CONTENTS

Preface vii

Audience vii
Document Conventions vii
Related Documentation for Cisco Nexus 9000 Series Switches viii
Documentation Feedback viii
Obtaining Documentation and Submitting a Service Request ix

CHAPTER 1 New and Changed Information 1

New and Changed Information 1

CHAPTER 2 Overview 3

About High Availability 3
Service-Level High Availability 4
Isolation of Processes 4
Process Restartability 4
System-Level High Availability 4
Physical Redundancy 4
Network-Level High Availability 5
Layer 2 HA Features 5
Layer 3 HA Features 5
Additional Management Tools for Availability 6
EEM 6
Smart Call Home 6
Software Image 6
Virtual Device Contexts 7

CHAPTER 3 Service-Level High Availability 9
## About Cisco NX-OS Service Restarts

Licensing Requirements  
Restartability Infrastructure  
System Manager  
Persistent Storage Service  
Message and Transaction Service  
HA Policies  
Process Restartability  
Types of Process Restarts  
Stateful Restarts  
Stateless Restarts  
Switchovers  
Restarts on Standby Supervisor Services  
Restarts on Switching Module Services  
Troubleshooting Restarts  
Additional References for Service-Level High Availability  
Related Documents  
MIBs  

## Network-Level High Availability

About Network-Level High Availability  
Licensing Requirements  
Spanning Tree Protocol  
Virtual Port Channels  
First-Hop Redundancy Protocols  
Nonstop Forwarding in Routing Protocols  
Additional References for Network-Level High Availability  
Related Documents  
MIBs  

## System-Level High Availability

About Cisco NX-OS System-Level High Availability  
Licensing Requirements  
Physical Redundancy  
Power Supply Redundancy
Power Modes 21
Fan Tray Redundancy 22
Switch Fabric Redundancy 22
Line Card Failures 23
Fabric Module Operating Modes 23
System Controller Redundancy 24
Supervisor Module Redundancy 24
Supervisor Modules 24
Configuring the Operating Mode for Fabric Modules 25
Supervisor Restarts and Switchovers 26
Restarts on Single Supervisors 26
Restarts on Dual Supervisors 26
Switchovers on Dual Supervisors 26
Switchover Characteristics 26
Switchover Mechanisms 26
Switchover Failures 27
Manually Initiating a Switchover 27
Switchover Guidelines 27
Verifying Switchover Possibilities 27
Replacing the Active Supervisor Module in a Dual Supervisor System 29
Replacing the Standby Supervisor Module in a Dual Supervisor System 30
Displaying HA Status Information 31
Additional References for System-Level High Availability 33
Related Documents 34
MIBs 34
Preface

This preface includes the following sections:

- **Audience**, page vii
- **Document Conventions**, page vii
- **Related Documentation for Cisco Nexus 9000 Series Switches**, page viii
- **Documentation Feedback**, page viii
- **Obtaining Documentation and Submitting a Service Request**, page ix

Audience

This publication is for network administrators who install, configure, and maintain Cisco Nexus switches.

Document Conventions

Command descriptions use the following conventions:

<table>
<thead>
<tr>
<th>Convention</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>bold</strong></td>
<td>Bold text indicates the commands and keywords that you enter literally as shown.</td>
</tr>
<tr>
<td><em>Italic</em></td>
<td>Italic text indicates arguments for which the user supplies the values.</td>
</tr>
<tr>
<td>[x]</td>
<td>Square brackets enclose an optional element (keyword or argument).</td>
</tr>
<tr>
<td>[x</td>
<td>y]</td>
</tr>
<tr>
<td>{x</td>
<td>y}</td>
</tr>
</tbody>
</table>
### Related Documentation for Cisco Nexus 9000 Series Switches

The entire Cisco Nexus 9000 Series switch documentation set is available at the following URL:


### Documentation Feedback

To provide technical feedback on this document, or to report an error or omission, please send your comments to nexus9k-docfeedback@cisco.com. We appreciate your feedback.
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 at: http://www.cisco.com/c/en/us/td/docs/general/whatsnew/whatsnew.html.

Subscribe to What's New in Cisco Product Documentation, which lists all new and revised Cisco technical documentation as an RSS feed and delivers content directly to your desktop using a reader application. The RSS feeds are a free service.
New and Changed Information

This chapter provides release-specific information for each new and changed feature in the Cisco Nexus 9000 Series NX-OS High Availability and Redundancy Guide, Release 6.x.

- New and Changed Information, page 1

New and Changed Information

This table summarizes the new and changed features for the Cisco Nexus 9000 Series NX-OS High Availability and Redundancy Guide, Release 6.x and tells you where they are documented.

Table 1: New and Changed Features for Cisco NX-OS Release 6.x

<table>
<thead>
<tr>
<th>Feature</th>
<th>Description</th>
<th>Changed in Release</th>
<th>Where Documented</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisco Nexus 9332PQ, 9372PX, 9372TX, and 9396TX chassis</td>
<td>Introduced these devices.</td>
<td>6.1(2)I3(1)</td>
<td>System-Level High Availability, on page 19</td>
</tr>
<tr>
<td>Supervisor B</td>
<td>Introduced this module.</td>
<td>6.1(2)I3(1)</td>
<td>System-Level High Availability, on page 19</td>
</tr>
<tr>
<td>Line cards</td>
<td>Added the ability to keep line cards powered down whenever they fail or crash.</td>
<td>6.1(2)I3(1)</td>
<td>System-Level High Availability, on page 19</td>
</tr>
<tr>
<td>Cisco Nexus 9516 chassis</td>
<td>Introduced this device.</td>
<td>6.1(2)I2(2a)</td>
<td>System-Level High Availability, on page 19</td>
</tr>
<tr>
<td>X9464PX, X9464TX, X9432PQ, and X9536PQ</td>
<td>Introduced these line cards.</td>
<td>6.1(2)I2(2a)</td>
<td>System-Level High Availability, on page 19</td>
</tr>
<tr>
<td>Feature</td>
<td>Description</td>
<td>Changed in Release</td>
<td>Where Documented</td>
</tr>
<tr>
<td>----------------------------------------</td>
<td>-----------------------------------------------------------------------------</td>
<td>--------------------</td>
<td>-------------------------------------------------------</td>
</tr>
<tr>
<td>Operating modes for fabric modules</td>
<td>Added the ability to configure the system fabric in full-rate mode (FRM) or oversubscribed mode (OSM).</td>
<td>6.1(2)I2(2a)</td>
<td>System-Level High Availability, on page 19</td>
</tr>
<tr>
<td>Cisco Nexus 9504, 9396PX, and 93128TX chassis</td>
<td>Introduced these devices.</td>
<td>6.1(2)I2(1)</td>
<td>System-Level High Availability, on page 19</td>
</tr>
<tr>
<td>X9564PX and X9564TX</td>
<td>Introduced these line cards.</td>
<td>6.1(2)I2(1)</td>
<td>System-Level High Availability, on page 19</td>
</tr>
<tr>
<td>HSRP</td>
<td>Introduced this feature.</td>
<td>6.1(2)I2(1)</td>
<td>Network-Level High Availability, on page 15</td>
</tr>
<tr>
<td>Spanning Tree Protocol</td>
<td>Introduced this feature.</td>
<td>6.1(2)I2(1)</td>
<td>Network-Level High Availability, on page 15</td>
</tr>
<tr>
<td>vPCs</td>
<td>Added support for virtual port channels.</td>
<td>6.1(2)I2(1)</td>
<td>Network-Level High Availability, on page 15</td>
</tr>
<tr>
<td>VRRP</td>
<td>Introduced this feature.</td>
<td>6.1(2)I2(1)</td>
<td>Network-Level High Availability, on page 15</td>
</tr>
<tr>
<td>Cisco Nexus 9508 chassis</td>
<td>Introduced this device.</td>
<td>6.1(2)I1(1)</td>
<td>System-Level High Availability, on page 19</td>
</tr>
<tr>
<td>X9636PQ</td>
<td>Introduced this line card.</td>
<td>6.1(2)I1(1)</td>
<td>System-Level High Availability, on page 19</td>
</tr>
</tbody>
</table>
Overview

Cisco NX-OS is a resilient operating system that is specifically designed for high availability at the network, system, and process level.

This chapter describes high-availability (HA) concepts and features for Cisco NX-OS devices and includes the following sections:

- About High Availability, page 3
- Service-Level High Availability, page 4
- System-Level High Availability, page 4
- Network-Level High Availability, page 5
- Additional Management Tools for Availability, page 6
- Software Image, page 6
- Virtual Device Contexts, page 7

About High Availability

To prevent or minimize traffic disruption during hardware or software failures, Cisco NX-OS has these features:

- Redundancy—Cisco NX-OS HA provides physical and software redundancy at every component level, spanning across the physical, environmental, power, and system software aspects of its architecture.

