contents of the database of a particular application.

Cisco Fabric Services provides a common infrastructure for automatic configuration synchronization in the fabric. This infrastructure provides the transport function as well as a rich set of CiscoWorks Common Services to the applications. Cisco Fabric Services has the ability to discover Cisco Fabric Services-capable switches in the fabric and can discover the application capabilities of all Cisco Fabric Services-capable switches.

Cisco Fabric Services OOB Distribution

- Switches attempt distribution over Fibre Channel first; upon failure, distribute over IP.
- Do not send duplicate messages if distribution over both IP and Fibre Channel is enabled.
- Cisco Fabric Services applications with the physical-all distribution scope can be used with IP distribution.



Cisco Fabric Services distribution can be performed out-of-band (OOB) over IP to synchronize Cisco Fabric Services-enabled applications for physically separate fabrics. Cisco Fabric Services distribution over IP supports these features:

- Physical distribution entirely over an IP network
- Physical distribution over a hybrid Fibre Channel and IP network, with the distribution reaching all switches that are reachable over either Fibre Channel or IP

Note	The switch attempts to distribute information first over Fibre Channel, and then over the IP network if the first attempt over Fibre Channel fails. Cisco Fabric Services does not send duplicate messages if distribution over both IP and Fibre Channel is enabled. Distribution can be over IPv4 or IPv6. Cisco Fabric Services cannot distribute over both IPv4 and IPv6 from the same switch.

The Cisco Fabric Services keepalive mechanism uses a configurable multicast address to detect network topology changes:

- Distribution for logical scope applications is not supported, because the VSAN implementation is limited to Fibre Channel.

Note	OOB communication is the exchange of control information in a separate band of the data channel.

Cisco Fabric Services Applications

Application	Distribution Mode	Scope	Cisco Fabric Services Enabled?
Call Home	Coordinated	Physical	No
Global Device Alias	Coordinated	Physical	Yes
DPVM	Coordinated	Physical	Yes
IVR	Both	Physical	No
iSNS	Uncoordinated	Physical	Yes
NTP	Coordinated	Physical	No
Port Security	Coordinated	Logical	No
RADIUS and TACACS+	Coordinated	Physical	No
Role (role-based access control)	Coordinated	Physical	No
Syslog	Coordinated	Physical	No
VSAN fctimer (fabric timeout values)	Coordinated	Physical	No
fcdomain allowed list	Coordinated	Logical	No
RSCN event timer	Coordinated	Logical	No
SCSI Flow Services	Coordinated	Selected	Yes
iSCSI Load Balancing	Coordinated	Physical	No
Fabric Startup Configuration Manager	Coordinated	Physical	Yes
FlexAttach Virtual pWWN	Coordinated	Physical	Yes

These Cisco NX-OS features use the Cisco Fabric Services infrastructure:

- NPV
- FlexAttach virtual pWWN
- Network Time Protocol (NTP)
- Dynamic Port VSAN Membership (DPVM)
- Distributing Device Alias Service (DDAS)
- IVR topology
- SAN device virtualization
- TACACS+ and RADIUS (authentication, authorization, and accounting [AAA] services)
- User and administrator roles
- Port security
- Internet Storage Name Service (iSNS)
- Call Home
- Syslog (system message log)
- Fibre Channel timer (fctimer)
- Small Computer Systems Interface (SCSI) flow services
- Saving startup configurations in the fabric using the Fabric Startup Configuration Manager (FSCM)
- Allowed domain ID lists
- Registered State Change Notification (RSCN) timer
- iSCSI server load balancing (iSLB)
- FlexAttach Virtual pWWN

The table in the figure is provided for reference and displays these items:

- Application name, as seen in the output of the **show Cisco Fabric Services application** command
- Distribution mode (coordinated, uncoordinated, or both)
- Distribution scope (logical or physical)
- Whether Cisco Fabric Services distribution is enabled by default when the application is enabled

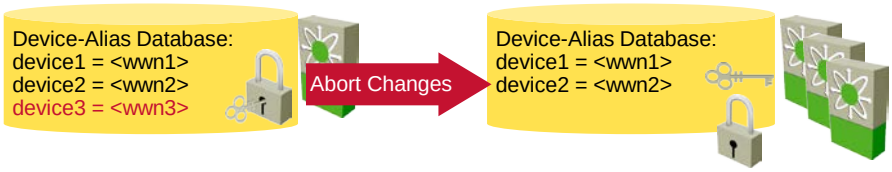
Note	IVR uses coordinated distribution mode and uncoordinated distribution mode for unrelated data.

Cisco Fabric Services Applications (Cont.)

Applications enable or disable Cisco Fabric Services distribution capability.

Pending application changes can be aborted:

- Abort is allowed only by the user that initiated the lock on the switch on which the pending database was created.
- The pending database is flushed and locks are released fabricwide.



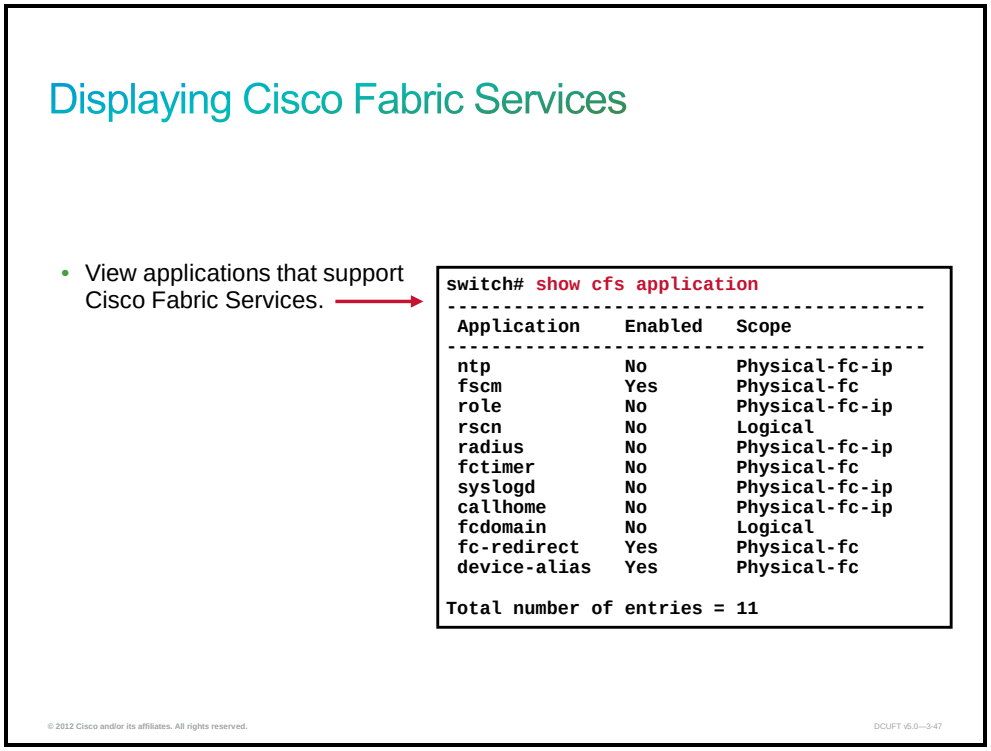
Locks can be cleared to recover from an inconsistent state:

