CHAPTER 4

Understanding System-Level High Availability

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

Information 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:
  - Supervisor
  - Switch fabric
  - Power supply
  - Fan trays

  For details about physical requirements and redundant hardware components, respectively, see the *Cisco Nexus 7000 Series Site Preparation Guide* and the *Cisco Nexus 7000 Series Hardware Installation and Reference Guide*.
- HA software features:
  - In-service software upgrades (ISSU) — For details about configuring and performing nondisruptive upgrades, see Chapter 5, "Understanding In-Service Software Upgrades."
  - Nonstop forwarding (NSF) — For details about nonstop forwarding, also known as *graceful restart*, see the *Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide, Release 6.x*.
  - Virtual device contexts (VDCs) — For details about VDCs and HA, see the *Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide, Release 4.1*.
  - Generic online diagnostics (GOLD) — For details about configuring GOLD, see the *Cisco Nexus 7000 Series NX-OS System Management Configuration Guide, Release 6.x*.
  - Embedded event manager (EEM) — For details about configuring EEM, see the *Cisco Nexus 7000 Series NX-OS System Management Configuration Guide, Release 6.x*.
  - Smart Call Home — For details about configuring Smart Call Home, see the *Cisco Nexus 7000 Series NX-OS System Management Configuration Guide, Release 6.x*.

Virtualization Support

For information about system-level high availability within a virtual device context (VDC), see the "VDC High Availability" section on page 4-12.

Note For complete information on VDCs, see the *Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide, Release 4.1*.

Licensing Requirements

The following table shows the licensing requirements for system-level high availability features:

Product	License Requirement
Cisco NX-OS	WIth the exception of VDC and Smart Call Home, the system-level high availability features require no license. Any feature not included in a license package is bundled with the Cisco NX-OS system images and is provided for free.
VDC	VDC requires an Advanced Services license.
Smart Call Home	Smart Call Home is available through Cisco SMARTnet Service and Cisco SP Base Service.

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 Nexus 7000 series includes the following physical redundancies:

For additional details about physical redundancies, see the *Cisco Nexus 7000 Series Site Preparation Guide* and the *Cisco Nexus 7000 Series Hardware Installation and Reference Guide*.

Power Supply Redundancy

The Nexus 7000 series supports up to three power supply modules on a Cisco Nexus 7010 switch and up to four power supplies on a Cisco Nexus 7018 switch. Each power supply module can deliver up to 7.5 KW, depending on the number of inputs and the input line voltage. By installing two or three modules, you can ensure that the failure of one module will 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 *Cisco Nexus 7000 Series Hardware Installation and Reference Guide.*

For further redundancy, each power supply module includes two internalized isolated power units, which give it two power paths per modular power supply, and six paths in total, per chassis, when fully populated. In addition, the power subsystem allows the three power supplies to be configured in any one of four redundancy modes.

Power Modes

Each of the four available power redundancy modes imposes different power budgeting and allocation models, which in turn deliver varying usable power yields and capacities. For more information regarding power budgeting, usable capacity, planning requirements, and redundancy configuration, see the *Cisco Nexus 7000 Series Hardware Installation and Reference Guide.*

The available power supply redundancy modes are described in Table 4-1.

Table 4-1 Power Redundancy Modes

Redundancy Mode	Description
Combined	This mode does not provide power redundancy. The available power is the total power capacity of all power supplies.
insrc-redundant	This mode utilizes two electrical grids, each one powering a half module within each power supply. If one power grid goes down, each power supply continues to draw power through its other half module. The available power is the amount of power by the lesser of the two grids through the power supplies.
ps-redundant	This mode provides an extra power supply in case an active power supply goes down. The power supply that can supply the most power operates in standby mode. The other one or two power supplies are active. The available power is the amount of power provided by the active power supply units.
redundant	This mode combines power supply redundancy and input source redundancy, which means that the chassis has an extra power supply and each half of each power supply is connected to one electrical grid while the other half of each power supply is connected to the other electrical grid. The available power is the lesser of the available power for power supply mode and input source mode.

Fan Tray Redundancy

The Cisco Nexus 7000 series chassis contains two redundant system fan trays for I/O module cooling and two additional fan trays for switch fabric module cooling. Only one of each pair of fan trays is required to provide system cooling.

