



Cisco Nexus 1000V for VMware vSphere High Availability and Redundancy Configuration Guide, Release 5.x

First Published: 2014-08-22

Last Modified: 2015-09-18

Americas Headquarters

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

THE SPECIFICATIONS AND INFORMATION REGARDING THE PRODUCTS IN THIS MANUAL ARE SUBJECT TO CHANGE WITHOUT NOTICE. ALL STATEMENTS, INFORMATION, AND RECOMMENDATIONS IN THIS MANUAL ARE BELIEVED TO BE ACCURATE BUT ARE PRESENTED WITHOUT WARRANTY OF ANY KIND, EXPRESS OR IMPLIED. USERS MUST TAKE FULL RESPONSIBILITY FOR THEIR APPLICATION OF ANY PRODUCTS.

THE SOFTWARE LICENSE AND LIMITED WARRANTY FOR THE ACCOMPANYING PRODUCT ARE SET FORTH IN THE INFORMATION PACKET THAT SHIPPED WITH THE PRODUCT AND ARE INCORPORATED HEREIN BY THIS REFERENCE. IF YOU ARE UNABLE TO LOCATE THE SOFTWARE LICENSE OR LIMITED WARRANTY, CONTACT YOUR CISCO REPRESENTATIVE FOR A COPY.

The Cisco implementation of TCP header compression is an adaptation of a program developed by the University of California, Berkeley (UCB) as part of UCB's public domain version of the UNIX operating system. All rights reserved. Copyright © 1981, Regents of the University of California.

NOTWITHSTANDING ANY OTHER WARRANTY HEREIN, ALL DOCUMENT FILES AND SOFTWARE OF THESE SUPPLIERS ARE PROVIDED "AS IS" WITH ALL FAULTS. CISCO AND THE ABOVE-NAMED SUPPLIERS DISCLAIM ALL WARRANTIES, EXPRESSED OR IMPLIED, INCLUDING, WITHOUT LIMITATION, THOSE OF MERCHANTABILITY, FITNESS FOR A PARTICULAR PURPOSE AND NON-INFRINGEMENT OR ARISING FROM A COURSE OF DEALING, USAGE, OR TRADE PRACTICE.

IN NO EVENT SHALL CISCO OR ITS SUPPLIERS BE LIABLE FOR ANY INDIRECT, SPECIAL, CONSEQUENTIAL, OR INCIDENTAL DAMAGES, INCLUDING, WITHOUT LIMITATION, LOST PROFITS OR LOSS OR DAMAGE TO DATA ARISING OUT OF THE USE OR INABILITY TO USE THIS MANUAL, EVEN IF CISCO OR ITS SUPPLIERS HAVE BEEN ADVISED OF THE POSSIBILITY OF SUCH DAMAGES.

Any Internet Protocol (IP) addresses and phone numbers used in this document are not intended to be actual addresses and phone numbers. Any examples, command display output, network topology diagrams, and other figures included in the document are shown for illustrative purposes only. Any use of actual IP addresses or phone numbers in illustrative content is unintentional and coincidental.

All printed copies and duplicate soft copies of this document are considered uncontrolled. See the current online version for the latest version.

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

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: [www.cisco.com go trademarks](http://www.cisco.com/go/trademarks). Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1721R)

© 2009–2019 Cisco Systems, Inc. All rights reserved.



CONTENTS

CHAPTER 1	New and Changed Information 1
	New and Changed Information for High Availability 1

CHAPTER 2	Overview 3
	Information About High Availability 3
	System Components 4
	Service-Level High Availability 4
	Isolation of Processes 4
	Process Restartability 5
	System-Level High Availability 5
	Network-Level High Availability 5
	VSM-to-VSM Heartbeats 5
	Control and Management Interface Redundancy 6
	Partial Communication 7
	Loss of Communication 7
	VSM-VEM Communication Loss 7
	One-Way Communication 8
	Split-Brain Resolution 8
	Checking the Accounting Logs and the Redundancy Traces 9
	VSM Role Collision Detection 9
	Displaying the Role Collision 10
	Enhancements for Domain ID Collision 11
	Displaying Domain ID Collision 11
	Recommended Reading 12

