Cisco MDS 9000 Series NX-OS High Availability and Redundancy Configuration Guide

First Published: 2016-08-29

Americas Headquarters
Cisco Systems, Inc.
170 West Tasman Drive
San Jose, CA 95134-1706
USA
http://www.cisco.com
Tel: 408 526-4000
800 553-NETS (6387)
Fax: 408 527-0883

CONTENTS

Preface

This preface describes the audience, organization of, and conventions used in the Cisco MDS 9000 Series Configuration Guides. It also provides information on how to obtain related documentation, and contains the following chapters:

Audience

To use this installation guide, you need to be familiar with electronic circuitry and wiring practices, and preferably be an electronic or electromechanical technician.

Document Conventions

This document uses the following conventions:

Note Means *reader take note*. Notes contain helpful suggestions or references to material not covered in the manual.

Caution Means *reader be careful*. In this situation, you might do something that could result in equipment damage or loss of data.

Warnings use the following conventions:

Warning This warning symbol means danger. You are in a situation that could cause bodily injury. Before you work on any equipment, be aware of the hazards involved with electrical circuitry and be familiar with standard practices for preventing accidents. Use the statement number provided at the end of each warning to locate its translation in the translated safety warnings that accompanied this device. Statement 1071.

Related Documentation

The documentation set for the Cisco MDS 9000 Series Switches includes the following documents.

Release Notes

http://www.cisco.com/c/en/us/support/storage-networking/mds-9000-nx-os-san-os-software/products-release-notes-list.html

Regulatory Compliance and Safety Information

http://www.cisco.com/c/en/us/td/docs/switches/datacenter/mds9000/hw/regulatory/compliance/RCSI.html

Compatibility Information

http://www.cisco.com/c/en/us/support/storage-networking/mds-9000-nx-os-san-os-software/products-device-support-tables-list.html

Installation and Upgrade

http://www.cisco.com/c/en/us/support/storage-networking/mds-9000-nx-os-san-os-software/products-installation-guides-list.html

Configuration

http://www.cisco.com/c/en/us/support/storage-networking/mds-9000-nx-os-san-os-software/products-installation-and-configuration-guides-list.html

CLI

http://www.cisco.com/c/en/us/support/storage-networking/mds-9000-nx-os-san-os-software/products-command-reference-list.html

Troubleshooting and Reference

http://www.cisco.com/c/en/us/support/storage-networking/mds-9000-nx-os-san-os-software/tsd-products-support-troubleshoot-and-alerts.html

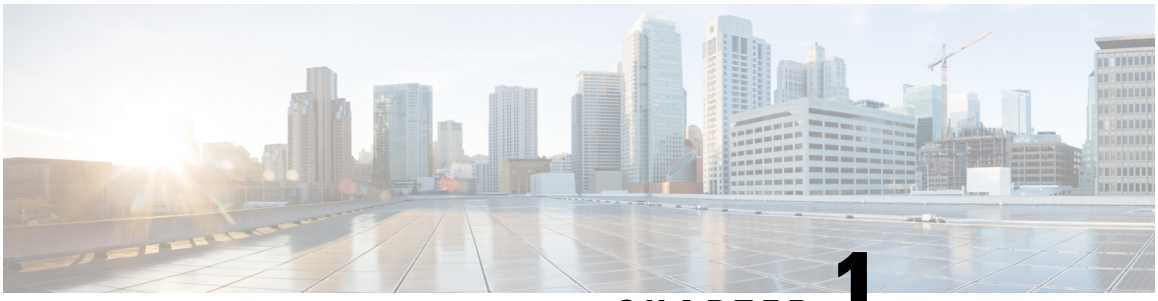
To find a document online, use the Cisco MDS NX-OS Documentation Locator at:

http://www.cisco.com/c/en/us/td/docs/storage/san_switches/mds9000/roadmaps/doclocater.html

Obtaining Documentation and Submitting a Service Request

For information on obtaining documentation, using the Cisco Bug Search Tool (BST), submitting a service request, and gathering additional information, see What's New in Cisco Product Documentation.