- Can be invoked from any switch in the fabric by using administrator privileges

All switches in the fabric must be Cisco Fabric Services-capable. Switches that are not Cisco Fabric Services-capable do not receive distributions. This results in a portion of the fabric not receiving the intended distribution.

Cisco Fabric Services has these requirements:

- **Implicit Cisco Fabric Services usage:** The first time you issue a Cisco Fabric Services task for a Cisco Fabric Services-enabled application, the configuration modification process begins, and the application locks the fabric.
- **Pending database:** The pending database is a temporary buffer to hold uncommitted information. The uncommitted changes are not applied immediately, to ensure that the database is synchronized with the database in the other switches in the fabric. When you commit the changes, the pending database overwrites the configuration database, which is also known as the active database or the effective database.
- **Cisco Fabric Services distribution enabled or disabled on a per-application basis:** The default (enable or disable) for the Cisco Fabric Services distribution state differs between applications. If Cisco Fabric Services distribution is disabled for an application, that application does not distribute any configuration and does not accept a distribution from other switches in the fabric.
- **Explicit Cisco Fabric Services commit:** Most applications require an explicit commit operation to copy the changes in the temporary buffer to the application database, to distribute the new database to the fabric, and to release the fabric lock. The changes in the temporary buffer are not applied if you do not perform the **commit** operation.



All Cisco Fabric Services-based applications provide an option to enable or disable the distribution capabilities.

Verifying Application Registration Status

The **show Cisco Fabric Services application** command displays the applications that are currently registered with Cisco Fabric Services. As the figure shows, the first column of output displays the application name. The second column indicates whether the application is enabled or disabled for distribution (yes or no). The last column indicates the scope of distribution for the application (logical, physical, or both).

Note	The **show Cisco Fabric Services application** command displays only applications that are registered with Cisco Fabric Services. Conditional services that use Cisco Fabric Services do not appear in the output, unless these services are running.

The **show Cisco Fabric Services application name** command displays the details for a particular application. The command displays the enabled or disabled state, timeout as registered with Cisco Fabric Services, merge capability (if the application has registered with Cisco Fabric Services for merge support), and the distribution scope.

```
switch# show cfs application name ntp
Enabled       : Yes
Timeout       : 5s
Merge Capable : Yes
Scope         : Physical
```


Verifying Cisco Fabric Services Peers

- Displays all Cisco Fabric Services-capable switches in the fabric
- Independent of individual application registrations
- Local switch indicated as [Local]

```
switch# show cfs peers
Physical Fabric
-------------------------------------------------------
 Switch WWN               IP Address
-------------------------------------------------------
 20:00:00:0d:ec:a0:7e:80 10.0.1.11             [Local]
                         P10-MDS-1
 20:00:00:0d:ec:94:64:80 10.0.1.12

Total number of entries = 2
```


The **show Cisco Fabric Services peers** command displays all the switches in the physical fabric, in terms of the switch WWN and the IP address. The local switch is indicated as [Local].

Verifying Cisco Fabric Services Lock

- If the application does show in the output, then the distribution has not completed yet.

```
switch# show cfs lock
Application: callhome
Scope      : Physical-fc-ip
------------------------------------------------------------------------
 Switch WWN               IP Address         User Name     User Type
------------------------------------------------------------------------
 20:00:00:22:55:79:a4:c1  172.28.230.85        admin       CLI/SNMP v3
                          switch
Total number of entries = 1
```

- If the application does not show in the output, then the distribution has completed.

Use the **show cfs lock** command to verify that a distribution is not in progress in the network for the application. If the application does not show in the output, the distribution has completed.

Verifying Cisco Fabric Services Merge Status

```
switch# show cfs merge status name port-security
Logical [VSAN 1] Merge Status: Failed
Local Fabric
-----------------------------------------------------------
 Domain Switch WWN               IP Address
-----------------------------------------------------------
 238    20:00:00:05:30:00:6b:9e  10.76.100.167 [Merge Master]
Remote Fabric
-----------------------------------------------------------
 Domain Switch WWN               IP Address
-----------------------------------------------------------
 236    20:00:00:0e:d7:00:3c:9e  10.76.100.169 [Merge Master]
Logical [VSAN 2] Merge Status: Success
Local Fabric
-----------------------------------------------------------
 Domain Switch WWN               IP Address
-----------------------------------------------------------
 211    20:00:00:05:30:00:6b:9e  10.76.100.167 Merge Master]
 1      20:00:00:0e:d7:00:3c:9e  10.76.100.169
```