CHAPTER 3	Understanding Service-Level High Availability 13
------------------	---

Information About Cisco NX-OS Service Restarts	13
Restartability Infrastructure	13
System Manager	13
Persistent Storage Service	14
Message and Transaction Service	14
High Availability Policies	14
Process Restartability	15
Stateful Restarts	15
Stateless Restarts	16
Switchovers	16
Restarts on Standby Supervisor Services	16
Restarts on Switching Module Services	17
Troubleshooting Restarts	17
MIBs	17
RFCs	17
Technical Assistance	17

CHAPTER 4**Configuring System-Level High Availability 19**

Information About System-Level High Availability	19
Information About Single and Dual Supervisor Roles	19
HA Supervisor Roles	20
Dual Supervisor Active and Standby Redundancy States	21
Dual Supervisor Synchronization	21
Information About VSM Restarts and Switchovers	22
Restarts on Standalone VSMs	22
Restarts on Dual VSMs	22
Switchovers on Dual VSMs	22
Switchover Characteristics	22
Automatic Switchovers	22
Manual Switchovers	22
Guidelines and Limitations	23
Configuring System-Level High Availability	23
Changing the VSM Role	23
Configuring a Switchover	25

Guidelines and Limitations for Configuring a Switchover	25
Verifying that a System is Ready for a Switchover	25
Manually Switching the Active VSM to Standby	25
Configuring the VSM-to-VSM Heartbeat Interval	27
Adding a Second VSM to a Standalone System	27
Adding a Second VSM to a Standalone System	27
Changing the Standalone VSM to a Primary VSM	28
Verifying the Change to a Dual VSM System	29
Replacing the Secondary VSM in a Dual VSM System	29
Replacing the Primary VSM in a Dual VSM System	30
Changing the Domain ID in a Dual VSM System	31
Changing the Domain ID in a Dual VSM System for VSMs Hosted on Cisco Nexus 1010	32
Disabling Domain ID Collision	35
Verifying the HA Status	36
Related Documents	37
Standards	37
MIBs	37
RFCs	37
Technical Assistance	37
Feature History for System-Level High Availability	38



CHAPTER 1

New and Changed Information

This chapter lists new and changed content in this document by software release.

- [New and Changed Information for High Availability, on page 1](#)

New and Changed Information for High Availability

This section lists new and changed content in this document by software release.

To find additional information about new features or command changes, see the *Cisco Nexus 1000V Release Notes* and *Cisco Nexus 1000V Command Reference*.

Table 1: New and Changed Features for the Cisco Nexus 1000V High Availability Configuration Guide

Feature	Description	Changed in Release	Where Documented
Enhancements for Domain ID Collision	The high availability functionality on Cisco Nexus 1000V is enhanced to detect domain ID collisions and keep high availability (HA) intact. Any rogue VSM with the same domain ID will not affect the HA.	5.2(1)SV3(1.1)	Enhancements for Domain ID Collision, on page 11



CHAPTER 2

Overview

This chapter contains the following sections:

- [Information About High Availability, on page 3](#)
- [System Components, on page 4](#)
- [Service-Level High Availability, on page 4](#)
- [System-Level High Availability, on page 5](#)
- [Network-Level High Availability, on page 5](#)
- [VSM-to-VSM Heartbeats, on page 5](#)
- [Split-Brain Resolution, on page 8](#)
- [Checking the Accounting Logs and the Redundancy Traces, on page 9](#)
- [VSM Role Collision Detection, on page 9](#)
- [Displaying the Role Collision, on page 10](#)
- [Enhancements for Domain ID Collision, on page 11](#)
- [Displaying Domain ID Collision, on page 11](#)
- [Recommended Reading, on page 12](#)

Information About High Availability

The purpose of high availability (HA) is to limit the impact of failures—both hardware and software—within a system. The Cisco NX-OS operating system is designed for high availability at the network, system, and service levels.