- Isolation of planes and processes—Cisco NX-OS HA provides isolation between control and data forwarding planes within the device and between software components so that a failure within one plane or process does not disrupt others.

- Restartability—Most system functions and services are isolated so that they can be restarted independently after a failure while other services continue to run. In addition, most system services can perform stateful restarts, which allow the service to resume operations transparently to other services.

- Supervisor stateful switchover—The Cisco Nexus 9504, 9508, and 9516 chassis support an active and standby dual supervisor configuration. State and configuration remain constantly synchronized between the two supervisor modules to provide seamless and stateful switchover in the event of a supervisor module failure.
Service-Level High Availability

Cisco NX-OS has a modularized architecture that compartmentalizes components for fault isolation, redundancy, and resource efficiency.

Isolation of Processes

In the Cisco NX-OS software, independent processes, known as services, perform a function or set of functions for a subsystem or feature set. Each service and service instance runs as an independent, protected process. This approach provides a highly fault-tolerant software infrastructure and fault isolation between services. A failure in a service instance (such as BGP) does not affect any other services that are running at that time (such as the Link Aggregation Control Protocol [LACP]). In addition, each instance of a service can run as an independent process, which means that two instances of a routing protocol (for example, two instances of the Open Shortest Path First [OSPF] protocol) can run as separate processes.

Process Restartability

Cisco NX-OS processes run in a protected memory space independently from each other and the kernel. This process isolation provides fault containment and enables rapid restarts. Process restartability ensures that process-level failures do not cause system-level failures. In addition, most services can perform stateful restarts, which allow a service that experiences a failure to be restarted and to resume operations transparently to other services within the platform and to neighboring devices within the network.

System-Level High Availability

The Cisco Nexus 9000 Series switches are protected from system failure by redundant hardware components and a high-availability software framework.

Physical Redundancy

The Cisco Nexus 9000 Series switches have the following physical redundancies:

- Power Supply Redundancy—To provide redundant power input to the chassis, the Cisco Nexus 9504 chassis supports up to four power supply modules, the Cisco Nexus 9508 chassis supports up to eight power supply modules, the Cisco Nexus 9516 chassis supports up to ten power supply modules, and the Cisco Nexus 9332, 9372, 9396, and 93128 chassis support up to two power supply modules.

- Fan Tray Redundancy—For cooling the system, the Cisco Nexus 9500 Series chassis and the Cisco Nexus 9396 and 93128 chassis each contain three fan trays. The Cisco Nexus 9332 and 9372 chassis each contain four fan trays.

- Fabric Redundancy—Cisco NX-OS provides switching fabric availability through redundant switch fabric modules. You can configure a single Cisco Nexus 9500 Series chassis with one to six switch fabric cards for capacity and redundancy.
The Cisco Nexus 9300 Series chassis do not have fabric modules.

- System Controller Redundancy—A pair of redundant system controllers in the Cisco Nexus 9500 Series chassis offloads chassis management functions from the supervisor modules.

Note

The Cisco Nexus 9300 Series chassis do not contain system controllers.

- Supervisor Module Redundancy—The Cisco Nexus 9500 Series chassis support dual supervisor modules to provide redundancy for the control and management plane.

Note

The Cisco Nexus 9300 Series chassis do not support supervisor module redundancy.

Network-Level High Availability

Network convergence is optimized by providing tools and functions to make both failover and fallback transparent and fast.

Layer 2 HA Features

Cisco NX-OS provides these Layer 2 HA features:

- Spanning Tree Protocol (STP) enhancements, such as Bridge Protocol Data Unit (BPDU) Guard, Loop Guard, Root Guard, BPDU Filters, and Bridge Assurance, to guarantee the health of the STP control plane
- Unidirectional Link Detection (UDLD) Protocol
- IEEE 802.3ad link aggregation

Note

Virtual port channels (vPCs) allow you to create redundant physical links between two systems that act as a logical single link.

Layer 3 HA Features

Cisco NX-OS provides these Layer 3 HA features:

- Nonstop forwarding (NSF) graceful restart extensions for routing protocols
  
  Open Shortest Path First version 2 (OSPFv2), OSPFv3, Intermediate System to Intermediate System (IS-IS), Enhanced Interior Gateway Routing Protocol (EIGRP), and Border Gateway Protocol (BGP)
utilize graceful restart extensions to the base protocols to provide nonstop forwarding and least obtrusive routing recovery for those environments.

- Shortest Path First (SPF) optimizations such as link-state advertisement (LSA) pacing and incremental SPF
- Protocol-based periodic refresh
- Millisecond timers for First-Hop Redundancy Protocols (FHRPs) such as the Hot Standby Router Protocol (HSRP) and the Virtual Router Redundancy Protocol (VRRP)

### Additional Management Tools for Availability

Cisco Nexus 9000 Series NX-OS incorporates several Cisco system management tools for monitoring and notification of system availability events.

#### EEM

Cisco Embedded Event Manager (EEM) consists of Event Detectors, the Event Manager, and an Event Manager Policy Engine. Using EEM, you can define policies to take specific actions when the system software recognizes certain events through the Event Detectors. The result is a flexible set of tools to automate many network management tasks and to direct the operation of Cisco Nexus 9000 Series NX-OS to increase availability, collect information, and notify external systems or personnel about critical events.

For information about configuring EEM, see the *Cisco Nexus 9000 Series NX-OS System Management Configuration Guide*.

#### Smart Call Home

Combining Cisco GOLD and Cisco EEM capabilities, Smart Call Home provides an e-mail-based notification of critical system events. Smart Call Home has message formats that are compatible with pager services, standard e-mail, or XML-based automated parsing applications. You can use this feature to page a network support engineer, e-mail a network operations center, or use Cisco Smart Call Home services to automatically generate a case with Cisco’s Technical Assistance Center (TAC).

For information about configuring Smart Call Home, see the *Cisco Nexus 9000 Series NX-OS System Management Configuration Guide*.

#### Software Image

The Cisco Nexus 9000 Series NX-OS software consists of one NXOS software image (for example, n9000-dk9.6.1.2.I1.1.bin). This image runs on all Cisco Nexus 9000 Series switches.
Virtual Device Contexts

Cisco NX-OS can segment operating system and hardware resources into virtual device contexts (VDCs) that emulate virtual devices. The Cisco Nexus 9000 Series switches currently do not support multiple VDCs. All switch resources are managed in the default VDC.
This chapter describes the Cisco NX-OS service restartability for service-level high availability (HA) and includes the following sections:

- About Cisco NX-OS Service Restarts, page 9
- Licensing Requirements, page 10
- Restartability Infrastructure, page 10
- Process Restartability, page 11
- Restarts on Standby Supervisor Services, page 13
- Restarts on Switching Module Services, page 13
- Troubleshooting Restarts, page 13
- Additional References for Service-Level High Availability, page 14

About Cisco NX-OS Service Restarts

The Cisco NX-OS service restart features allow you to restart a faulty service without restarting the supervisor to prevent process-level failures from causing system-level failures. You can restart a service depending on current errors, failure circumstances, and the high-availability policy for the service. A service can undergo either a stateful or stateless restart. Cisco NX-OS allows services to store run-time state information and messages for a stateful restart. In a stateful restart, the service can retrieve this stored state information and resume operations from the last checkpoint service state. In a stateless restart, the service can initialize and run as if it had just been started with no prior state.

Not all services are designed for a stateful restart. For example, Cisco NX-OS does not store run-time state information for Layer 3 routing protocols (such as Open Shortest Path First [OSPF] and Routing Information Protocol [RIP]). Their configuration settings are preserved across a restart, but these protocols are designed to rebuild their operational state using information obtained from neighbor routers.

Note

This chapter refers to processes and services interchangeably. A process is considered to be a running instance of a service.
Licensing Requirements

<table>
<thead>
<tr>
<th>Product</th>
<th>License Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisco NX-OS</td>
<td>The service-level high availability features require no license. Any feature not included in a license package is bundled with the nx-os image and is provided for free.</td>
</tr>
</tbody>
</table>

For a complete explanation of the Cisco NX-OS licensing scheme, see the *Cisco NX-OS Licensing Guide*.

Restartability Infrastructure

Cisco NX-OS allows stateful restarts of most processes and services. Back-end management and orchestration of processes, services, and applications within a platform are handled by a set of high-level system-control services.

System Manager

The System Manager directs overall system function, service management, and system health monitoring and enforces high-availability policies. The System Manager is responsible for launching, stopping, monitoring, and restarting services as well as initiating and managing the synchronization of service states and supervisor states for a stateful switchover.