The fan speeds are variable and are automatically adjusted to one of 16 levels in order to optimize system cooling while minimizing overall system noise and power draw. A detected failure of a fan within a given fan tray will trigger an increase in the speed of the remaining fans to compensate for the failure. A detected removal of an entire fan tray, without replacement, will initiate a system shutdown after a three-minute warning period.

Starting with Cisco NX-OS Release 5.0(2a), the fan shutdown policy for the 10-slot chassis is as follows:

- If a system fan is removed: Earlier releases shut off the other fan in 3 minutes. The new policy is to increase the speed of the other fan based on the table mapping.

- If a fabric fan is removed: Earlier releases shut off the other fan in 3 minutes. The new policy is to increase the speed of the other fan to the maximum.

Caution In the case of a fan tray failure, in the Nexus 7009 or the Nexus 7018 device, you must leave the failed unit in place to ensure proper airflow until a replacement is made available. The fan trays are hot swappable, but you must complete the removal and replacement within three minutes to avoid an automatic system shutdown.

Switch Fabric Redundancy

Cisco NX-OS provides switching fabric availability through redundant switch fabric module implementation. You can configure a single Nexus 7000 series with one to five switch fabric cards for capacity and redundancy. Each I/O module 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.

Supervisor Module Redundancy

The Nexus 7000 series supports 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 statefu1 switchover in the event of a supervisor module failure.

Cisco NX-OS's Generic On-Line 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.

For additional details on supervisor switchovers, see the "Supervisor Restarts and Switchovers" section on page 4-4.

Supervisor Restarts and Switchovers

This section includes the following topics:

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 will perform a switchover rather than a restart to maintain stateful operation. In some cases, however, a switchover may 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 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.
- It does not reload the Connectivity Management Processor (CMP).

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

If a switchover does not complete successfully within 28 seconds, the supervisors will 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 will 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

When you manually initiate a switchover, it takes place immediately. Follow these guidelines when performing a switchover:

- A switchover can only be performed 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. For information about the **show system redundancy status** command, see the "Displaying HA Status Information" section on page 4-10.

- 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
---  -----  ----------------------------------- ------------------ ----------
1    0      Supervisor module-1X                N7K-SUP1           active *
2    0      Supervisor module-1X                N7K-SUP1           ha-standby
3    32     1/10 Gbps Ethernet Module           N7K-D132XP-15      ok
4    48     1/10 Gbps Ethernet Module           N7K-F248XP-24      ok
5    48     10/100/1000 Mbps Ethernet XL Module N7K-M148GT-11L     ok
6    32     1/10 Gbps Ethernet Module           N7K-F132XP-15      ok
9    32     1/10 Gbps Ethernet Module           N7K-F132XP-15      ok

Mod  Sw              Hw
---  --------------  ------
1    6.0(1)          1.8
2    6.0(1)          1.1
3    6.0(1)          0.405
4    6.0(1)          0.500
5    6.0(1)          1.0
6    6.0(1)          0.617
9    6.0(1)          0.616


Mod  MAC-Address(es)                         Serial-Num
---  --------------------------------------  ----------
1    f0-25-72-ab-a3-f8 to f0-25-72-ab-a4-00  JAF1446BMRR
2    00-22-55-77-bc-48 to 00-22-55-77-bc-50  JAB122901WK
3    00-24-f7-1b-69-70 to 00-24-f7-1b-69-b4  JAF1321ARLQ
4    40-55-39-25-c8-00 to 40-55-39-25-c8-34  JAF1530AAAF
5    e8-b7-48-00-03-60 to e8-b7-48-00-03-94  JAF1513BPCH
6    f8-66-f2-02-a1-f8 to f8-66-f2-02-a2-3c  JAF1427DETN
9    a8-b1-d4-57-bc-bc to a8-b1-d4-57-bd-00  JAF1424CFMH

Mod  Online Diag Status
---  ------------------
1    Pass
2    Pass
3    Pass
4    Pass
5    Pass
6    Pass
9    Pass

Xbar Ports  Module-Type                          Model              Status
---  -----  ----------------------------------- ------------------ ----------
2    0      Fabric Module 2                     N7K-C7009-FAB-2    ok
4    0      Fabric Module 2                     N7K-C7009-FAB-2    ok
5    0      Fabric Module 2                     N7K-C7009-FAB-2    ok

Xbar Sw              Hw
---  --------------  ------
2    NA              0.201
4    NA              0.203
5    NA              0.201


Xbar MAC-Address(es)                         Serial-Num
---  --------------------------------------  ----------
2    NA                                      JAF1406ATRH
4    NA                                      JAF1422AHCP
5    NA                                      JAF1406ATRQ
```

```
* this terminal session
switch#
```

The Status column in the output should display an OK status for switching modules and an active or ha-standby status for supervisor modules.

- 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 are as follows:

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

Replacing the Active Supervisor Module in a Dual Supervisor System

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

To replace the active supervisor module, follow these steps:

Step 1 Initiate a manual switchover to the standby supervisor.