The following Cisco NX-OS features minimize or prevent traffic disruption in the event of a failure:

- **Redundancy**—Redundancy at every aspect of the software architecture.
- **Isolation of processes**—Isolation between software components to prevent a failure within one process disrupting other processes.
- **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**—Active/standby dual supervisor configuration. The state and configuration remain constantly synchronized between two Virtual Supervisor Modules (VSMs) to provide a seamless and stateful switchover in the event of a VSM failure.

Starting with Release 4.2(1)SV2(1.1), the high availability functionality is enhanced to support the split active and standby Cisco Nexus 1000V Virtual Supervisor Modules (VSMs) across two data centers to implement the cross-DC clusters and the VM mobility while ensuring high availability.

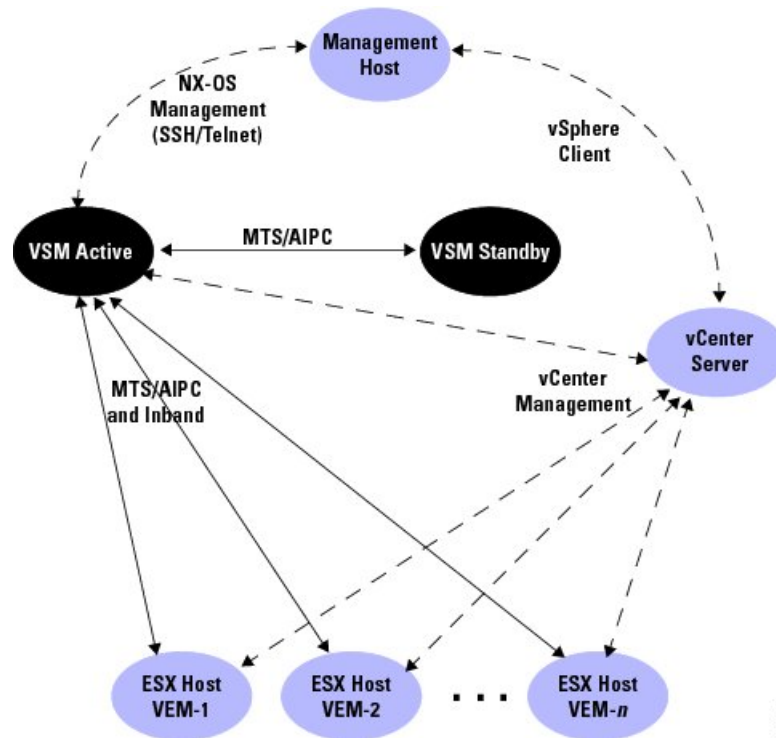
System Components

The Cisco Nexus 1000V system is made up of the following:

- One or two VSMs that run within Virtual Machines (VMs).
- Virtual Ethernet Modules (VEMs) that run within virtualization servers. VEMs are represented as modules within the VSM.
- A remote management component. VMware vCenter Server.

The following figure shows the HA components and the communication links between them.

Figure 1: HA Components and Communication Links



Service-Level High Availability

Isolation of Processes

The Cisco NX-OS software has independent processes, known as services, that 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 way of operating provides a highly fault-tolerant software infrastructure and fault isolation between services. A failure in a service instance does not affect any other services that are running at that time. Additionally, each instance of a service can run as an independent process, which means that two instances of a routing protocol can run as separate processes.

Process Restartability

Cisco NX-OS processes run in a protected memory space independently of 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. These stateful restarts 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 1000V supports redundant VSM virtual machines—a primary and a secondary—running as an HA pair. Dual VSMs operate in an active/standby capacity in which only one of the VSMs is active at any given time, while the other acts as a standby backup. The VSMs are configured as either primary or secondary as a part of the Cisco Nexus 1000V installation.

The state and configuration remain constantly synchronized between the two VSMs to provide a stateful switchover if the active VSM fails.