Persistent Storage Service

Cisco NX-OS services use the persistent storage service (PSS) to store and manage the operational run-time information. The PSS component works with system services to recover states in the event of a service restart. PSS functions as a database of state and run-time information, which allows services to make a checkpoint of their state information whenever needed. A restarting service can recover the last known operating state that preceded a failure, which allows for a stateful restart.

Each service that uses PSS can define its stored information as private (it can be read only by that service) or shared (the information can be read by other services). If the information is shared, the service can specify that it is local (the information can be read only by services on the same supervisor) or global (it can be read by services on either supervisor or on modules). For example, if the PSS information of a service is defined as shared and global, services on other modules can synchronize with the PSS information of the service that runs on the active supervisor.

Message and Transaction Service

The message and transaction service (MTS) is a high-performance interprocess communications (IPC) message broker that specializes in high-availability semantics. MTS handles message routing and queuing between services on and across modules and between supervisors. MTS facilitates the exchange of messages such as event notification, synchronization, and message persistency between system services and system components. MTS can maintain persistent messages and logged messages in queues for access even after a service restart.
HA Policies

Cisco NX-OS allows each service to have an associated set of internal HA policies that define how a failed service is restarted. Each service can have four defined policies—a primary and secondary policy when two supervisors are present and a primary and secondary policy when only one supervisor is present. If no HA policy is defined for a service, the default HA policy to be performed upon a service failure is a switchover if two supervisors are present or a supervisor reset if only one supervisor is present.

Each HA policy specifies three parameters:

- **Action to be performed by the System Manager:**
  - Stateful restart
  - Stateless restart
  - Supervisor switchover (or restart)

- **Maximum retries**—Spefies the number of restart attempts to be performed by the System Manager. If the service has not restarted successfully after this number of attempts, the HA policy is considered to have failed, and the next HA policy is used. If no other HA policy exists, the default policy is applied, resulting in a supervisor switchover or reset.

- **Minimum lifetime**—Spefies the time that a service must run after a restart attempt to consider the restart attempt as successful. The minimum lifetime is no less than 4 minutes.

Process Restartability

Process restartability ensures that a failed service can recover and resume operations without disrupting the data plane or other services. Depending on the service HA policies, previous restart failures, and the health of other services running on the same supervisor, the System Manager determines the action to be taken when a service fails.

The action taken by the System Manager for various failure conditions is described in the following table.

<table>
<thead>
<tr>
<th>Failure</th>
<th>System Manager Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Service/process exception</td>
<td>Service restart</td>
</tr>
<tr>
<td>Service/process crash</td>
<td>Service restart</td>
</tr>
<tr>
<td>Unresponsive service/process</td>
<td>Service restart</td>
</tr>
<tr>
<td>Repeated service failure</td>
<td>Supervisor reset (single) or switchover (dual)</td>
</tr>
<tr>
<td>Unresponsive System Manager</td>
<td>Supervisor reset (single) or switchover (dual)</td>
</tr>
<tr>
<td>Supervisor hardware failure</td>
<td>Supervisor reset (single) or switchover (dual)</td>
</tr>
</tbody>
</table>
Types of Process Restarts

A failed service is restarted by one of the methods described in this section, depending on the service’s HA implementation and HA policies,

### Stateful Restarts

When a restartable service fails, it is restarted on the same supervisor. If the new instance of the service determines that the previous instance was abnormally terminated by the operating system, the service then determines whether a persistent context exists. The initialization of the new instance attempts to read the persistent context to build a run-time context that makes the new instance appear like the previous one. After the initialization is complete, the service resumes the tasks that it was performing when it stopped. During the restart and initialization of the new instance, other services are unaware of the service failure. Any messages that are sent by other services to the failed service are available from the MTS when the service resumes.

Whether or not the new instance survives the stateful initialization depends on the cause of failure of the previous instance. If the service is unable to survive a few subsequent restart attempts, the restart is considered as failed. In this case, the System Manager performs the action specified by the HA policy of the services, forcing either a stateless restart, no restart, or a supervisor switchover or reset.

During a successful stateful restart, there is no delay while the system reaches a consistent state. Stateful restarts reduce the system recovery time after a failure.

The events before, during, and after a stateful restart are as follows:

1. The running services make a checkpoint of their run-time state information to the PSS.
2. The System Manager monitors the health of the running services that use heartbeats.
3. The System Manager restarts a service instantly when it crashes or hangs.
4. After restarting, the service recovers its state information from the PSS and resumes all pending transactions.
5. If the service does not resume a stable operation after multiple restarts, the System Manager initiates a reset or switchover of the supervisor.
6. Cisco NX-OS collects the process stack and core for debugging purposes with an option to transfer core files to a remote location.

When a stateful restart occurs, Cisco NX-OS sends a syslog message of level LOG_ERR. If SNMP traps are enabled, the SNMP agent sends a trap. If the Smart Call Home service is enabled, the service sends an event message.

<table>
<thead>
<tr>
<th>Failure</th>
<th>Action</th>
</tr>
</thead>
<tbody>
<tr>
<td>Kernel failure</td>
<td>Supervisor reset (single) or switchover (dual)</td>
</tr>
<tr>
<td>Watchdog timeout</td>
<td>Supervisor reset (single) or switchover (dual)</td>
</tr>
</tbody>
</table>
Stateless Restarts

Cisco NX-OS infrastructure components manage stateless restarts. During a stateless restart, the System Manager identifies the failed process and replaces it with a new process. The service that failed does not maintain its run-time state upon the restart. The service can either build the run-time state from the running configuration or if necessary, exchange information with other services to build a run-time state.

When a stateless restart occurs, Cisco NX-OS sends a syslog message of level LOG_ERR. If SNMP traps are enabled, the SNMP agent sends a trap. If the Smart Call Home service is enabled, the service sends an event message.

Switchovers

If a standby supervisor is available, Cisco NX-OS performs a supervisor switchover rather than a supervisor restart whenever multiple failures occur at the same time, because these cases are considered unrecoverable on the same supervisor. For example, if more than one HA application fails, that is considered an unrecoverable failure.

Restarts on Standby Supervisor Services

When a service fails on a supervisor that is in the standby state, the System Manager does not apply the HA policies and restarts the service after a delay of 30 seconds. The delay ensures that the active supervisor is not overloaded by repeated standby service failures and synchronizations. If the service being restarted requires synchronization with a service on the active supervisor, the standby supervisor is taken out of hot standby mode until the service is restarted and synchronized. Services that are not restartable cause the standby supervisor to reset.

When a standby service restart occurs, Cisco NX-OS sends a syslog message of level LOG_ERR. If SNMP traps are enabled, the SNMP agent sends a trap. If the Smart Call Home service is enabled, the service sends an event message.

Restarts on Switching Module Services

When services fail on a switching module or another nonsupervisor module, the recovery action is determined by HA policies for those services. Because service failures on nonsupervisor module services do not require a supervisor switchover, the recovery options are a stateful restart, a stateless restart, or a module reset. If the module is capable of a nondisruptive upgrade, it is also capable of a nondisruptive restart.

When a nonsupervisor module service restart occurs, Cisco NX-OS sends a syslog message of level LOG_ERR. If SNMP traps are enabled, the SNMP agent sends a trap. If the Smart Call Home service is enabled, the service sends an event message.

Troubleshooting Restarts

When a service fails, the system generates information that can be used to determine the cause of the failure. The following sources of information are available:

- Every service restart generates a syslog message of level LOG_ERR.
If the Smart Call Home service is enabled, every service restart generates a Smart Call Home event.

If SNMP traps are enabled, the SNMP agent sends a trap when a service is restarted.

When a service failure occurs on a local module, you can view a log of the event by entering the `show processes log` command in that module. The process logs are persistent across supervisor switchovers and resets.

When a service fails, a system core image file is generated. You can view recent core images by using the `show cores` command on the active supervisor. Core files are not persistent across supervisor switchovers and resets, but you can configure the system to export core files to an external server by using a file transfer utility such as the Trivial File Transfer Protocol (TFTP).

CISCO-SYSTEM-EXT-MIB contains a table for cores (cseSwCoresTable).

For information on collecting and using the generated information relating to service failures, see the Cisco Nexus 9000 Series NX-OS Troubleshooting Guide.

## Additional References for Service-Level High Availability

This section describes additional information related to service-level high availability.