To receive new and revised Cisco technical content directly to your desktop, you can subscribe to the What's New in Cisco Product Documentation RSS feed. RSS feeds are a free service.



CHAPTER 1

New and Changed Information

As of Cisco MDS NX-OS Release 4.2(1), software configuration information is available in new feature-specific configuration guides for the following information:

- System management
- Interfaces
- Fabric
- Quality of service
- Security
- IP services
- High availability and redundancy

The information in these new guides previously existed in the *Cisco MDS 9000 Series CLI Configuration Guide* and in the *Cisco MDS 9000 Series Fabric Manager Configuration Guide* . Those configuration guides remain available on Cisco.com and should be used for all software releases prior to MDS NX-OS Release 4.2(1). Each guide addresses the features introduced in or available in a particular release. Select and view the configuration guide that pertains to the software installed in your switch.

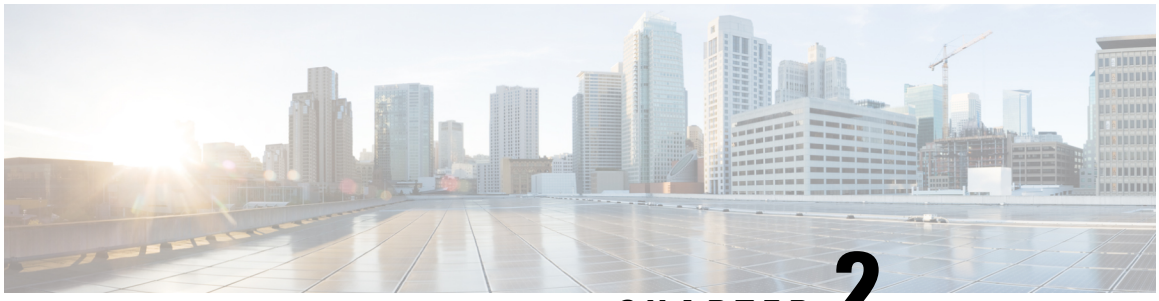
For a complete list of document titles, see the list of Related Documentation in the "Preface."

To find additional information about Cisco MDS NX-OS Release 4.2(x), see the *Cisco MDS 9000 Series Release Notes* available at the Cisco Systems website.

About this Guide

The information in the new *Cisco MDS 9000 NX-OS High Availability and Redundancy Configuration Guide* previously existed in Part 2: Installation and Switch Management of the *Cisco MDS 9000 Series CLI Configuration Guide.*

There are no new or changed CLI features for high availability and redundancy in MDS NX-OS Release 4.2(1).



CHAPTER 2

High Availability Overview

You can configure the high availability (HA) software framework and redundancy features using CLI. These features include application restartability and nondisruptive supervisor switchability. Cisco high availability is a technology delivered in Cisco NX-OS software that enables network-wide resilience to increase IP network availability.

The Cisco MDS 9500 Series of multilayer directors and switches support application restartability and nondisruptive supervisor switchability. The switches are protected from system failure by redundant hardware components and a high availability software framework.

The high availability (HA) software framework enables the following features:

- Ensures nondisruptive software upgrade capability.
- Provides redundancy for supervisor module failure by using dual supervisor modules.
- Performs nondisruptive restarts of a failed process on the same supervisor module. A service running on the supervisor modules and on the switching module tracks the HA policy defined in the configuration and takes action based on this policy. This feature is also available in switches in the Cisco MDS 9200 Series and the Cisco MDS 9100 Series.
- Protects against link failure using the PortChannel (port aggregation) feature. This feature is also available in switches in the Cisco MDS 9200 Series and in the Cisco MDS 9100 Series.
- Provides management redundancy using the Virtual Router Redundancy Protocol (VRRP). This feature is also available in switches in the Cisco MDS 9200 Series and in the Cisco MDS 9100 Series.
- Provides switchovers if the active supervisor fails. The standby supervisor, if present, takes over without disrupting storage or host traffic.
- Ensures Internal Cyclic Redundancy Check (CRC) detection and isolation on the Cisco MDS 9700 series switches.