Network-Level High Availability

The Cisco Nexus 1000V high availability at the network level includes port channels and the Link Aggregation Control Protocol (LACP). A port channel bundles physical links into a channel group to create a single logical link that provides the aggregate bandwidth of up to eight physical links. If a member port within a port channel fails, the traffic previously carried over the failed link switches to the remaining member ports within the port channel.

Additionally, LACP allows you to configure up to 16 interfaces into a port channel. A maximum of eight interfaces can be active, and a maximum of eight interfaces can be placed in a standby state.

VSM-to-VSM Heartbeats

The primary and secondary VSMs use a VSM-to-VSM heartbeat to do the following within their domain:

- Broadcast their presence
- Detect the presence of another VSM
- Negotiate active and standby redundancy states

When a VSM first boots up, it broadcasts discovery frames to the domain to detect the presence of another VSM. If no other VSM is found, the booting VSM becomes active. If another VSM is found to be active, the booting VSM becomes the standby VSM. If another VSM is found to be initializing (for example, during a system reload), the primary VSM has priority over the secondary to become the active VSM.



Note Starting with Release 5.2(1)SV3(1.1) and later releases, the VSM validates the source MAC address of the high availability (HA) packets that it receives on control and management interfaces. During initial contact, the VSM learns the peer VSM MAC addresses and stores them in a permanent location. Only the HA packets that are learned from the VSM are accepted.

After the initial contact and role negotiation, the active and standby VSMs unicast the following in heartbeat messages:

- Redundancy state
- Control flags requesting action by the other VSM

The following intervals apply when sending heartbeat messages.

Interval	Description
Inter-VSM heartbeat Default: 1 second	Interval at which heartbeat requests are sent.
Inter-VSM maximum heartbeat loss Range : 6-30 seconds Default : 15 seconds	<p>At the active VSM</p> <ul style="list-style-type: none"> • Half of the inter VSM maximum heartbeat loss is the interval after which missed heartbeats indicate degraded communication on the control interface so that heartbeats are also sent on the management interface. This is known as degraded mode. • Inter-VSM maximum heartbeat loss is the interval after which communication over the control interface with the standby VSM is considered down. <p>At the standby VSM</p> <ul style="list-style-type: none"> • Inter-VSM maximum heartbeat loss is the interval after if no communication is received from the active VSM (over the control or management interface), the standby VSM interprets it as an active VSM failure and it becomes active.
Varies	The standby VSM is reset by the active VSM when the active VSM is no longer able to synchronize with it, which means that the interval varies depending on how long the active VSM can buffer the data to be synchronized when the communication is interrupted.

Control and Management Interface Redundancy

The VSM communicates with the peer VSM over layer 2 only on the control and management interfaces. If the active VSM does not receive a heartbeat response over the control interface for a period of half of the inter-VSM maximum heartbeat loss interval (eight heartbeats by default), communication is seen as degraded and the VSM begins sending requests over the management interface in addition to the control interface. In this case, the management interface provides redundancy by preventing both VSMs from becoming active. This process is called an active-active or split-brain situation.



Note The communication is not fully redundant, however, because the management interface only handles heartbeat requests and responses.

AIPC and the synchronization of data between VSMs is done through the control interface only.

Partial Communication

The secondary VSM is not immediately rebooted when communication over the control interface is interrupted because the HA mechanism tolerates brief interruptions in communication. When communication is first interrupted on the control interface, the heartbeat messages are sent over the management interface. If communication over the management interface is successful, the VSMs enter into a degraded mode, as displayed in the **show system internal redundancy trace** command output. If communication is interrupted on both interfaces for too long, the two VSMs get out of synchronization and the standby VSM is forced to reboot.



Note A transition from active to standby always requires a reload in both the Cisco Nexus 1000V and the Cisco Nexus Cloud Services Platform.