### Related Documents

<table>
<thead>
<tr>
<th>Related Topic</th>
<th>Document Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Troubleshooting</td>
<td><em>Cisco Nexus 9000 Series NX-OS Troubleshooting Guide</em></td>
</tr>
<tr>
<td>Cisco NX-OS fundamentals</td>
<td><em>Cisco Nexus 9000 Series NX-OS Fundamentals Configuration Guide</em></td>
</tr>
<tr>
<td>Licensing</td>
<td><em>Cisco NX-OS Licensing Guide</em></td>
</tr>
</tbody>
</table>

### MIBs

<table>
<thead>
<tr>
<th>MIBs</th>
<th>MIBs Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIBs related to service-level high availability</td>
<td>To locate and download supported MIBs, go to the following URL: ftp://ftp.cisco.com/pub/mibs/supportlists/nexus9000/Nexus9000MIBSupportList.html</td>
</tr>
</tbody>
</table>
Network-Level High Availability

This chapter describes Cisco NX-OS network high availability (HA) and includes the following sections:

- About Network-Level High Availability, page 15
- Licensing Requirements, page 15
- Spanning Tree Protocol, page 16
- Virtual Port Channels, page 16
- First-Hop Redundancy Protocols, page 17
- Nonstop Forwarding in Routing Protocols, page 17
- Additional References for Network-Level High Availability, page 18

About Network-Level High Availability

Network-level HA is optimized by tools and functionality that provide failovers and fallbacks transparently and quickly. The features described in this chapter ensure high availability at the network level.

Licensing Requirements

<table>
<thead>
<tr>
<th>Product</th>
<th>License Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisco NX-OS</td>
<td>The network-level high availability features require no license. Any feature not included in a license package is bundled with the nx-os image and is provided for free.</td>
</tr>
<tr>
<td>BGP</td>
<td>Border Gateway Protocol (BGP) requires an Enterprise Services license.</td>
</tr>
</tbody>
</table>

For a complete explanation of the Cisco NX-OS licensing scheme and how to obtain and apply licenses, see the Cisco NX-OS Licensing Guide.
Spanning Tree Protocol

The Spanning Tree Protocol (STP) refers to IEEE 802.1w and IEEE 802.1s. If this publication is referring to the IEEE 802.1D STP, 802.1D is stated specifically.

When you create fault-tolerant internetworks, you must have a loop-free path between all nodes in a network. Multiple active paths between end stations cause loops in the network that result in network devices learning end station MAC addresses on multiple Layer 2 LAN ports. This condition can result in a broadcast storm, which creates an unstable network.

STP provides a loop-free network at the Layer 2 level. Layer 2 LAN ports send and receive STP frames, which are called Bridge Protocol Data Units (BPDUs), at regular intervals. Network devices do not forward these frames but use the frames to determine the network topology and to construct a loop-free path within that topology. Using the spanning tree topology, STP forces redundant data paths into a blocked state. If a network segment in the spanning tree fails and a redundant path exists, the STP algorithm recalculates the spanning tree topology and activates the blocked path.

Cisco NX-OS also supports the Multiple Spanning Tree Protocol (MSTP). The multiple independent spanning tree topology enabled by MSTP provides multiple forwarding paths for data traffic, enables load balancing, and reduces the number of STP instances required to support a large number of VLANs.

MST incorporates the Rapid Spanning Tree Protocol (RSTP), which allows rapid convergence. MST improves the fault tolerance of the network because a failure in one instance (forwarding path) does not affect other instances (forwarding paths).

You can configure spanning tree parameters only on Layer 2 interfaces; a spanning tree configuration is not allowed on a Layer 3 interface. For information on creating Layer 2 interfaces, see the Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide.

For details about STP behavior and configuration, see the Cisco Nexus 9000 Series NX-OS Layer 2 Switching Configuration Guide.

Virtual Port Channels

The major 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 gained with port channel technology to a single path. To address this limitation, Cisco NX-OS 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.

For more information on vPCs, see the Cisco Nexus 9000 Series NX-OS Interfaces Configuration Guide.
First-Hop Redundancy Protocols

Within a group of two or more routers, first-hop redundancy protocols (FHRPs) allow a transparent failover of the first-hop IP router. Cisco NX-OS supports the following FHRPs:

- Hot Standby Router Protocol (HSRP)—HSRP provides first-hop routing redundancy for IP hosts on Ethernet networks configured with a default gateway IP address. An HSRP router group of two or more routers chooses an active gateway and a standby gateway. The active gateway routes packets while the standby gateway remains idle until the active gateway fails or when preset conditions are met.

Many host implementations do not support any dynamic router discovery mechanisms but can be configured with a default router. Running a dynamic router discovery mechanism on every host is not feasible for a number of reasons, including administrative overhead, processing overhead, and security issues. HSRP provides failover services to these hosts.

- Virtual Router Redundancy Protocol (VRRP)—VRRP dynamically assigns responsibility for one or more virtual routers to the VRRP routers on a LAN, which allows several routers on a multi-access link to use the same virtual IP address. A VRRP router is configured to run VRRP with one or more other routers attached to a LAN. One router is elected as the virtual router master, while the other routers act as backups if the virtual router master fails.

For configuration details about FHRPs, see the Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration Guide.

Nonstop Forwarding in Routing Protocols

Cisco NX-OS provides a multilevel high-availability architecture. Open Shortest Path First version 2 (OSPFv2) supports stateful restart, which is also referred to as nonstop routing (NSR). If OSPFv2 experiences problems, it attempts to restart from its previous runtime state. The neighbors would not register any neighbor event in this case.

If the first restart is not successful and another problem occurs, OSPFv2 attempts a graceful restart. A graceful restart, or nonstop forwarding (NSF), allows OSPFv2 to remain in the data forwarding path through a process restart. When OSPFv2 needs to do a graceful restart, it first sends a link-local opaque (type 9) link-state advertisement (LSA), called a grace LSA. (For more information about opaque LSAs, see the Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration Guide.) The restarting of the OSPFv2 platform is called NSF capable. The grace LSA includes a grace period, which is a specified time that the neighbor OSPFv2 interfaces hold onto the LSAs from the restarting OSPFv2 interface. (Typically, OSPFv2 tears down the adjacency and discards all LSAs from a down or restarting OSPFv2 interface.) The participating neighbors, which are called NSF helpers, keep all LSAs that originate from the restarting OSPFv2 interface as if the interface were still adjacent. When the restarting OSPFv2 interface is operational again, it rediscovers its neighbors, establishes adjacency, and starts sending its LSA updates again. At this point, the NSF helpers recognize that graceful restart has finished.

Scenarios where a stateful restart is used:

- First recovery attempt after a process experiences problems.
- User-initiated switchover using the system switchover command.

Scenarios where graceful restart is used:
- Second recovery attempt after a process experiences problems within a 4-minute interval.
- Manual restart of the process using the `restart ospfv3` command.
- Active supervisor removal.
- Active supervisor reload using the `reload module active-sup` command.

### Additional References for Network-Level High Availability

This section describes additional information related to network-level high availability.

#### Related Documents

<table>
<thead>
<tr>
<th>Related Topic</th>
<th>Document Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Graceful restart</td>
<td><em>Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration Guide</em></td>
</tr>
<tr>
<td>Licensing</td>
<td><em>Cisco NX-OS Licensing Guide</em></td>
</tr>
</tbody>
</table>

#### MIBs

<table>
<thead>
<tr>
<th>MIBs</th>
<th>MIBs Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIBs related to network-level high availability</td>
<td>To locate and download supported MIBs, go to the following URL: ftp://ftp.cisco.com/pub/mibs/supportlists/nexus9000/Nexus9000MIBSupportList.html</td>
</tr>
</tbody>
</table>
System-Level High Availability

This chapter describes the Cisco NX-OS high availability (HA) system and application restart operations and includes the following sections:

- About Cisco NX-OS System-Level High Availability, page 19
- Licensing Requirements, page 20
- Physical Redundancy, page 20
- Configuring the Operating Mode for Fabric Modules, page 25
- Supervisor Restarts and Switchovers, page 26
- Displaying HA Status Information, page 31
- Additional References for System-Level High Availability, page 33

About Cisco NX-OS System-Level High Availability

Cisco NX-OS system-level HA mitigates the impact of hardware or software failures and is supported by the following features:

- Redundant hardware components:
  - Power supply
  - Fan tray
  - Switch fabric (Cisco Nexus 9504, 9508, and 9516 chassis only)
  - System controller (Cisco Nexus 9504, 9508, and 9516 chassis only)
  - Supervisor (Cisco Nexus 9504, 9508, and 9516 chassis only)
  For details about physical requirements and redundant hardware components, see the Site Preparation and Hardware Installation Guide for your specific Cisco Nexus 9000 Series chassis.