Directors in the Cisco MDS 9500 Series have two supervisor modules (Supervisor-1 and Supervisor-2) in slots 5 and 6 (Cisco MDS 9509 and 9506 Switches) or slots 7 and 8 (Cisco MDS 9513 Switch). When the switch powers up and both supervisor modules are present, the supervisor module that comes up first enters the active mode, and the supervisor module that comes up second enters the standby mode. If both supervisor modules come up at the same time, Supervisor-1 becomes active. The standby supervisor module constantly monitors the active supervisor module. If the active supervisor module fails, the standby supervisor module takes over without any impact to user traffic.

Note For high availability, you need to connect the Ethernet port for both active and standby supervisors to the same network or virtual LAN. The active supervisor owns the one IP address used by these Ethernet connections. On a switchover, the newly activated supervisor takes over this IP address.

Internal CRC Detection and Isolation

Beginning with the Cisco MDS NX-OS Release 6.2(13), the Internal Cyclic Redundancy Check (CRC) detection and isolation functionality is supported on the Cisco MDS 9700 Series switches.

This functionality enables the Cisco MDS switches to detect CRC errors that occur internally within a switch and isolate the source of these errors.

Note Internal CRC Detection and Isolation is supported only on the Cisco MDS 9700 Series Multilayer Directors.

By default, the internal CRC detection and isolation is disabled.

The modules that support this functionality are:

- Cisco MDS 9700 48-Port 16-Gbps Fibre Channel Switching Module
- Cisco MDS 9700 48-Port 10-Gbps Fibre Channel over Ethernet Switching Module
- Cisco MDS 9700 Fabric Module 1
- Cisco MDS 9700 Supervisor Module 3

Note *Module* refers either a switching module or a supervisor module.

These errors are a separate class of CRC errors when compared to frames that arrive from outside the switch, with CRC errors. In store mode and forward mode, frames with CRC errors are dropped at the ingress port and do not propagate through the system. Internal CRC errors occur when frames are received without errors, but get corrupted when they pass through the switching path.

Internal CRC errors are usually caused by a fault in the system. Such faults may be transient, such as an ungracefully removed module, or permanent, such as a badly seated module, or, in rare cases, a failing or failed hardware component. The rate of errors depends on many factors and may range from very high to very low.

The error-rate threshold is configurable as a system-wide value, but separate error counts are maintained for each module to identify an error source.

Note The counters are reset at 24 hours from the time the feature, the Internal Cyclic Redundancy Check (CRC) detection and isolation was first configured.

Stages of Internal CRC Detection and Isolation

The five possible stages at which internal CRC errors may occur in a switch:

Stage 1—Ingress buffer of a module

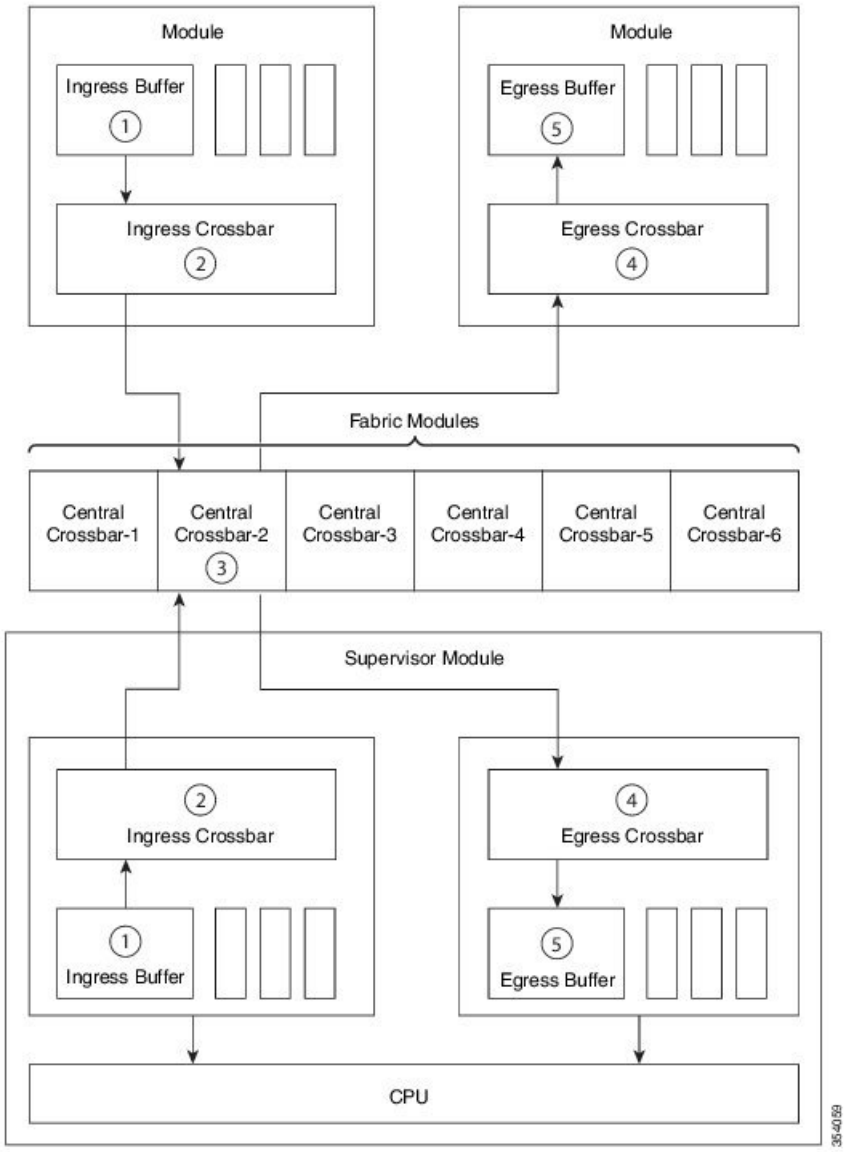
Stage 2—Ingress crossbar of a module

Stage 3—Crossbar of a fabric module

Stage 4—Egress crossbar of a module

Stage 5—Egress buffer of a module

Figure 1: Stages of Internal CRC Detection and Isolation



Errors on each module are handled individually when the error count exceeds the threshold.

Note A total of errors on all applicable ASICs on the module must exceed the threshold.

When errors cross the specified threshold, XBAR_MONITOR_INTERNAL_CRC_ERR is the syslog message that is logged. This syslog message specifies the location of the error and the type of action taken.

Example: Error Messages

```
switch# show logging logfile | inc MONITOR_INTERNAL_CRC_ERR
2015 May 25 21:20:41 switch %XBAR-2-XBAR_MONITOR_INTERNAL_CRC_ERR: Module-1 detects CRC
Error:4 at Egress Q-engine, putting it in failure state
2015 May 25 21:15:35 switch %XBAR-2-XBAR_MONITOR_INTERNAL_CRC_ERR: Fab_slot-12 detects CRC
 error:1 at ingress stage2, putting it in failure state
2015 May 25 15:47:10 switch %XBAR-2-XBAR_MONITOR_INTERNAL_CRC_ERR: Module-5 detects CRC
error:2 at Ingress Qengine, Only one Sup is present, bringing down the active VSAN
2015 May 25 15:08:17 switch %XBAR-2-XBAR_MONITOR_INTERNAL_CRC_ERR: Module-5 detects CRC
error:1 at Ingress Qengine, putting it in failure state
```

Stage 1—Ingress Buffer of a Module

There are multiple ingress buffers on each module. When the CRC error rate of an ingress buffer on a switching module reaches the threshold, the entire module is shut down. See Actions Taken on a Supervisor when the Threshold Exceeded, on page 6 for more information.

Stage 2—Ingress Crossbar of a Module

Ingress crossbar is an ASIC complex on an ingress module that switches traffic from ingress buffers to fabric modules. When the CRC error rate of an ingress switching module crossbar reaches the threshold, the entire module is shut down. See Actions Taken on a Supervisor when the Threshold Exceeded, on page 6 for more information.