```
switch# system switchover
 Raw time read from Hardware Clock: Y=2009 M=2 D=2 07:35:48
 writing reset reason 7,

NX7 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: IOFPGA 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)# 2009 Feb  2 07:35:46 switch %$ VDC-1 %$ %KERN-2-SYSTEM_MSG: Switchover
started by redundancy driver - kernel
```

```
2009 Feb  2 07:35:47 switch %$ VDC-1 %$ %SYSMGR-2-HASWITCHOVER_PRE_START: This supervisor
is becoming active (pre-start phase).
2009 Feb  2 07:35:47 switch %$ VDC-1 %$ %SYSMGR-2-HASWITCHOVER_START: This supervisor is
becoming active.
2009 Feb  2 07:35:48 switch %$ VDC-1 %$ %SYSMGR-2-SWITCHOVER_OVER: Switchover completed.
```

Note Wait until the the switchover completes and the standby supervisor becomes active.

Step 2 Power down the supervisor module you are replacing.

```
switch# out-of-service module 6
```

Step 3 Replace the supervisor module. For information on replacing a supervisor module, see the *Cisco Nexus 7000 Series Hardware Installation and Reference Guide*.

Note Boot the replacement supervisor module from the active supervisor module after six minutes of insertion. If you don't want to wait six minutes, use the **reload module** *slot-number* **force** command to boot the replacement supervisor immediately.

Step 4 Copy the kickstart image from the active supervisor module to the standby supervisor module.

```
switch# copy bootflash:n7000-s1-kickstart.6.0.1.gbin.S30
bootflash://sup-remote/n7000-s1-kickstart.6.0.1.gbin.S30
```

Step 5 Copy the system image from the active supervisor module to the standby supervisor module.

```
switch# copy bootflash:n7000-s1-dk9.6.0.1.gbin.S30
bootflash://sup-remote/n7000-s1-dk9.6.0.1.gbin.S30
```

Step 6 Configure the standby supervisor boot variables.

```
switch# config t
switch# boot kickstart bootflash://sup-remote/n7000-s1-kickstart.6.0.1.gbin.S30
switch# boot system bootflash://sup-remote/n7000-s1-dk9.6.0.1.gbin.S30
```

Step 7 Save these changes to the startup configuration.

```
switch# copy running-config startup-config
```

Replacing the Standby Supervisor Module in a Dual Supervisor System

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

To replace the standby supervisor module, follow these steps:

Step 1 Power down the standby supervisor module.

```
switch# out-of-service module 6
```

Step 2 Replace the supervisor module. For information on replacing a supervisor module, see the *Cisco Nexus 7000 Series Hardware Installation and Reference Guide*.

Note Boot the replacement supervisor module from the active supervisor module after six minutes of insertion. If you don't want to wait six minutes, use the **reload module** *slot-number* **force** command to boot the replacement supervisor immediately.

Step 3 Copy the kickstart image from the active supervisor module to the standby supervisor module.

```
switch# copy bootflash:n7000-s1-kickstart.6.0.1.bin
bootflash://sup-remote/n7000-s1-kickstart.6.0.1.bin
```

Step 4 Copy the system image from the active supervisor module to the standby supervisor module.

```
switch# copy bootflash:n7000-s1-dk9.6.0.1.bin
bootflash://sup-remote/n7000-s1-dk9.6.0.1.bin
```

Step 5 Configure the standby supervisor boot variables.

```
switch# config t
switch# boot kickstart bootflash://sup-remote/n7000-s1-kickstart.6.0.1.bin
switch# boot system bootflash://sup-remote/n7000-s1-dk9.6.0.1.bin
```

Step 6 Save these changes to the startup configuration.

```
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. Tables 4-2 to 4-4 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 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.

Table 4-2 lists the possible values for the redundancy states.