- HA software features:
  - Nonstop forwarding (NSF)—For details about nonstop forwarding, also known as graceful restart, see the Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration Guide.
* Embedded Event Manager (EEM)—For details about configuring EEM, see the *Cisco Nexus 9000 Series NX-OS System Management Configuration Guide*.

* Smart Call Home—For details about configuring Smart Call Home, see the *Cisco Nexus 9000 Series NX-OS System Management Configuration Guide*.

## Licensing Requirements

<table>
<thead>
<tr>
<th>Product</th>
<th>License Requirement</th>
</tr>
</thead>
<tbody>
<tr>
<td>Cisco NX-OS</td>
<td>With the exception of Smart Call Home, the system-level high availability features require no license. Any feature not included in a license package is bundled with the nx-os image and is provided for free.</td>
</tr>
<tr>
<td>Smart Call Home</td>
<td>Smart Call Home is available through Cisco SMARTnet Service and Cisco SP Base Service.</td>
</tr>
</tbody>
</table>

For a complete explanation of the Cisco NX-OS licensing scheme and how to obtain and apply licenses, see the *Cisco NX-OS Licensing Guide*.

## Physical Redundancy

The Cisco Nexus 9504, 9508, and 9516 chassis include the following physical redundancies:

- Power supply
- Fan tray
- Switch fabric
- System controller
- Supervisor module

The Cisco Nexus 9332, 9372, 9396, and 93128 chassis include the following physical redundancies:

- Power supply
- Fan tray

For additional details about physical redundancies, see the *Hardware Installation Guide* for your specific Cisco Nexus 9000 Series chassis.

## Power Supply Redundancy

The Cisco Nexus 9504 chassis supports up to four power supply modules, the Cisco Nexus 9508 chassis supports up to eight power supply modules, and the Cisco Nexus 9516 chassis supports up to ten power supply modules.
Each 9500 Series power supply module can deliver up to 3 kW.

The power subsystem allows the power supplies to be configured in one of the available redundancy modes. By installing more modules, you can ensure that the failure of one module does not disrupt system operations. You can replace the failed module while the system is operating. For information on power supply module installation and replacement, see the Hardware Installation Guide for your specific Cisco Nexus 9000 Series chassis.

**Power Modes**

Each of the power redundancy modes imposes different power budgeting and allocation models, which deliver varying usable power yields and capacities. For more information about power budgeting, usable capacity, planning requirements, and redundancy configuration, see the Hardware Installation Guide for your specific Cisco Nexus 9000 Series chassis.

The available power supply redundancy modes are described in the following table.

<table>
<thead>
<tr>
<th>Redundancy Mode</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combined (nonredundant)</td>
<td>This mode does not provide power redundancy. The available power is the total power capacity of all power supplies.</td>
</tr>
<tr>
<td>insrc-redundant (grid redundancy)</td>
<td>This mode provides grid redundancy when you connect half of the power supplies to one grid and the other half of the power supplies to the second grid. The available power is the amount of power available through a grid. To enable grid redundancy, you must connect the power supplies to the correct power grid slots. For example, on the Cisco Nexus 9508 switch, slots 1, 2, 3, and 4 are in grid A, and slots 5, 6, 7, and 8 are in grid B. To configure and operate in grid redundancy mode, you must connect half of your power supplies to the slots in grid A and the rest of your power supplies to the slots in grid B. For more information on power grid slot assignments for your power supplies, see the Hardware Installation Guide for your specific Cisco Nexus 9000 Series platform.</td>
</tr>
<tr>
<td>ps-redundant (N+1 redundancy)</td>
<td>This mode provides an extra power supply if an active power supply goes down. One power supply of all the available power supplies is considered an extra power supply, and the total available power is the amount provided by the active power supply units.</td>
</tr>
</tbody>
</table>

Use the `power redundancy-mode {combined | insrc_redundant | ps-redundant}` command to specify one of these power modes.
Fan Tray Redundancy

The Cisco Nexus 9500 Series chassis and the Cisco Nexus 9396 and 93128 chassis each contain three redundant system fan trays for cooling the system. The Cisco Nexus 9332 and 9372 chassis each contain four redundant fan trays.

The fan speeds are variable and depend on the temperature of the ASICs in the system. If fans are removed or go bad, the other fan modules can start running at a higher speed to compensate for the missing or failed fans. If the system temperature increases above the thresholds, the system shuts down.

- If a single fan fails within a fan tray, the fan speed of the other fans in the tray does not increase.
- If multiple fans fail within a fan tray, the fan speed increases to 100% on all the fan trays.
- If an entire fan tray is removed, the fan speed for the other two fan trays increases to 100% as soon as the tray is removed.
- If multiple fan trays are removed and not replaced within 2 minutes, the device reloads. When the device comes back, if it still detects the multiple fan tray failure, it reloads again after 2 minutes. If desired, you can use EEM to overwrite this policy.
- If a fan tray fails, leave the failed unit in place to ensure proper airflow until you can replace it. The fan trays are hot swappable, but you must replace one fan tray at a time. Otherwise, the device reboots after 2 minutes if multiple fan trays are missing.

Note

There is no time limit for replacing a single fan tray, but to ensure proper airflow, replace the fan tray as soon as possible.

Switch Fabric Redundancy

Cisco NX-OS provides switching fabric availability through redundant switch fabric module implementation. You can configure a single Cisco Nexus 9504, 9508, or 9516 chassis with one to six switch fabric modules for capacity and redundancy. Each line card installed in the system automatically connects to and uses all functionally installed switch fabric modules. A failure of a switch fabric module triggers an automatic reallocation and balancing of traffic across the remaining active switch fabric modules. Replacing the failed fabric module reverses this process. After you insert the replacement fabric module and bring it online, traffic is again redistributed across all installed fabric modules and redundancy is restored.

Fabric modules are hot swappable. Hot swapping can temporarily disrupt traffic. To prevent the disruption of traffic when you hot-swap fabric modules, use the `poweroff module slot-number` command before you remove a fabric module and the `no poweroff module slot-number` command after you reinsert the fabric module.

To get line rate traffic with X9464PX, X9464TX, and X9432PQ line cards, you must insert four fabric modules into four fabric module slots (FM2, FM3, FM4, and FM6). Additional fabric modules will not provide additional redundancy for these line cards.

To get line rate traffic with X9564PX, X9564TX, and X9536PQ line cards, you must insert three fabric modules into the even fabric module slots (FM2, FM4, and FM6). Additional fabric modules will provide additional redundancy for these line cards. Each even fabric module provides redundancy for each odd fabric
module failure (FM2 provides redundancy for FM1, FM4 provides redundancy for FM3, and FM6 provides redundancy for FM5).

To get line rate traffic with X9636PQ line cards, you must insert all six fabric modules. This line card is supported only in the 4-slot and 8-slot Cisco Nexus 9500 Series chassis.

### Line Card Failures

To keep line cards powered down whenever they fail or crash, use the `system module failure-action shutdown` command to prevent the cards from rebooting. This command is useful if your topology is configured for network-level redundancy and you want to prevent a second disruption from occurring in the network because a line card is trying to come up.

You can use the `show module module` command to verify that the line card has been powered down. If desired, use the `no poweroff module module` command to manually bring the line card back up.

```
switch(config)# system module failure-action shutdown
2014 Sep 8 23:31:51 switch %$ VDC-1 %$ %SYSMGR-SLOT1-2-SERVICE_CRASHED:
Service "ipfib" (PID 2558) hasn't caught signal 11 (core will be saved).

2014 Sep 8 23:32:25 switch %$ VDC-1 %$ %PLATFORM-2-MOD_PWRDN:
Module 1 powered down (Serial number SAL1815Q1DP)

switch(config)# show module 1
Mod Ports Module-Type Model Status
--- ----- ------------------------------------- --------------- ----------
1 52 48x1/10G-T 4x40G Ethernet Module N9K-X9564TX powered-dn

switch(config)# no poweroff module 1
2014 Sep 8 23:34:31 switch %$ VDC-1 %$ %PLATFORM-2-PFM_MODULE_POWER_ON:
Manual power-on of Module 1 from Command Line Interface

2014 Sep 8 23:34:31 switch %$ VDC-1 %$ %PLATFORM-2-MOD_DETECT:
Module 1 detected (Serial number SAL1815Q1DP) Module-Type 48x1/10G-T
4x40G Ethernet Module Model N9K-X9564TX

2014 Sep 8 23:34:31 switch %$ VDC-1 %$ %PLATFORM-2-MOD_PWRUP:
Module 1 powered up (Serial number SAL1815Q1DP)
```

### Fabric Module Operating Modes

Fabric modules with the network forwarding engine (NFE) can operate in two modes: full-rate mode (FRM) or oversubscribed mode (OSM). In OSM, all 32 40G ports are enabled on the NFE for a total I/O bandwidth of 1280G. In FRM, only 24 ports are enabled on the NFE for a total bandwidth of 960G at all packet sizes. In OSM, the NFE cannot run at line rate for a packet size of less than 200 bytes.