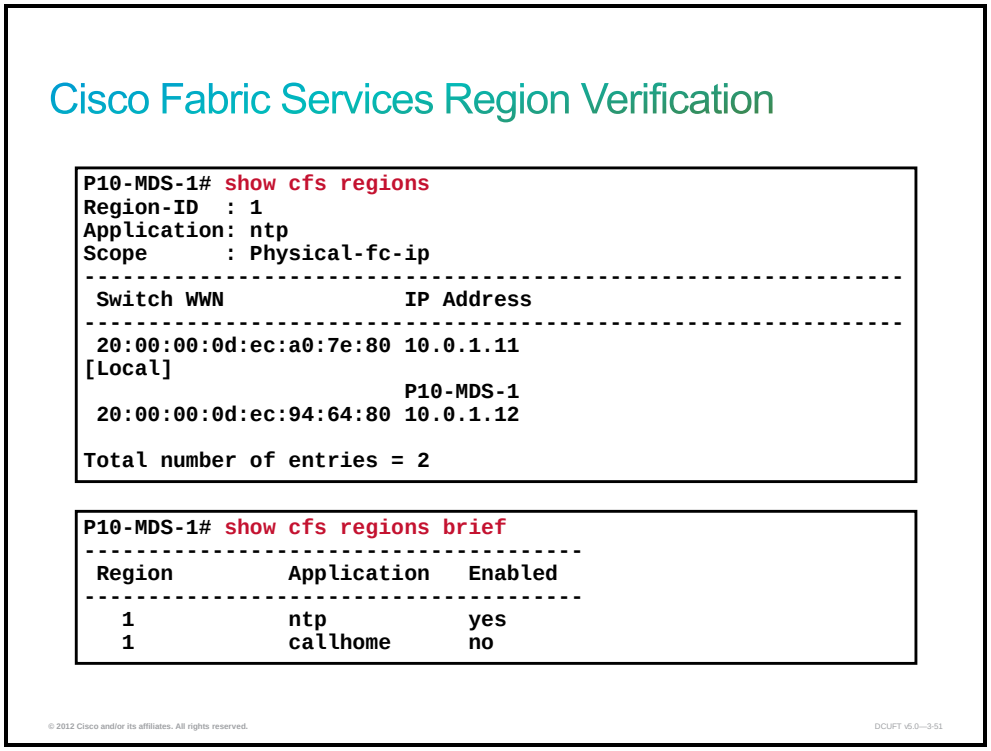
VSAN 1

VSAN 2

The **show Cisco Fabric Services merge status name** *application* command displays the merge status for a given application. For an application that distributes in logical scope, this command shows the merge status in all the valid VSANs on the switch.

The merge status is Success, Waiting, Failure, or In Progress. In a successful merge, all the switches in the fabric are shown under the local fabric. In a merge failure or when a merge is in progress, the local fabric and the remote fabric that are involved in the merge are indicated separately.

The application server in each fabric that is responsible for the merge is indicated by the term merge master. The merge master is selected based on the lowest sWWN in that fabric or SAN.



Use the **show Cisco Fabric Services regions** command to verify Cisco Fabric Services region status. The command displays only distribution-enabled applications.

This command displays the complete Cisco Fabric Services region database:

```
switch# show cfs regions
```

This command restricts the display of the Cisco Fabric Services region database:

```
switch# show cfs regions brief region id
```

Other helpful **show** commands for Cisco Fabric Services regions include these:

```
switch# show cfs regions region id
switch# show cfs regions name app-name
```

Summary

This topic summarizes the key points that were discussed in this lesson.

Summary

- Use the **show vsan** command to verify VSAN configuration and the **show vsan membership** command to verify VSAN membership.
- In a Fibre Channel network, the principal switch issues domain IDs when a new switch is added to an existing fabric. Use the **show fcdomain** command to display the domain database.
- Use the **show fcns** and **show fcns details** commands to display the FCNS and statistical information for a specific VSAN or for all VSANs.
- To display the current zoning configuration on the local switch, use the **show zoneset** command; to verify the current active zone set, use the **show zoneset active** command.
- The **show cfs application** command displays the applications that are currently registered with Cisco Fabric Services.

Lesson 3

Troubleshooting NPV Mode

Overview

This lesson is designed to provide you with examples of common issues that relate to SAN switching when the switch is in node port (N-Port) Virtualization (NPV) mode. The lesson also shows you how to identify and resolve these issues.

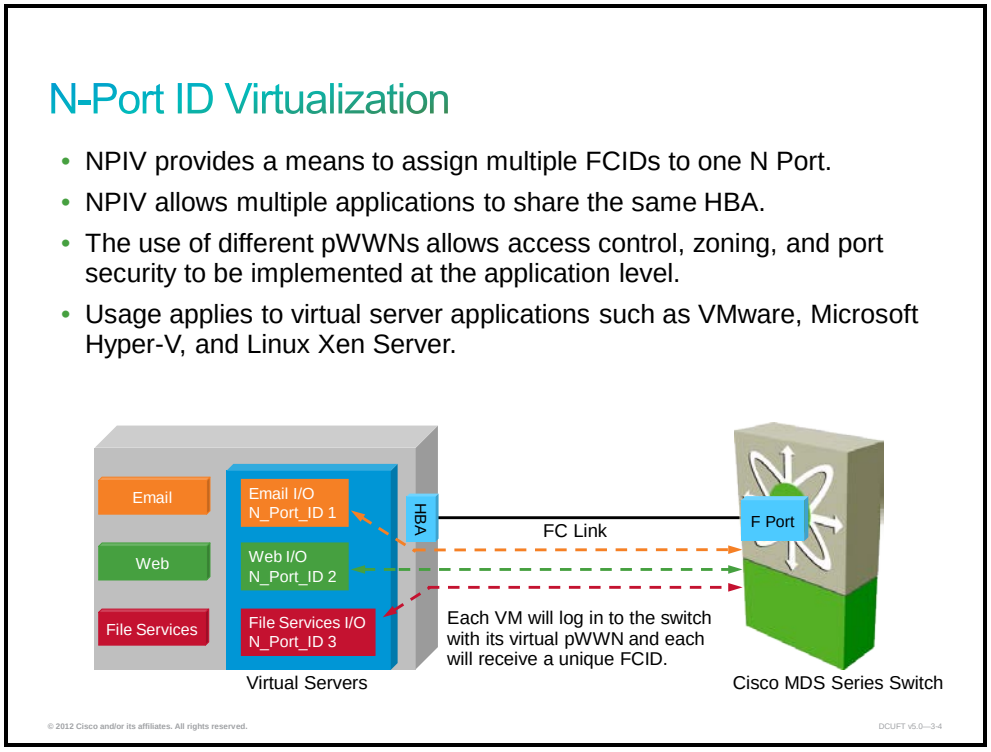
Objectives

Upon completing this lesson, you will be able to identify and resolve issues that relate to Fibre Channel switching when the Cisco Nexus Operating System (NX-OS) switch is used in NPV mode. This ability includes being able to meet these objectives:

- Explain the FLOGI process, both when multiple VMs are hosted on a server that is connected to a Cisco Nexus or Cisco MDS Series switch running NPIV and when a Cisco Nexus or Cisco MDS Series switch is running in NPV mode
- Explain how to troubleshoot issues that relate to NPV mode on a Cisco Nexus or Cisco MDS Series switch

NPIV and NPV Mode FLOGI Processes

This topic explains the fabric login (FLOGI) process when multiple virtual machines (VMs) are hosted on a server that is connected to a Cisco Nexus or Cisco MDS Series switch running N-Port ID Virtualization (NPIV) and explain the FLOGI process when a Cisco Nexus or Cisco MDS Series switch is running in NPV mode.

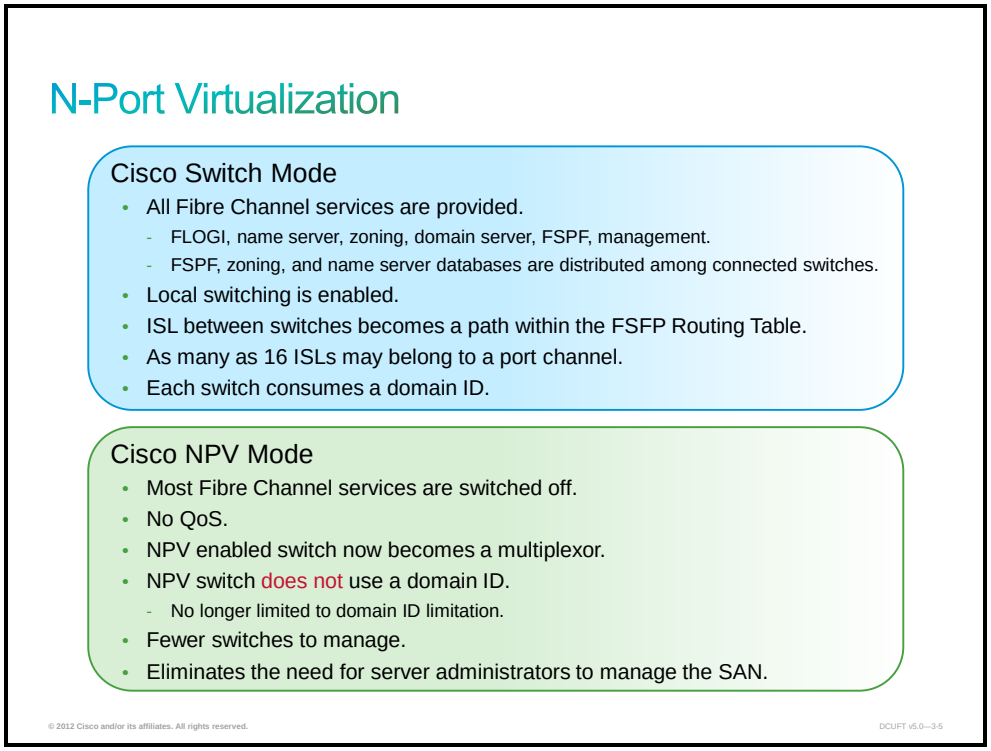


A single N Port on the switch can support a FLOGI from only one N Port device. The N Port sends a FLOGI and receives a unique Fibre Channel ID (FCID) from the switch.

NPIV provides a means to assign multiple FCIDs to one N Port. Therefore, NPIV allows multiple applications to share the same host bus adapter (HBA) port.

The use of different port world wide names (pWWNs) allows access control, zoning, and port security to be implemented at the application level.

Usage applies to virtual server applications such as VMware, Microsoft Hyper-V, and Linux Xen Servers.

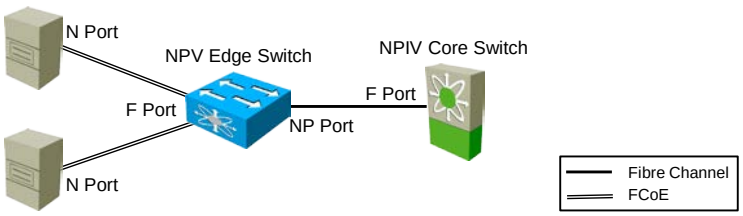


There are two operating modes for switches that support NPV:

- Switch mode
  - All Fibre Channel services are provided:
    - FLOGI, name server, zoning, domain server, Fabric Shortest Path First (FSPF), and management
    - FSPF, zoning, and name server databases are distributed among connected switches.
  - Local switching is enabled.
  - The Inter-Switch Link (ISL) between switches becomes a path within the FSPF routing table.
  - As many as 16 ISLs may belong to a single port channel.
  - Each switch consumes a domain ID.
- NPV mode
  - Most Fibre Channel services are switched off.
  - The NPV-enabled switch now becomes a multiplexor.
  - The NPV-enabled switch does not use a domain ID; therefore, you are no longer limited to the domain ID limitation.
  - You have fewer switches to manage.
  - NVP mode eliminates the need for server administrators to manage the SAN.

NPV Mode

- Provides physical port-level virtualization of multiple Fibre Channel end nodes to one F Port off a Fibre Channel switch:
  - Cisco Nexus 5000 Series Switch operates in N-Port proxy mode (not in Fibre Channel switch mode)
  - Simplifies multivendor interoperation
- Eliminates the Fibre Channel domain on the Cisco Nexus 5000 Series Switch
- Simplifies management
- Used in conjunction with NPIV on the core Fibre Channel switch



Each Fibre Channel switch typically has one domain ID and connects to the upper-layer Fibre Channel switch through an expansion port (E Port). The host that is connected to the switch first sends the FLOGI to the switch. The switch intercepts the FLOGI and assigns the FCID to the host by sending an Accept frame. The Fibre Channel switch also provides other Fibre Channel services, such as name registration, binding check, zoning check, and so on. The Fibre Channel switch runs the FSPF routing protocol. The switch uses the Fibre Channel switch table to compute the path to reach the remote domains. The Fibre Channel switch maintains the FCID to the port mapping table (station table) for a locally attached host. When the Fibre Channel switch receives a frame, it checks the domain ID of the destination ID (DID) first. If that domain ID is the same as the domain ID of the switch, then the local station table is used to figure out the destination port. If that domain ID is different from the domain ID of the switch, then the Fibre Channel switch searches the Fibre Channel switch table to identify the egress port.