Loss of Communication

When there is no communication between redundant VSMs or Cisco Nexus Cloud Services Platforms, they cannot detect the presence of the other. The standby VSM will be removed from the list of inserted modules at the active VSM. The standby interprets the lack of heartbeats as a sign that the active has failed and it also becomes active. This process is what is referred to as active-active or split-brain, as both are trying to control the system by connecting to vCenter and communicating with the VEMs.

Because redundant VSMs or Cisco Nexus Cloud Services Platforms use the same IP address for their management interface, remote Secure Shell (SSH)/Telnet connections might fail, as a result of the path to this IP address changing in the network. For this reason, we recommend that you use the consoles during a split-brain conflict.

The following parameters are used to select the VSM to be rebooted during the split-brain resolution: the module count, the vCenter Server connectivity status, the last configuration time, and the last active time.

VSM-VEM Communication Loss

Depending on the specific network failure that caused it, each VSM might reach a different, possibly overlapping, subset of Virtual Ethernet Modules (VEMs). When the VSM that was in the standby state becomes a new active VSM, it broadcasts a request to all VEMs to switch to it as the current active device. Whether a VEM switches to the new active VSM, depends on the following:

- The connectivity between each VEM and the two VSMs.
- Whether the VEM receives the request to switch.

A VEM remains attached to the original active VSM even if it receives heartbeats from the new active VSM. However, if the VEM also receives a request to switch from the new active VSM, it detaches from the original active VSM and attaches to the new VSM.

If a VEM loses connectivity to the original active device and only receives heartbeats from the new one, it ignores those heartbeats until it goes into headless mode, which occurs approximately 15 seconds after it stops receiving heartbeats from the original, active VSM. At that point, the VEM attaches to the new active VSM if it has connectivity to it.



Note If a VEM loses the connection to its VSM, VMotionsthat particular VEM is blocked. The VEM shows vCenter Server a degraded (yellow) status.

One-Way Communication

If a network communication failure occurs where the standby VSM receives heartbeat requests but the active VSM does not receive a response, the following occurs:

- The active VSM declares that the standby VSM is not present.
- The standby VSM remains in a standby state and continues receiving heartbeats from the active VSM.

In this scenario, the redundancy state is inconsistent (**show system redundancy state**) and the two VSMs lose synchronization. When two-way communication is resumed, the standby VSM replies to the active VSM and asks to be reset.



Note If a one-way communication failure occurs in the active to standby direction, it is equivalent to a total loss of communication because a standby VSM sends heartbeats only in response to active VSM requests.

Split-Brain Resolution

When the connectivity between two Virtual Supervisor Modules (VSMs) is broken, this loss of communication can cause both VSMs to take the active role. This condition is called active-active or split-brain condition. When the communication is restored between the VSMs, both VSMs exchange information to decide which one would have a lesser impact on the system, if rebooted.

Both primary and secondary VSMs process the same data to select the VSM (primary or secondary) that needs to be rebooted. When the selected VSM is rebooted and attaches itself back to the system, high availability is back to normal. The VSM uses the following parameters in order of precedence to select the VSM to be rebooted during the split-brain resolution:

1. Module count—The number of modules that are attached to the VSM.
2. vCenter status— Status of the connection between the VSM and vCenter.
3. Last configuration time—The time when the last configuration is done on the VSM.
4. Last standby-active switch—The time when the VSM last switched from the standby state to the active state. (The VSM with a longer active time gets higher priority.)

Checking the Accounting Logs and the Redundancy Traces

During the split-brain resolution, when a VSM reboots, the accounting logs that are stored on the VSM are lost. Starting with Release 4.2(1)SV2(1.1), you can display the accounting logs that were backed up during the split-brain resolution. You can also check the redundancy traces that are stored on the local and remote VSMs.