Fabric module NFEs run in FRM for Cisco NX-OS Releases 6.1(2)I1(x), 6.1(2)I2(1), and 6.1(2)I2(2). Starting with Cisco NX-OS Release 6.1(2)I2(2a), fabric module NFEs can run in FRM or OSM, but OSM is the default mode.

Some line cards cannot be powered up when the system fabric is operating in FRM. This table lists the line card types that are supported in each mode:

<table>
<thead>
<tr>
<th>Operating Mode</th>
<th>Supported Line Cards</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSM</td>
<td>Allows a mix of all three line card types (X9400, X9500, X9600)</td>
</tr>
</tbody>
</table>
System Controller Redundancy

Two redundant system controllers in the Cisco Nexus 9504, 9508, and 9516 chassis offload chassis management functions from the supervisor modules. The controllers are responsible for managing power supplies and fan trays and act as a central point for the Gigabit Ethernet Out-of-Band Channel (EOBC) between the supervisors, fabric modules, and line cards.

Supervisor Module Redundancy

The Cisco Nexus 9504, 9508, and 9516 chassis support dual supervisor modules to provide 1+1 redundancy for the control and management plane. A dual supervisor configuration operates in an active or standby capacity in which only one of the supervisor modules is active at any given time, while the other acts as a standby backup. The state and configuration remain constantly synchronized between the two supervisor modules to provide stateful switchover in the event of a supervisor module failure.

A Cisco NX-OS generic online diagnostics (GOLD) subsystem and additional monitoring processes on the supervisor trigger a stateful failover to the redundant supervisor when the processes detect unrecoverable critical failures, service restartability errors, kernel errors, or hardware failures.

If a supervisor-level unrecoverable failure occurs, the currently active failed supervisor triggers a switchover. The standby supervisor becomes the new active supervisor and uses the synchronized state and configuration while the failed supervisor is reloaded. If the failed supervisor is able to reload and pass self-diagnostics, it initializes, becomes the new standby supervisor, and then synchronizes its operating state with the newly active unit.

Supervisor Modules

Two supervisor modules are available for the Cisco Nexus 9000 Series switches: Supervisor A (SUP A) and Supervisor B (SUP B). The following table lists the differences between the two modules.

<table>
<thead>
<tr>
<th></th>
<th>Supervisor A</th>
<th>Supervisor B</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPU</td>
<td>4 core, 1.8 GHz</td>
<td>6 core, 2.1 GHz</td>
</tr>
<tr>
<td>Memory</td>
<td>16 GB</td>
<td>24 GB</td>
</tr>
<tr>
<td>SSD storage</td>
<td>64 GB</td>
<td>256 GB</td>
</tr>
<tr>
<td>Software release</td>
<td>6.1(2)I3(1) or later release</td>
<td>6.1(2)I3(1) or later release</td>
</tr>
</tbody>
</table>
SUP A and SUP B are not compatible and should not be installed in the same chassis, except for migration purposes. For dual supervisor systems, you should install either two SUP A modules or two SUP B modules (and not a combination of the two) to ensure supervisor module redundancy.

In a dual supervisor system, Cisco NX-OS checks the memory size of both the active and standby supervisors. If the memory size is different for each supervisor (because both SUP A and SUP B are installed), a message appears instructing you to replace SUP A with a second SUP B.

To migrate from SUP A to SUP B, insert SUP B into the device and enter the `system switchover` command. SUP B becomes the active supervisor, and SUP A becomes the standby supervisor, which is not a supported configuration. A warning message appears every hour until you remove SUP A or replace it with a second SUP B.

### Configuring the Operating Mode for Fabric Modules

You can configure the operating mode for fabric module network forwarding engines (NFEs).

#### Procedure

<table>
<thead>
<tr>
<th>Command or Action</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 1</strong></td>
<td>Enters global configuration mode.</td>
</tr>
<tr>
<td><code>configure terminal</code></td>
<td></td>
</tr>
<tr>
<td><strong>Example:</strong></td>
<td></td>
</tr>
<tr>
<td>switch# configure terminal switch(config)#</td>
<td></td>
</tr>
<tr>
<td><strong>Step 2</strong></td>
<td>Configures the system fabric in full-rate mode (FRM). You can use the <code>no</code> form of this command to return to oversubscribed mode (OSM).</td>
</tr>
<tr>
<td><code>[no] system fabric-mode full-rate</code></td>
<td></td>
</tr>
<tr>
<td><strong>Example:</strong></td>
<td></td>
</tr>
<tr>
<td>switch(config)# system fabric-mode full-rate</td>
<td></td>
</tr>
<tr>
<td><strong>Step 3</strong></td>
<td>(Optional) Displays the configured system fabric mode and the applied system fabric mode.</td>
</tr>
<tr>
<td><code>show system fabric-mode</code></td>
<td></td>
</tr>
<tr>
<td><strong>Example:</strong></td>
<td></td>
</tr>
<tr>
<td>switch(config)# show system fabric-mode</td>
<td></td>
</tr>
<tr>
<td><strong>Step 4</strong></td>
<td>Copies the running configuration to the startup configuration.</td>
</tr>
<tr>
<td><code>copy running-config startup-config</code></td>
<td></td>
</tr>
<tr>
<td><strong>Example:</strong></td>
<td></td>
</tr>
<tr>
<td>switch# copy running-config startup-config</td>
<td></td>
</tr>
<tr>
<td><strong>Step 5</strong></td>
<td>Reboots the entire device. A reboot is required for the configuration setting to take effect.</td>
</tr>
<tr>
<td><code>reload</code></td>
<td></td>
</tr>
</tbody>
</table>
Supervisor Restarts and Switchovers

Restarts on Single Supervisors

In a system with only one supervisor, when all HA policies have been unsuccessful in restarting a service, the supervisor restarts. The supervisor and all services reset and start with no prior state information.

Restarts on Dual Supervisors

When a supervisor-level failure occurs in a system with dual supervisors, the System Manager performs a switchover rather than a restart to maintain stateful operation. In some cases, however, a switchover might not be possible at the time of the failure. For example, if the standby supervisor module is not in a stable standby state, a restart rather than a switchover is performed.

Switchovers on Dual Supervisors

A dual supervisor configuration allows nonstop forwarding (NSF) with a stateful switchover (SSO) when a supervisor-level failure occurs. The two supervisors operate in an active/standby capacity in which only one of the supervisor modules is active at any given time, while the other acts as a standby backup. The two supervisors constantly synchronize the state and configuration in order to provide a seamless and stateful switchover of most services if the active supervisor module fails.

Switchover Characteristics

An HA switchover has the following characteristics:

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

Switchover Mechanisms

Switchovers occur by one of the following two mechanisms:

- 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.

When a switchover process begins, another switchover process cannot be started on the same switch until a stable standby supervisor module is available.
Switchover Failures

Supervisor switchovers are generally hitless and occur without traffic loss. If for some reason a switchover does not complete successfully, the supervisors reset. A reset prevents loops in the Layer 2 network if the network topology was changed during the switchover. For optimal performance of this recovery function, we recommend that you do not change the Spanning Tree Protocol (STP) default timers.

If three system-initiated switchovers occur within 20 minutes, all nonsupervisor modules shut down to prevent switchover cycling. The supervisors remain operational to allow you to collect system logs before resetting the switch.

Manually Initiating a Switchover

To manually initiate a switchover from an active supervisor module to a standby supervisor module, use the `system switchover` command. After you run this command, you cannot start another switchover process on the same system until a stable standby supervisor module is available.

Note

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

To ensure that an HA switchover is possible, use 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, you can manually initiate a switchover.

Switchover Guidelines

Follow these guidelines when performing a switchover:

• When you manually initiate a switchover, it takes place immediately.

• A switchover can be performed only when two supervisor modules are functioning in the switch.

• The modules in the chassis must be functioning.