The fact that each Fibre Channel switch requires one domain ID poses scalability problems for the fabric. Each virtual storage area network (VSAN) can have as many as 239 domain IDs. Unfortunately, many storage vendors do not support as many as 239 domain IDs. This restriction means that no more than 40 domain IDs are typically found within a single fabric (or VSAN).

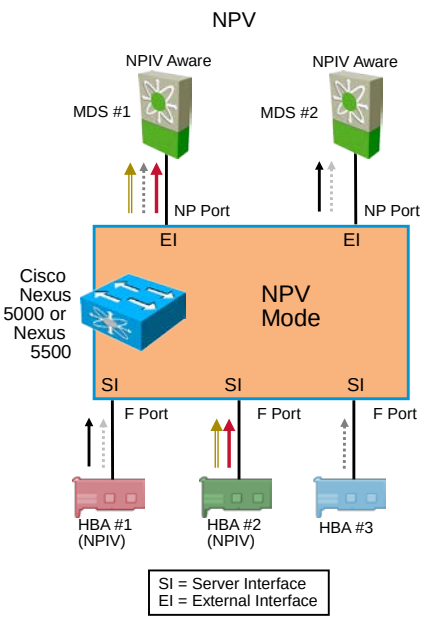
With Cisco NPV, Cisco Nexus 5000 Series or 5500 Platform switches relay the FLOGI and discover Fabric Service parameters (fabric discovery [FDISC]) to the upstream Fibre Channel switch. In this mode, a Cisco Nexus 5000 Series or 5500 Platform switch operates as a N-Port proxy (NP Port) mode. The switch does not perform any Fibre Channel switching itself. There are no local switching and zoning checks on Cisco Nexus switches. The Cisco Nexus switch appears as a host to the upper-layer switches and as a Fibre Channel switch to the server attached to it.

Because Cisco Nexus 5000 Series and 5500 Platform switches do not provide Fibre Channel services, the switches do not need a Fibre Channel domain ID. As a result, the Cisco Nexus 5000 Series and 5500 Platform switches increase the scalability and eliminate the switch-to-switch interoperability issues. The switches also make management easier because the switches do not get involved in the FSPF and Fibre Channel policy.

The ports that connect to the upstream Fibre Channel switch need to support NPIV.

NPV Mode (Cont.)

- NPIV proxy captures all login-associated packets from the HBA or converged network adapter and external interfaces.
- Hosts are pinned to external interfaces.
- Supports NPIV over server interfaces.
- Relies on NPIV on external interfaces.
- Retries failed login requests from one external interface on a different interface.
- Handles events by generating proxy LOGOs.



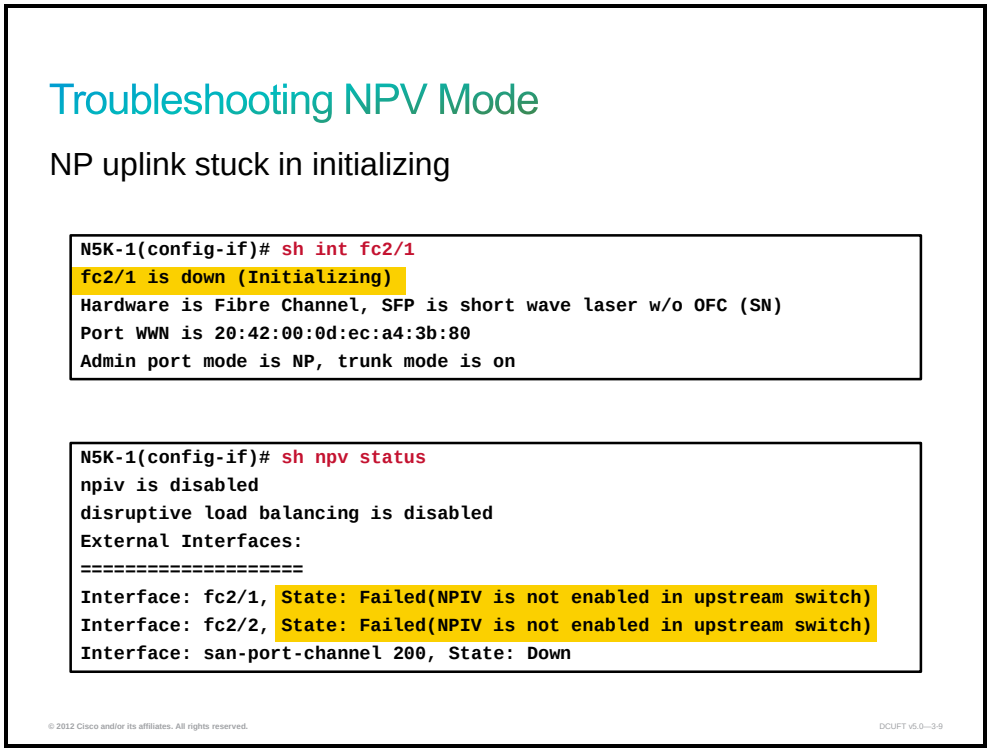
The NPIV proxy module in the Cisco NX-OS provides the proxy functionality of distributing FLOGI requests from the servers over the available border interfaces. The Fibre Channel HBAs in the servers and the upstream Fibre Channel switch act as though they are directly connected to each other through a physical cable. NPIV-proxy functionality allows NPIV to be used between the Cisco Nexus 5000 Series and 5500 Platform switches and other Fibre Channel switches. This use occurs even if some or all HBAs implement basic N-Port functionality.