Stage 3—Crossbar of a Fabric Module

Crossbar is an ASIC complex on a fabric module that switches traffic from an ingress module to an egress module.

When the CRC error rate of a crossbar reaches the threshold, if there is more than one fabric module in the corresponding switch, the host fabric module is shut down. If the switch has only one fabric module, the module connected to the fabric module link on which the errors occurred is shut down.

Stage 4—Egress Crossbar of a Module

Egress crossbar is an ASIC complex on an egress module that switches traffic from fabric modules to egress buffers. When the CRC error rate of an egress switching module crossbar reaches the threshold, the entire module is shut down. See Actions Taken on a Supervisor when the Threshold Exceeded, on page 6 for more information.

Stage 5—Egress Buffer of a Module

There are multiple egress buffers on each module. When the CRC error rate of an egress buffer on a switching module reaches the threshold, the entire module is shut down. See Actions Taken on a Supervisor when the Threshold Exceeded, on page 6 for more information.

Actions Taken on a Supervisor when the Threshold Exceeded

The actions taken on a supervisor when the threshold is exceeded during the following stages of internal CRC detection and isolation:

Stage 1—Ingress Buffer of a Module

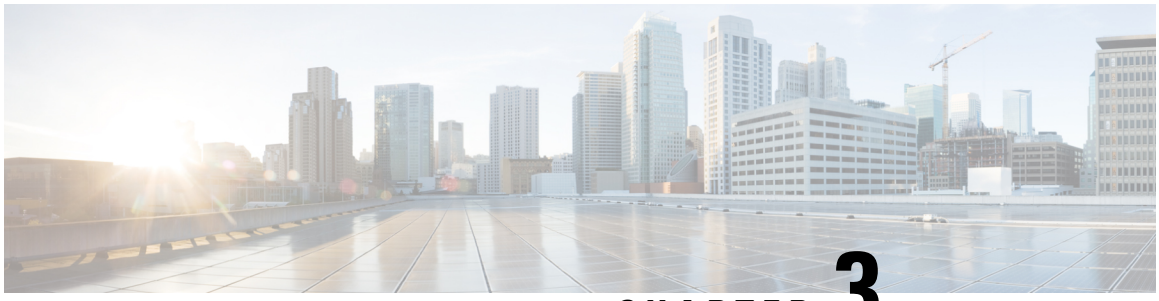
Stage 2— Ingress Crossbar of a Module

Stage 3—Egress Crossbar of a Module

Stage 5—Egress Buffer of a Module

Note

- When both active and standby supervisors are present in the switch, the active supervisor is brought down and the standby takes over.
- When a single fabric module is present and Stage 2 error occurs, the line card connected to the fabric module is powered down; as a result the switch is brought down. This mechanism helps in isolating the faulty spine port or link as the line card connected to the spine which experienced the error is brought down.
- For information on configuring the Internal CRC Detection and Isolation feature, see "Configuring Internal CRC Detection and Isolation, on page 11."



CHAPTER 3

Configuring High Availability

This chapter describes how to configure high availability, and describes the switchover processes.

About High Availability

Process restartability provides the high availability functionality in Cisco MDS 9000 Series switches. This process ensures that process-level failures do not cause system-level failures. It also restarts the failed processes automatically. This process is able to restore its state prior to the failure and continues executing from the failure point going forward.

An HA switchover has the following characteristics:

- It is stateful (nondisruptive) because control traffic is not impacted.
- It does not disrupt data traffic because the switching modules are not impacted.
- Switching modules are not reset.

Note Switchover is not allowed if auto-copy is in progress.

Switchover Processes

Switchovers occur by one of the following two processes:

- The active supervisor module fails and the standby supervisor module automatically takes over.
- You manually initiate a switchover from an active supervisor module to a standby supervisor module.

Once a switchover process has started another switchover process cannot be started on the same switch until a stable standby supervisor module is available.

Caution If the standby supervisor module is not in a stable state (ha-standby), a switchover is not performed.