Command	Purpose
switch# show system internal active-active accounting logs	Displays the accounting logs that are stored on a local VSM during the last split-brain resolution.
switch# show system internal active-active redundancy traces	Displays the redundancy traces that are stored on a local VSM during the last split-brain resolution.
switch# show system internal active-active remote accounting logs	Displays the remote accounting logs that are stored on a remote VSM during the last split-brain resolution.
switch# show system internal active-active remote redundancy traces	Displays the remote redundancy traces that are stored on a remote VSM during the last split-brain resolution.
switch# clear active-active accounting logs	Clears the accounting logs that are stored on a local VSM during the split-brain resolution.
switch# clear active-active redundancy traces	Clears the redundancy traces that are stored on a local VSM during the split-brain resolution.
switch# clear active-active remote accounting logs	Clears the remote accounting logs that are stored on a remote VSM during the split-brain resolution.
switch# clear active-active remote redundancy traces	Clears the remote redundancy traces that are stored on a remote VSM during the split-brain resolution.

VSM Role Collision Detection

In the Cisco Nexus 1000V, if a secondary VSM is configured or installed with the same role as the primary VSM and with the same domain ID, the secondary VSM and the primary VSM exchange heartbeats to discover each other. Both VSMs detect and report a role collision when they exchange heartbeats. When a collision is detected, the VSMs report the MAC address of the VSM with which the local VSM is colliding.

Due to this issue, the HA-paired VSM cannot communicate with the correct VSM. This problem can occur on a primary VSM or a secondary VSM depending on whether the newly configured or the installed VSM has the primary or the secondary role assigned to it.

The collisions are detected on the control and the management interfaces. The maximum number of colliding VSMs reported is eight.



Note

After the eighth role collision, the problem is still logged and the MAC address entry is overwritten. The **show system redundancy status** command displays the overwrite details.



Note The colliding VSMS might also report a collision detection from the original VSM. If the colliding VSMS use the same IP address for their management interfaces, the remote SSH/Telnet connections might fail. Therefore, we recommend that you use the consoles during a role collision detection.

Enter the **show system redundancy status** command on both the primary and secondary VSM consoles to display the MAC addresses of the detected VSMS with the same role and domain ID, if any. When the VSM stops communicating in the domain, the collision time is not updated anymore. After an hour elapses since the last collision, the collision MAC entries are removed.

Displaying the Role Collision

Use the **show system redundancy status** CLI command to display the VSM role collision:

Command	Purpose
n1000v# show system redundancy status	Displays a detected role collision. A warning is highlighted in the CLI output. Along with the MAC addresses, the latest collision time is also displayed in the output. If no collisions are detected, the highlighted output does not appear.

Example

This example shows how to display the detected traffic collision:

```
n1000v# show system redundancy status

Redundancy role
-----
      administrative:  secondary
      operational:    secondary

Redundancy mode
-----
      administrative:  HA
      operational:    HA

This supervisor (sup-2)
-----
      Redundancy state:  Active
      Supervisor state:  Active
      Internal state:    Active with HA standby

Other supervisor (sup-1)
-----
      Redundancy state:  Standby
      Supervisor state:  HA standby
      Internal state:    HA standby

WARNING! Conflicting sup-2(s) detected in same domain
-----
      MAC                Latest Collision Time
00:50:56:97:02:3b      2012-Sep-11 18:59:17
00:50:56:97:02:3c      2012-Sep-11 18:59:17
```



```

00:50:56:97:02:2f      2012-Sep-11 18:57:42
00:50:56:97:02:35      2012-Sep-11 18:57:46
00:50:56:97:02:29      2012-Sep-11 18:57:36
00:50:56:97:02:30      2012-Sep-11 18:57:42
00:50:56:97:02:36      2012-Sep-11 18:57:46
00:50:56:97:02:2a      2012-Sep-11 18:57:36

```

NOTE: Please run the same command on sup-1 to check for conflicting(if any) sup-1(s) in the same domain.

Enhancements for Domain ID Collision

If you configure or install a VSM with the same domain ID when a VSM pair is already in place, a domain ID collision occurs. In Releases earlier than 5.2(1)SV3(1.1), this collision breaks high availability (HA) between the existing HA pair. In Release 5.2(1)SV3(1.1), the rogue VSM does not affect the existing VSM HA pair because the source MAC addresses of the VSM HA packets are validated and the rogue VSM packets are dropped.