The main responsibilities of the NPIV-proxy module include the following:

- Capture all the login-associated packets such as FLOGI, logout (LOGO), Link Service Accept (LS_ACC), and Link Service Reject (LS_RJT) from the servers and border interfaces that are enabled for N Port or NPIV functionality.
- Perform intelligent load balancing of login sessions coming from servers over the set of border interfaces so that logins are uniformly distributed.
- Support NPIV functionality over server interfaces to support HBAs that are NPIV-capable.
- Support NPIV on border interfaces to support Fibre Channel switches that support NPIV logins.
- If a login request on one border interface fails, provide the retry functionality for sending the login request on a different border interface.
- Perform events such as border interfaces, server interfaces, and VSANs going up and down, by generating proxy LOGOs.
- Provide high availability from process restart and switchovers.

Troubleshooting NPV Mode

This topic explains how to troubleshoot issues related to NPV mode on a Cisco Nexus or Cisco MDS Series switch.



If an uplink NP Port is stuck initializing, then a possible cause could be that the upstream switch is not enabled for NPIV.

To identify a port that is stuck initializing, use the **show interface fc** *x/y* command. Check the status of the NPV external interfaces. Use the **show npv status** command to determine whether NPIV is enabled on the core switch.

Troubleshooting NPV Mode (Cont.)

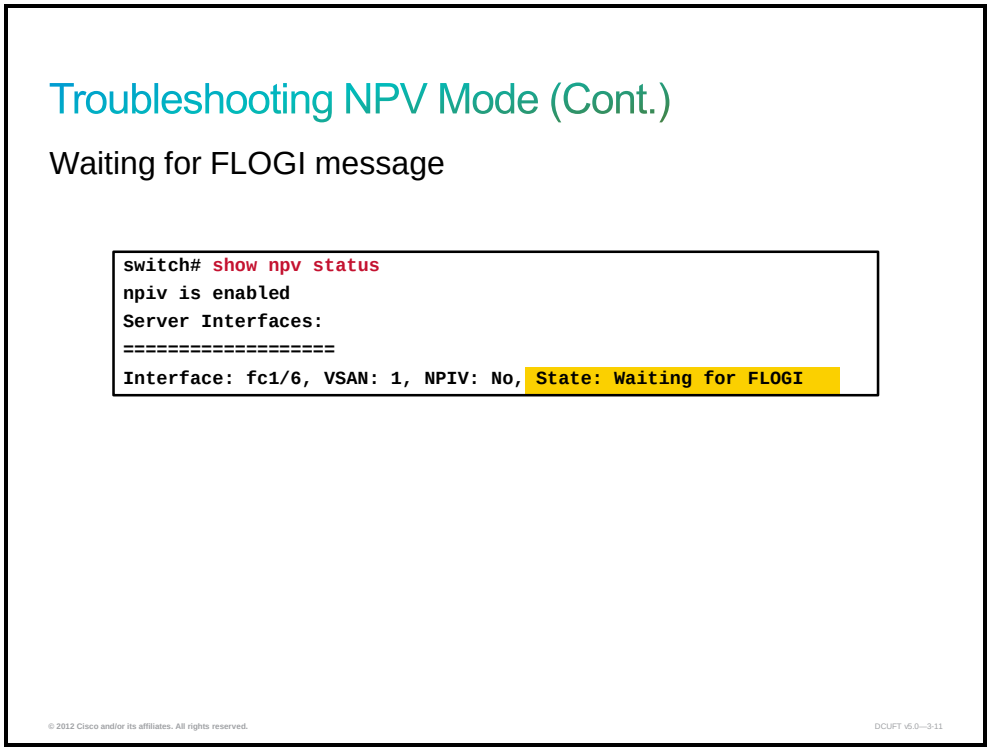
NPV upstream port unavailable

```
N5K-1# sh int fc2/7
fc2/7 is down (NPV upstream port not available)
Hardware is Fibre Channel, SFP is short wave laser w/o OFC (SN)
Port WWN is 20:47:00:0d:ec:a4:3b:80
Admin port mode is F, trunk mode is off
snmp link state traps are enabled
Port vsan is 99
Receive data field Si ze is 2112
```

```
N5K-1# show vsan membership
vsan 1 interfaces:
fc2/1 fc2/2 fc2/3 fc2/4
fc2/5 fc2/6 san-port-channel 200
vsan 99 interfaces:
fc2/7 fc2/8
```

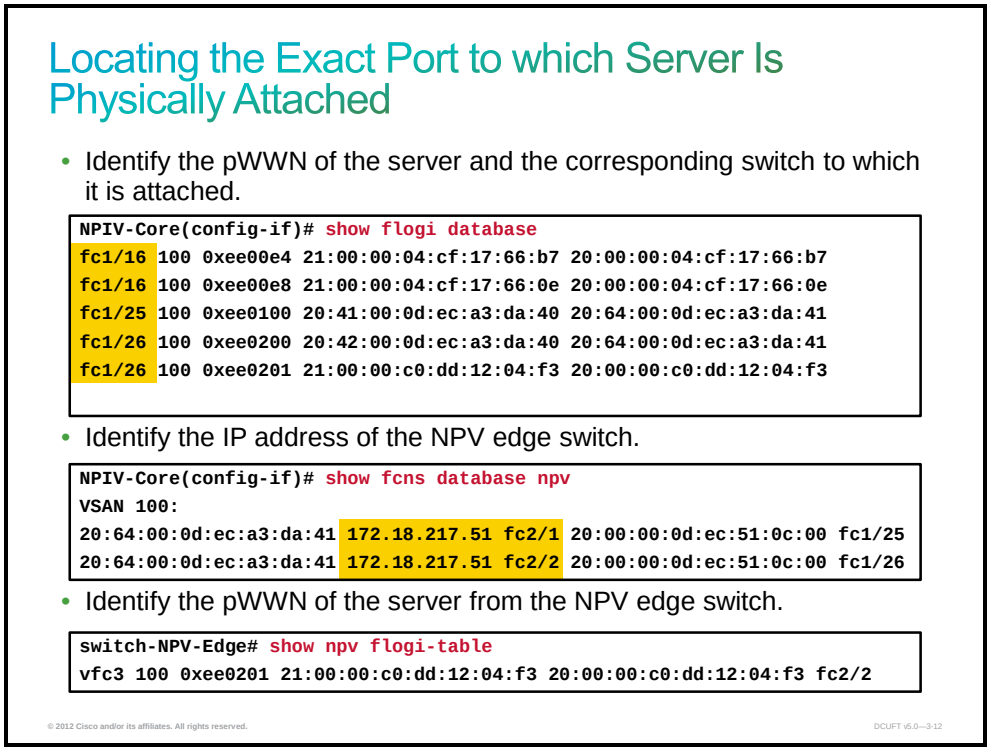

A server port that is connected to an NPV edge switch cannot come online if the status of the upstream port is “NPV upstream port not available.” If the VSAN does not match at each end of the connection, then this error can occur.