Synchronizing Supervisor Modules

The running image is automatically synchronized in the standby supervisor module by the active supervisor module. The boot variables are synchronized during this process.

The standby supervisor module automatically synchronizes its image with the running image on the active supervisor module.

Note The image a supervisor module is booted up from cannot be deleted from bootflash. This is to ensure that the new standby supervisor module ia able to synchronize during the process.

Manual Switchover Guidelines

Be aware of the following guidelines when performing a manual switchover:

- When you manually initiate a switchover, system messages indicate the presence of two supervisor modules.
- A switchover can only be performed when two supervisor modules are functioning in the switch.
- The modules in the chassis are functioning as designed.

Manually Initiating a Switchover

To manually initiate a switchover from an active supervisor module to a standby supervisor module, use the active supervisor module using Device Manager**system switchover** command. After you enter this command, another switchover process cannot be started on the same switch until a stable standby supervisor module is available.

To ensure that an HA switchover is possible, enter the **show system redundancy status** command or the **show module** command. If the command output displays the HA standby state for the standby supervisor module, then the switchover is possible. See "Verifying Switchover Possibilities, on page 10" for more information.

Verifying Switchover Possibilities

This section describes how to verify the status of the switch and the modules before a manual switchover.

- Use the s**how system redundancy status** command to ensure that the system is ready to accept a switchover.
- Use the **show module** command to verify the status (and presence) of a module at any time. A sample output of the **show module** command follows:

```
switch# show module
Mod  Ports  Module-Type                       Model              Status
---  -----  -------------------------------- ------------------ ------------
2    8      IP Storage Services Module       DS-X9308-SMIP      ok
```

```
5    0      Supervisor/Fabric-1             DS-X9530-SF1-K9    active *
6    0      Supervisor/Fabric-1             DS-X9530-SF1-K9    ha-standby
8    0      Caching Services Module         DS-X9560-SMAP      ok
9    32     1/2 Gbps FC Module              DS-X9032           ok
Mod  MAC-Address(es)                         Serial-Num
---  --------------------------------------  ----------
2    00-05-30-00-9d-d2 to 00-05-30-00-9d-de  JAB064605a2
5    00-05-30-00-64-be to 00-05-30-00-64-c2  JAB06350B1R
6    00-d0-97-38-b3-f9 to 00-d0-97-38-b3-fd  JAB06350B1R
8    00-05-30-01-37-7a to 00-05-30-01-37-fe  JAB072705ja
9    00-05-30-00-2d-e2 to 00-05-30-00-2d-e6  JAB06280ae9
* this terminal session
```

The Status column in the output should display an OK status for switching modules and an active or HA-standby status for supervisor modules. If the status is either OK or active, you can continue with your configuration.

- Use the **show boot auto-copy** command to verify the configuration of the auto-copy feature and if an auto-copy to the standby supervisor module is in progress. Sample outputs of the **show boot auto-copy** command follow:

```
switch# show boot auto-copy
Auto-copy feature is enabled
switch# show boot auto-copy list
No file currently being auto-copied
```

Configuring Internal CRC Detection and Isolation

Note

- This functionality is disabled by default.
- Cisco MDS 9500 Series Multilayer Directors do not support the Internal CRC Detection and Isolation feature.

To configure internal CRC detection and isolation, perform these steps:

Step 1 Enter configuration mode:

switch# **configure terminal**

Step 2 Enable internal CRC detection and isolation:

switch(config)# **hardware fabric crc** [**threshold** *count*]

The error rate is measured over sequential 24-hour window. The range of threshold is 1 to 100. If the threshold is not specified, the default is 3.

(Optional) Disable internal CRC detection and isolation:

switch(config)# **no hardware fabric crc**

(Optional) Save the configuration change:

switch(config)# **copy running-config startup-config**

Default Settings for Internal CRC Detection and Isolation

The table below lists the default settings for interface parameters.