Verifying Switchover Possibilities

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

• Use the `show 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:

<table>
<thead>
<tr>
<th>Mod</th>
<th>Ports</th>
<th>Module-Type</th>
<th>Model</th>
<th>Status</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>32</td>
<td>32p 40G Ethernet Module</td>
<td>N9K-X9432PQ</td>
<td>ok</td>
</tr>
<tr>
<td>2</td>
<td>52</td>
<td>48x1/10G SFP+ 4x40G Ethernet Module</td>
<td>N9K-X9564PX</td>
<td>ok</td>
</tr>
<tr>
<td>5</td>
<td>52</td>
<td>48x1/10G SFP+ 4x40G Ethernet Module</td>
<td>N9K-X9464PX</td>
<td>ok</td>
</tr>
<tr>
<td>6</td>
<td>36</td>
<td>36p 40G Ethernet Module</td>
<td>N9K-X9536PQ</td>
<td>ok</td>
</tr>
<tr>
<td>7</td>
<td>36</td>
<td>36p 40G Ethernet Module</td>
<td>N9K-X9536PQ</td>
<td>ok</td>
</tr>
<tr>
<td>10</td>
<td>32</td>
<td>32p 40G Ethernet Module</td>
<td>N9K-X9432PQ</td>
<td>ok</td>
</tr>
</tbody>
</table>
Switchovers on Dual Supervisors

<table>
<thead>
<tr>
<th>Mod</th>
<th>Sw</th>
<th>Hw</th>
<th>Slot</th>
</tr>
</thead>
<tbody>
<tr>
<td>11</td>
<td>S2</td>
<td>48x1/10G-T 4x40G Ethernet Module</td>
<td>N9K-X9564TX ok</td>
</tr>
<tr>
<td>12</td>
<td>S2</td>
<td>48x1/10G-T 4x40G Ethernet Module</td>
<td>N9K-X9464TX ok</td>
</tr>
<tr>
<td>15</td>
<td>S2</td>
<td>48x1/10G SFP+ 4x40G Ethernet Module</td>
<td>N9K-X9464PX ok</td>
</tr>
<tr>
<td>21</td>
<td>0</td>
<td>Fabric Module</td>
<td>N9K-C9516-FM ok</td>
</tr>
<tr>
<td>22</td>
<td>0</td>
<td>Fabric Module</td>
<td>N9K-C9516-FM ok</td>
</tr>
<tr>
<td>23</td>
<td>0</td>
<td>Fabric Module</td>
<td>N9K-C9516-FM ok</td>
</tr>
<tr>
<td>24</td>
<td>0</td>
<td>Fabric Module</td>
<td>N9K-C9516-FM ok</td>
</tr>
<tr>
<td>25</td>
<td>0</td>
<td>Fabric Module</td>
<td>N9K-C9516-FM ok</td>
</tr>
<tr>
<td>26</td>
<td>0</td>
<td>Fabric Module</td>
<td>N9K-C9516-FM ok</td>
</tr>
<tr>
<td>27</td>
<td>0</td>
<td>Supervisor Module</td>
<td>N9K-SUP-A ha-standby</td>
</tr>
<tr>
<td>28</td>
<td>0</td>
<td>Supervisor Module</td>
<td>N9K-SUP-A active *</td>
</tr>
<tr>
<td>29</td>
<td>0</td>
<td>System Controller</td>
<td>N9K-SC-A active</td>
</tr>
<tr>
<td>30</td>
<td>0</td>
<td>System Controller</td>
<td>N9K-SC-A standby</td>
</tr>
</tbody>
</table>

Mod  Sw  Hw  Slot
---  ----------  ------  ----
1    6.1(2)T3(1) 0.1050 LC1
2    6.1(2)T3(1) 0.2010 LC2
5    6.1(2)T3(1) 0.1010 LC5
6    6.1(2)T3(1) 0.2060 LC6
7    6.1(2)T3(1) 0.2060 LC7
10   6.1(2)T3(1) 0.1010 LC10
11   6.1(2)T3(1) 0.2100 LC11
12   6.1(2)T3(1) 0.1010 LC12
15   6.1(2)T3(1) 0.1050 LC15
21   6.1(2)T3(1) 0.3010 FM1
22   6.1(2)T3(1) 0.3040 FM2
23   6.1(2)T3(1) 0.3040 FM3
24   6.1(2)T3(1) 0.3040 FM4
25   6.1(2)T3(1) 0.3010 FM5
26   6.1(2)T3(1) 0.3040 FM6
27   6.1(2)T3(1) 1.1 SUP1
28   6.1(2)T3(1) 1.1 SUP2
29   6.1(2)T3(1) 1.2 SC1
30   6.1(2)T3(1) 1.2 SC2

Mod  MAC-Address(es)  Serial-Num
---  ----------------------  ----------
1    74-26-ac-10-cb-0c to 74-26-ac-10-cb-9f  SAL1817REX2
2    00-22-bd-fd-93-57 to 00-22-bd-fd-93-9a  SAL1733B92R
5    74-26-ac-eb-99-0c to 74-26-ac-eb-99-4f  SAL1814P7LM
6    c0-8c-60-62-60-98 to c0-8c-60-62-61-2b  SAL1812NTG1
7    c0-8c-60-62-62-f7-70 to c0-8c-60-62-60-03  SAL1812NTFD
10   74-26-ac-e9-32-68 to 74-26-ac-e9-32-f2  SAL1811NH4K
12   74-26-ac-ec-2b-50 to 74-26-ac-ec-2b-93  SAL1816UQX
15   c0-8c-60-62-a3-b4 to c0-8c-60-62-a3-f7  SAL1816QGXH
21   NA  SAL1801K507
22   NA  SAL1813P9Y2
23   NA  SAL1813P9YM
24   NA  SAL1813P9Y9
25   NA  SAL1801K50F
26   NA  SAL1813NZN3
27   c0-67-af-a1-0e-d6 to c0-67-af-a1-0e-e7  SAL1803KXYX
28   c0-67-af-a1-0d-a4 to c0-67-af-a1-0d-b5  SAL1804L578
29   NA  SAL1801JU2Z
30   NA  SAL1801JU4V

Mod  Online Diag Status
---  ---------
1    Pass
2    Pass
5    Pass
6    Pass
7    Pass
10   Pass
11   Pass
12   Pass
15   Pass
21   Pass
22   Pass
23   Pass
24   Pass
Replacing the Active Supervisor Module in a Dual Supervisor System

You can nondisruptively replace the active supervisor module in a dual supervisor system.

**Procedure**

<table>
<thead>
<tr>
<th>Step</th>
<th>Command or Action</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td>switch # system switchover</td>
<td>Initiates a manual switchover to the standby supervisor. <strong>Note</strong> Wait until the switchover completes and the standby supervisor becomes active.</td>
</tr>
<tr>
<td>Step 2</td>
<td>switch# reload module slot-number force</td>
<td>Boots the supervisor module replacement immediately. <strong>Note</strong> If you do not force the boot, the replacement supervisor module should be booted by the active supervisor module 6 minutes after insertion. For information on replacing a supervisor module, see the Site Preparation and Hardware Installation Guide for your specific Cisco Nexus 9000 Series chassis.</td>
</tr>
<tr>
<td>Step 3</td>
<td>switch# copy bootflash:nx-os-image bootflash:nx-os-image</td>
<td>Copies the nx-os image from the active supervisor module to the standby supervisor module.</td>
</tr>
<tr>
<td>Step 4</td>
<td>switch# configure terminal</td>
<td>Enters global configuration mode.</td>
</tr>
<tr>
<td>Step 5</td>
<td>switch (config)# boot nxos bootflash:nx-os-image [sup-number]</td>
<td>Configures the standby supervisor boot variables.</td>
</tr>
</tbody>
</table>
Replacing the Standby Supervisor Module in a Dual Supervisor System

You can nondisruptively replace the standby supervisor module in a dual supervisor system.

<table>
<thead>
<tr>
<th>Command or Action</th>
<th>Purpose</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Step 6</strong> switch(config)# copy running-config startup-config</td>
<td>(Optional) Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration.</td>
</tr>
</tbody>
</table>

This example shows how to replace the active supervisor module in a dual supervisor system:

```
switch# system switchover
Raw time read from Hardware Clock: Y=2013 M=2 D=2 07:35:48
writing reset reason ?,
NX9 SUP Ver 3.17.0
Serial Port Parameters from CMOS
PMCON_1: 0x200
PMCON_2: 0x0
PMCON_3: 0x3a
PM1_STS: 0x1
Performing Memory Detection and Testing
Testing 1 DRAM Patterns
Total mem found : 4096 MB
Memory test complete.
NumCpus = 2.
Status 61: PCI DEVICES Enumeration Started
Status 62: PCI DEVICES Enumeration Ended
Status 9F: Dispatching Drivers
Status 9E: IOPPGA Found
Status 9A: Booting From Primary ROM
Status 98: Found Cisco IDE
Status 98: Found Cisco IDE
Status 90: Loading Boot Loader
Reset Reason Registers: 0x1 0x10
Filesystem type is ext2fs, partition type 0x83
Filesystem type is ext2fs, partition type 0x83

GNU GRUB version 0.97
Loader Version 3.17.0

current standby sup
----------------------------
switch(standby)# 2014 Aug 2 07:35:46 switch %$ VDC-1 %$ %KERN-2-SYSTEM_MSG: Switchover started by redundancy driver - kernel
2014 Aug 2 07:35:47 switch %$ VDC-1 %$ %SYSMGR-2-HASWITCHOVER_PRE_START: This supervisor is becoming active (pre-start phase).
2014 Aug 2 07:35:47 switch %$ VDC-1 %$ %SYSMGR-2-HASWITCHOVER_START: This supervisor is becoming active.
2014 Aug 2 07:35:48 switch %$ VDC-1 %$ %SYSMGR-2-SWITCHOVER_OVER: Switchover completed.

switch# reload module 27 force
switch# copy bootflash:n9000-dk9.6.1.2.I3.1.bin bootflash:n9000-dk9.6.1.2.I3.1.bin
switch# config terminal
switch# boot nxos bootflash:n9000-dk9.6.1.2.I3.1.bin sup-1
switch# copy running-config startup-config
```
Procedure