To identify whether an upstream port is in this state, use the **show interface fc** *x/y* command. If this state is shown, then use the **show vsan membership** command to verify the VSAN membership of the interfaces. Correct any incorrect VSAN memberships in the **vsan database** of the switch.



If the server on the downstream NPV edge switch does not log in to the fabric, or if you see the "waiting for FLOGI" message, then check the following:

- Verify the configuration of both the NPV edge and core switches. If you are not running the fabric port (F-Port) trunking feature, then verify that there are no VSAN mismatches. Verify that the server ports, NPV NP ports, NPIV core F Ports, and storage ports are all in the same VSAN and are online.
- If the configuration is correct, then collect an Ethanalyzer trace. Verify that the FLOGI frame is being received and sent to the NPIV core as an FDISC command.
- Perform **shutdown** and **no shutdown** commands on the NPV-attached server port to re-create the problem. The trace will be written to bootflash and can be copied off the switch to an external server.
- Verify the flow by using Wireshark.



NPIV switches lose visibility into the physical port to which a downstream NPV-connected server is attached. When you have an NPIV core switch that has several downstream NPV edge switches attached, you might want to locate the exact port to which a server is physically attached. Use these steps to locate the physical port:

Step 1 Identify the pWWN of the server and the corresponding switch to which it is attached; for example:

```
NPIV-Core(config-if)# show flogi database
fc1/16 100 0xee00e4 21:00:00:04:cf:17:66:b7 20:00:00:04:cf:17:66:b7
fc1/16 100 0xee00e8 21:00:00:04:cf:17:66:0e 20:00:00:04:cf:17:66:0e
fc1/25 100 0xee0100 20:41:00:0d:ec:a3:da:40 20:64:00:0d:ec:a3:da:41
fc1/26 100 0xee0200 20:42:00:0d:ec:a3:da:40 20:64:00:0d:ec:a3:da:41
fc1/26 100 0xee0201 21:00:00:c0:dd:12:04:f3 20:00:00:c0:dd:12:04:f3
```

In the example, the server is identified by this address:

```
fc1/26 100 0xee0201 21:00:00:c0:dd:12:04:f3 20:00:00:c0:dd:12:04:f3
```

The switch is identified by this address:

```
fc1/26 100 0xee0200 20:42:00:0d:ec:a3:da:40 20:64:00:0d:ec:a3:da:41
```

Step 2 Identify the IP address of the NPV edge switch; for example:

```
NPIV-Core(config-if)# show fcns database npv
VSAN 100:
20:64:00:0d:ec:a3:da:41 172.18.217.51 fc2/1 20:00:00:0d:ec:51:0c:00 fc1/25
20:64:00:0d:ec:a3:da:41 172.18.217.51 fc2/2 20:00:00:0d:ec:51:0c:00 fc1/26
```

Step 3 Connect by Telnet to the NPV edge switch; for example:

```
NPIV-Core(config-if)# telnet 172.18.217.51
```

Step 4 Identify the pWWN of the server for example:

```
switch-NPV-Edge# show npv flogi-table
vfc3 100 0xee0201 21:00:00:c0:dd:12:04:f3 20:00:00:c0:dd:12:04:f3 fc2/2
```

Step 5 If the interface is a Fibre Channel over Ethernet (FCoE) virtual Fibre Channel (vFC) interface as shown in the previous example, then use the **show interface vfc** *id* command to see to which port the vFC is physically bound.

Summary

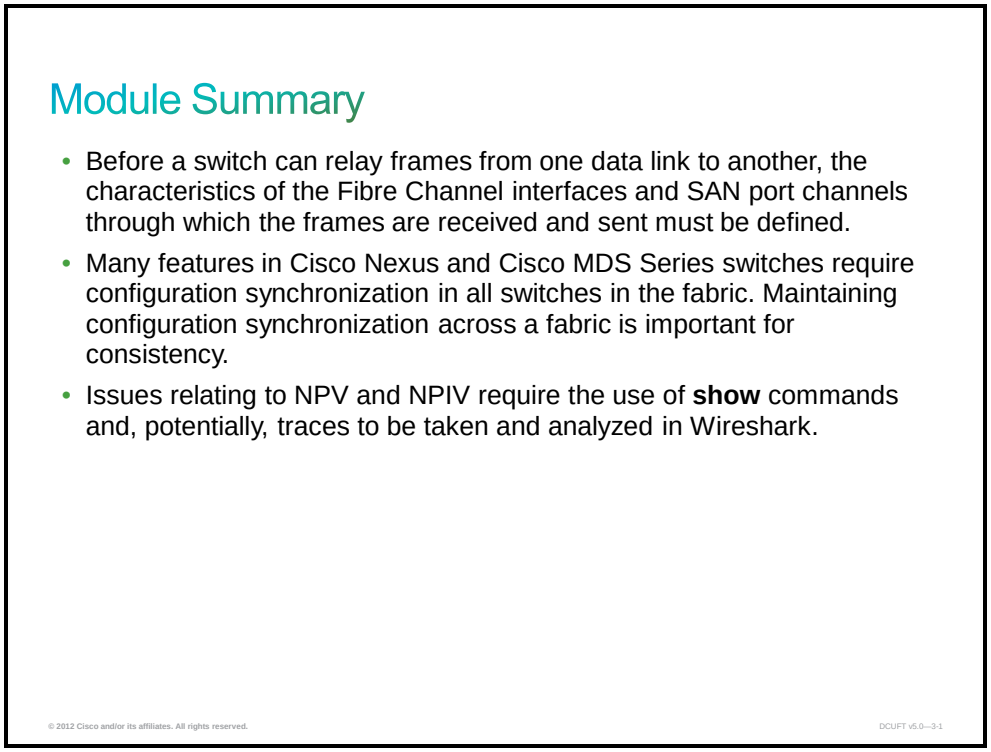
This topic summarizes the key points that were discussed in this lesson.