Table 1: Default Settings for Internal CRC Detection and Isolation

Parameters	Default
Internal CRC Error Handling	Disabled

Copying Boot Variable Images to the Standby Supervisor Module

You can copy the boot variable images that are in the active supervisor module (but not in the standby supervisor module) to the standby supervisor module. Only those KICKSTART and SYSTEM boot variables that are set for the standby supervisor module can be copied. For module (line card) images, all boot variables are copied to the corresponding standby locations (bootflash: or slot0:) if not already present.

Enabling Automatic Copying of Boot Variables

To enable or disable automatic copying of boot variables, follow these steps:

SUMMARY STEPS

1. switch# **config t**
2. switch(config)# **boot auto-copy**
3. switch(config)# no **boot auto-copy**

DETAILED STEPS

	Command or Action	Purpose
Step 1	switch# **config t** **Example:** `switch(config)#`	Enters configuration mode.
Step 2	switch(config)# **boot auto-copy** **Example:** `Auto-copy administratively enabled`	Enables (default) automatic copying of boot variables from the active supervisor module to the standby supervisor module.
Step 3	switch(config)# no **boot auto-copy** **Example:** `Auto-copy administratively disabled`	Disables the automatic copy feature.

Verifying the Copied Boot Variables

Use the **show boot auto-copy** command to verify the current state of the copied boot variables. This example output shows that automatic copying is enabled:

```
switch# show boot auto-copy
Auto-copy feature enabled
```

This example output shows that automatic copying is disabled:

```
switch# show boot auto-copy
Auto-copy feature disabled
```

Use the **show boot auto-copy list** command to verify what files are being copied. This example output displays the image being copied to the standby supervisor module's bootflash. Once this is successful, the next file will be image2.bin.

Note This command only displays files on the active supervisor module.

```
switch# show boot auto-copy list
File: /bootflash:/image1.bin
Bootvar: kickstart
File:/bootflash:/image2.bin
Bootvar: system
```

This example output displays a typical message when the **auto-copy** option is disabled or if no files are copied:

```
switch# show boot auto-copy list
No file currently being auto-copied
```

Displaying HA Status Information

Use the **show system redundancy status** command to view the HA status of the system. Tables Redundancy States to Internal States explain the possible output values for the redundancy, supervisor, and internal states.

```
switch# show system redundancy status
Redundancy mode
---------------
      administrative:   HA
         operational:   HA
This supervisor (sup-1)
-----------------------
    Redundancy state:   Active
    Supervisor state:   Active
      Internal state:   Active with HA standby
Other supervisor (sup-2)
------------------------
    Redundancy state:   Standby
    Supervisor state:   HA standby
      Internal state:   HA standby
```

The following conditions identify when automatic synchronization is possible:

- If the internal state of one supervisor module is Active with HA standby and the other supervisor module is HA standby, the switch is operationally HA and can do automatic synchronization.
- If the internal state of one of the supervisor modules is none, the switch cannot do automatic synchronization.

Following table lists the possible values for the redundancy states.

Following table lists the possible values for the supervisor module states.

Following table lists the possible values for the internal redundancy states.

Displaying the System Uptime

The system uptime refers to the time that the chassis was powered on and has at least one supervisor module controlling the switch. Use the **reset** command to reinitialize the system uptime. Nondisruptive upgrades and switchovers do not reinitialize the system uptime, which means that the system uptime is contiguous across such upgrades and switchovers.

The kernel uptime refers to the time since the NX-OS software was loaded on the supervisor module. Use the **reset** and **reload** commands to reinitialize the kernel uptime.

The active supervisor uptime refers to the time since the NX-OS software was loaded on the active supervisor module. The active supervisor uptime can be lower than the kernel uptime after nondisruptive switchovers.

You can use the **show system uptime** command to view the start time of the system, uptime of the kernel, and the active supervisor.

This example shows how to display the supervisor uptime:

```
switch# show system uptime
System start time:          Fri Aug 27 09:00:02 2004
System uptime:              1546 days, 2 hours, 59 minutes, 9 seconds
Kernel uptime:              117 days, 1 hours, 22 minutes, 40 seconds
Active supervisor uptime:   117 days, 0 hours, 30 minutes, 32 seconds
```

For more information on high availability, see chapter 1, High Availability Overview.