<table>
<thead>
<tr>
<th>Step</th>
<th>Command or Action</th>
<th>Purpose</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>Step 1</td>
<td><code>switch# reload module slot-number force</code></td>
<td>Boots the supervisor module replacement immediately.</td>
<td>If you do not force the boot, the replacement supervisor module should be booted by the active supervisor module 6 minutes after insertion. For information on replacing a supervisor module, see the Site Preparation and Hardware Installation Guide for your specific Cisco Nexus 9000 Series chassis.</td>
</tr>
<tr>
<td>Step 2</td>
<td><code>switch# copy bootflash:nx-os-image bootflash:nx-os-image</code></td>
<td>Copies the nx-os image from the active supervisor module to the standby supervisor module.</td>
<td></td>
</tr>
<tr>
<td>Step 3</td>
<td><code>switch# configure terminal</code></td>
<td>Enters global configuration mode.</td>
<td></td>
</tr>
<tr>
<td>Step 4</td>
<td><code>switch# (config)# boot nxos bootflash:nx-os-image [sup-number]</code></td>
<td>Configures the standby supervisor boot variables.</td>
<td></td>
</tr>
<tr>
<td>Step 5</td>
<td><code>switch(config)# copy running-config startup-config</code></td>
<td>(Optional) Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration.</td>
<td></td>
</tr>
</tbody>
</table>

This example shows how to replace the standby supervisor module in a dual supervisor system:

```
switch# reload module 27 force
switch# copy bootflash:n9000-dk9.6.1.2.I3.1.bin bootflash:n9000-dk9.6.1.2.I3.1.bin
switch# configure terminal
switch# boot nxos bootflash:n9000-dk9.6.1.2.I3.1.bin sup-1
switch# copy running-config startup-config
```

**Displaying HA Status Information**

Use the `show system redundancy status` command to view the HA status of the system.

```
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 system is operationally HA and can perform automatic synchronization.
- If the internal state of one of the supervisor modules is none, the system cannot perform automatic synchronization.

The following table lists the possible values for the redundancy states.

**Table 4: Redundancy States**

<table>
<thead>
<tr>
<th>State</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Not present</td>
<td>The supervisor module is not present or is not plugged into the chassis.</td>
</tr>
<tr>
<td>Initializing</td>
<td>The diagnostics have passed, and the configuration is being downloaded.</td>
</tr>
<tr>
<td>Active</td>
<td>The active supervisor module and the switch are ready to be configured.</td>
</tr>
<tr>
<td>Standby</td>
<td>A switchover is possible.</td>
</tr>
<tr>
<td>Failed</td>
<td>The system detects a supervisor module failure on initialization and automatically attempts to power-cycle the module three times. After the third attempt, it continues to display a failed state.</td>
</tr>
<tr>
<td>Offline</td>
<td>The supervisor module is intentionally shut down for debugging purposes.</td>
</tr>
<tr>
<td>At BIOS</td>
<td>The system has established a connection with the supervisor, and the supervisor module is performing diagnostics.</td>
</tr>
<tr>
<td>Unknown</td>
<td>The system is in an invalid state. If it persists, call TAC.</td>
</tr>
</tbody>
</table>

The following table lists the possible values for the supervisor module states.

**Table 5: Supervisor States**

<table>
<thead>
<tr>
<th>State</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Active</td>
<td>The active supervisor module in the switch is ready to be configured.</td>
</tr>
<tr>
<td>HA standby</td>
<td>A switchover is possible.</td>
</tr>
<tr>
<td>Offline</td>
<td>The system is intentionally shut down for debugging purposes.</td>
</tr>
<tr>
<td>Unknown</td>
<td>The system is in an invalid state and requires a support call to TAC.</td>
</tr>
</tbody>
</table>

The following table lists the possible values for the internal redundancy states.
Table 6: Internal States

<table>
<thead>
<tr>
<th>State</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>HA standby</td>
<td>The HA switchover mechanism in the standby supervisor module is enabled.</td>
</tr>
<tr>
<td>Active with no standby</td>
<td>A switchover is impossible.</td>
</tr>
<tr>
<td>Active with HA standby</td>
<td>The active supervisor module in the switch is ready to be configured. The standby supervisor module is in the ha-standby state.</td>
</tr>
<tr>
<td>Shutting down</td>
<td>The system is being shut down.</td>
</tr>
<tr>
<td>HA switchover in progress</td>
<td>The system is in the process of entering the active state.</td>
</tr>
<tr>
<td>Offline</td>
<td>The system is intentionally shut down for debugging purposes.</td>
</tr>
<tr>
<td>HA synchronization in progress</td>
<td>The standby supervisor module is in the process of synchronizing its state with the active supervisor modules.</td>
</tr>
<tr>
<td>Standby (failed)</td>
<td>The standby supervisor module is not functioning.</td>
</tr>
<tr>
<td>Active with failed standby</td>
<td>The active supervisor module and the second supervisor module are present, but the second supervisor module is not functioning.</td>
</tr>
<tr>
<td>Other</td>
<td>The system is in a transient state. If it persists, call TAC.</td>
</tr>
</tbody>
</table>

Additional References for System-Level High Availability

This section describes additional information related to system-level high availability.
## Related Documents

<table>
<thead>
<tr>
<th>Related Topic</th>
<th>Document Title</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hardware</td>
<td><em>Cisco Nexus 9504 NX-OS-Mode Switch Hardware Installation Guide</em></td>
</tr>
<tr>
<td></td>
<td><em>Cisco Nexus 9508 NX-OS-Mode Switch Hardware Installation Guide</em></td>
</tr>
<tr>
<td></td>
<td><em>Cisco Nexus 9516 NX-OS-Mode Switch Hardware Installation Guide</em></td>
</tr>
<tr>
<td></td>
<td><em>Cisco Nexus 9332PQ NX-OS-Mode Switch Hardware Installation Guide</em></td>
</tr>
<tr>
<td></td>
<td><em>Cisco Nexus 9372PX NX-OS-Mode Switch Hardware Installation Guide</em></td>
</tr>
<tr>
<td></td>
<td><em>Cisco Nexus 9372TX NX-OS-Mode Switch Hardware Installation Guide</em></td>
</tr>
<tr>
<td></td>
<td><em>Cisco Nexus 9396PX NX-OS-Mode Switch Hardware Installation Guide</em></td>
</tr>
<tr>
<td></td>
<td><em>Cisco Nexus 9396TX NX-OS-Mode Switch Hardware Installation Guide</em></td>
</tr>
<tr>
<td></td>
<td><em>Cisco Nexus 93128TX NX-OS-Mode Switch Hardware Installation Guide</em></td>
</tr>
<tr>
<td>Power mode configuration and Cisco NX-OS</td>
<td><em>Cisco Nexus 9000 Series NX-OS Fundamentals Configuration Guide</em></td>
</tr>
<tr>
<td>fundamentals</td>
<td></td>
</tr>
<tr>
<td>Nonstop forwarding (NSF)</td>
<td><em>Cisco Nexus 9000 Series NX-OS Unicast Routing Configuration Guide</em></td>
</tr>
<tr>
<td>EEM and Smart Call Home</td>
<td><em>Cisco Nexus 9000 Series NX-OS System Management Configuration Guide</em></td>
</tr>
<tr>
<td>Licensing</td>
<td><em>Cisco NX-OS Licensing Guide</em></td>
</tr>
</tbody>
</table>

## MIBs

<table>
<thead>
<tr>
<th>MIBs</th>
<th>MIBs Link</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIBs related to system-level high availability</td>
<td>To locate and download supported MIBs, go to the following URL:</td>
</tr>
</tbody>
</table>