Summary

- When NPIV is enabled, multiple FLOGIs are accepted through a single F Port. NPV mode provides physical port-level virtualization of multiple Fibre Channel end nodes to one F Port off a Fibre Channel switch.
- Issues relating to NPV and NPIV require the use of **show** commands and, potentially, traces to be taken and analyzed in Wireshark.

Module Summary

This topic summarizes the key points that were discussed in this module.



Before a switch can relay frames from one data link to another, the characteristics of the interfaces through which the frames are received and sent must be defined. Each physical Fibre Channel interface in a switch can operate in one of several port modes: expansion port (E Port), fabric port (F Port), fabric loop port (FL Port), translative loop port (TL Port), trunking expansion port (TE Port), Switched Port Analyzer (SPAN) destination port (SD Port), and bridge port (B Port). In addition to these modes, each interface can be configured in auto or Fx Port mode. These modes determine the port type during interface initialization. SAN port channels refer to the aggregation of multiple physical Fibre Channel interfaces into one logical interface, to provide higher aggregated bandwidth, load balancing, and link redundancy.

VSANs provide isolation among devices that are physically connected to the same fabric. With VSANs, you can create multiple logical SANs over a common physical infrastructure. The Fibre Channel Domain (fcdomain) feature performs principal switch selection, domain ID distribution, Fibre Channel ID (FCID) allocation, and fabric reconfiguration functions. The name server stores name entries for all hosts in the Fibre Channel Name Server (FCNS) database. In a multiswitch fabric configuration, the name server instances running on each switch share information in a distributed database. Zoning enables you to set up access control between storage devices or user groups. Many features in Cisco Nexus and Cisco MDS Series switches require configuration synchronization in all switches in the fabric. Maintaining configuration synchronization across a fabric is important for consistency.

Node port (N-Port) ID Virtualization (NPIV) provides a means to assign multiple port IDs to one N Port. This feature allows multiple applications on the N Port to use different identifiers and allows access control, zoning, and port security to be implemented at the application level. N-Port Virtualization (NPV) makes use of NPIV to get multiple FCIDs allocated from the core switch on the NP Port. In NPV mode, the switch relays all traffic from server-side ports to the core switch. The core switch provides F Port functionality (such as login and port security) and all the Fibre Channel switching capabilities. The switch appears as a Fibre Channel host to the core switches, and as a regular Fibre Channel switch to its connected devices.

Module Self-Check

Use the questions here to review what you learned in this module. The correct answers and solutions are found in the Module Self-Check Answer Key.

Q1) Which four modes are possible Fibre Channel interface modes? (Choose four.) (Source: Troubleshooting Fibre Channel Interfaces)

A) S Port
B) F Port
C) TP Port
D) TE Port
E) NP Port
F) SD Port

Q2) Which states are possible Fibre Channel interface operational states? (Choose three.) (Source: Troubleshooting Fibre Channel Interfaces)

A) auto
B) up
C) trunking
D) down
E) active

Q3) When using Cisco Device Manager, what does the red box port display in Device View mean? (Source: Troubleshooting Fibre Channel Interfaces)

A) An SFP is not present.
B) The port is administratively disabled.
C) An SFP is present, but FLOGI has failed.
D) An SFP is present, but there is no connection.

Q4) Which command displays configuration information for all port channels? (Source: Troubleshooting Fibre Channel Interfaces)

A) **show interface port-channel**
B) **show port-channel database**
C) **show port-channel internal event-history msgs**
D) **channel-mode active**

Q5) What is the purpose of VSAN 4079? (Source: Troubleshooting Fibre Channel Fabric Service)

A) isolated VSAN
B) extended VSAN
C) EVFP VSAN
D) standard VSAN

Q6) Regarding Fibre Channel domain services, the **auto-reconfigure** option is enabled by default. (Source: Troubleshooting Fibre Channel Fabric Service)

A) true
B) false

Q7) Which two of these commands or tools can you use to recover from a zone merge failure in VSAN 10? (Choose two.) (Source: Troubleshooting Fibre Channel Fabric Service)

A) switch(config)# **zoneset import interface fc 1/3 vsan 10**
B) Fabric Manager > Zone > Merge Fail Recovery
C) Fabric Manager > Zone > Merge Analysis
D) switch(config)# **import zoneset interface fc 1/3 vsan 10**
E) switch# **zoneset import interface fc 1/3 vsan 10**
F) switch# **import zoneset interface fc 1/3 vsan 10**

Q8) Which command displays all of the switches in the physical fabric in terms of the switch WWN and the IP address? (Source: Troubleshooting Fibre Channel Fabric Service)

A) **show cfs application**
B) **show fcns details**
C) **show cfs regions**
D) **show cfs peers**

Q9) What must be enabled on the upstream switch if a Cisco Nexus 5000 Series Switch is in NPV mode? (Source: Troubleshooting NPV Mode)

A) NPV
B) NPIV
C) NPV and NPIV
D) trunking

Q10) Which mode relays the FLOGI and FDISC requests to an upstream Fibre Channel switch? (Source: Troubleshooting NPV Mode)

A) NPV
B) NPIV
C) switch
D) NPV and NPIV

Q11) Which type of port connects a Cisco Nexus 5000 Series Switch in NPV mode to an upstream Fibre Channel switch? (Source: Troubleshooting NPV Mode)

A) N Port
B) F Port
C) E Port
D) NP Port

Q12) What might indicate that an upstream switch is not enabled for NPIV? (Source: Troubleshooting NPV Mode)

A) port down
B) port stuck initializing
C) NPV upstream port not available
D) waiting for FLOGI message

Module Self-Check Answer Key

Q1) B, D, E, F

Q2) B, C, D

Q3) C

Q4) B

Q5) C

Q6) B

Q7) B, E

Q8) D

Q9) B

Q10) A

Q11) D

Q12) B