Table 4-2 Redundancy States

State	Description
Not present	The supervisor module is not present or is not plugged into the chassis.
Initializing	The diagnostics have passed and the configuration is being downloaded.
Active	The active supervisor module and the switch are ready to be configured.
Standby	A switchover is possible.
Failed	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.
Offline	The supervisor module is intentionally shut down for debugging purposes.
At BIOS	The system has established connection with the supervisor and the supervisor module is performing diagnostics.
Unknown	The system is in an invalid state. If it persists, call TAC.

Table 4-3 lists the possible values for the supervisor module states.

Table 4-3 Supervisor States

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

Table 4-4 lists the possible values for the internal redundancy states.

Table 4-4 Internal States

State	Description
HA standby	The HA switchover mechanism in the standby supervisor module is enabled (see the "Manually Initiating a Switchover" section on page 4-6).
Active with no standby	A switchover is impossible.
Active with HA standby	The active supervisor module in the switch is ready to be configured. The standby supervisor module is in the ha-standby state.
Shutting down	The system is being shut down.
HA switchover in progress	The system is in the process of entering the active state.
Offline	The system is intentionally shut down for debugging purposes.
HA synchronization in progress	The standby supervisor module is in the process of synchronizing its state with the active supervisor modules.
Standby (failed)	The standby supervisor module is not functioning.

Table 4-4 Internal States (continued)

State	Description
Active with failed standby	The active supervisor module and the second supervisor module are present but the second supervisor module is not functioning.
Other	The system is in a transient state. If it persists, call TAC.

VDC High Availability

The Cisco NX-OS software incorporates high availability (HA) features that minimize the impact on the data plane if the control plane fails or a switchover occurs. The different HA service levels provide data plane protection, including service restarts, stateful supervisor module switchovers, and in-service software upgrades (ISSUs). All of these high availability features support VDCs.

If unrecoverable errors occur in a VDC, the Cisco NX-OS software provides HA policies that you can specify for each VDC. These HA policies include the following:

- Bringdown—Puts the VDC in the failed state. To recover from the failed state, you must reload the physical device. This is the behavior for default VDC. For non-default VDC, there is no need to reload the physical device.
- Reset— Initiates a supervisor module switchover for a Cisco NX-OS device with two supervisor modules, or reloads a Cisco NX-OS device with one supervisor module.
- Restart—Deletes the VDC and recreates it by using the startup configuration.

For details about VDCs and HA, see the *Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide, Release 4.1*.

Additional References

For additional information related to implementing system-level HA features, see the following sections:

Related Documents

Related Topic	Document Title
Virtual device context (VDC)	*Cisco Nexus 7000 Series NX-OS Virtual Device Context Configuration Guide, Release 4.1*
Redundant hardware	*Cisco Nexus 7000 Series Site Preparation Guide* and the *Cisco Nexus 7000 Series Hardware Installation and Reference Guide*
Power mode configuration and Cisco NX-OS fundamentals	*Cisco Nexus 7000 Series NX-OS Fundamentals Configuration Guide, Release 6.x*
Nonstop forwarding (NSF)	*Cisco Nexus 7000 Series NX-OS Unicast Routing Configuration Guide, Release 6.x*
In-service software upgrades (ISSU)	Chapter 5, "Understanding In-Service Software Upgrades"
GOLD, EEM, and Smart Call Home	*Cisco Nexus 7000 Series NX-OS System Management Configuration Guide, Release 6.x*
Licensing	*Cisco NX-OS Licensing Guide*

Standards

Standards	Title
No new or modified standards are supported by this feature, and support for existing standards has not been modified by this feature.	—

MIBs

MIBs	MIBs Link
• CISCO-SYSTEM-EXT-MIB: ciscoHaGroup, cseSwCoresTable, cseHaRestartNotify, cseShutDownNotify, cseFailSwCoreNotify, cseFailSwCoreNotifyExtended • CISCO-PROCESS-MIB • CISCO-RF-MIB	To locate and download MIBs, go to the following URL: http://www.cisco.com/public/sw-center/netmgmt/cmtk/mibs.shtml

RFCs

RFCs	Title
No RFCs are supported by this feature	—

Technical Assistance

Description	Link
Technical Assistance Center (TAC) home page, containing 30,000 pages of searchable technical content, including links to products, technologies, solutions, technical tips, and tools. Registered Cisco.com users can log in from this page to access even more content.	http://www.cisco.com/cisco/web/support/index.html