By default, the VSM domain ID collision detection feature is enabled.

This example shows how to verify if the check is enabled or disabled:

```

switch# show peer-sup mac-addresses details
Peer MAC Address Check = Enabled
Peer HA0 MAC Address = 00:50:56:b5:3a:99
Peer HA1 MAC Address = 00:50:56:b5:5e:05
switch#

```

Displaying Domain ID Collision

Use the **show system redundancy status** CLI command to display the VSM domain ID collision:

Command	Purpose
switch# show system redundancy status	Displays a detected domain ID collisions. In the "Peer Sup Mac Addresses Learnt" section of the output, the peer VSM's control and management MAC addresses are displayed. The "HA Packet Drops Due to Domain id Collision" section is displayed only if a domain ID collision occurs. This section also displays the number of packets that the VSM has dropped on both the control and management interfaces.

Example

This example shows how to display the detected domain ID collision:

```

switch# show system redundancy status

Redundancy role
-----
      administrative:  primary

```

```
operational: primary

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

Peer Sup Mac Addresses Learnt
-----
Control Interface: 00:50:56:91:44:c8
Mgmt Interface: 00:50:56:91:1f:6f

HA Packet Drops Due to Domain id Collision
-----
Control Interface: 109
Mgmt Interface: 93
```

Recommended Reading

- *Cisco Nexus 1000V Installation and Upgrade Guide*
- *Cisco Nexus 1000V Port Profile Configuration Guide*



CHAPTER 3

Understanding Service-Level High Availability

This chapter contains the following sections:

- [Information About Cisco NX-OS Service Restarts, on page 13](#)
- [Restartability Infrastructure, on page 13](#)
- [Process Restartability, on page 15](#)
- [Restarts on Standby Supervisor Services , on page 16](#)
- [Restarts on Switching Module Services, on page 17](#)
- [Troubleshooting Restarts, on page 17](#)
- [MIBs, on page 17](#)
- [RFCs, on page 17](#)
- [Technical Assistance, on page 17](#)

Information About Cisco NX-OS Service Restarts

The Cisco NX-OS service restart feature restarts a faulty service without restarting the supervisor to prevent process-level failures from causing system-level failures. A service is restarted 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.

Restartability Infrastructure

Cisco NX-OS allows stateful restarts of most processes and services. The 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 the 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 and for initiating and managing the synchronization of service states and supervisor states for stateful switchovers.

The System Manager directs the overall system function, service management, and system health monitoring. The System Manager is responsible for launching, stopping, monitoring, and restarting services.

The System Manager directs the 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 and for initiating and managing the synchronization of service states and supervisor states for stateful switchovers.

Persistent Storage Service

Cisco NX-OS services use the persistent storage service (PSS) to store and manage the operational run-time information and configuration of platform services. 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 one of the following:

- Private—It can be read only by that service.
- Shared—The information can be read by other services.

The service can specify that it is one of the following:

- Local—The information can be read only by services on the same supervisor.
- Global—It can be read by services on either 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.

High Availability Policies

Cisco NX-OS allows each service to have an associated set of internal HA policies that define how a failed service will be 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 will be 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**—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, which results in a supervisor switchover or restart.
- **Minimum lifetime**—The time that a service must run after a restart attempt in order to consider the restart attempt as successful. The minimum lifetime is no less than four minutes.

Process Restartability

Cisco NX-OS processes run in a protected memory space independently of 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. These stateful restarts 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.

A failed service is restarted by different methods depending on the service's HA implementation and HA policies.

The following table describes the action taken by the System Manager for various failure conditions.

Failure	Action
Service/process exception	Service restart
Service/process crash	Service restart
Unresponsive service/process	Service restart
Repeated service failure	Supervisor reset (single) or switchover (dual)
Unresponsive System Manager	Supervisor reset (single) or switchover (dual)
Kernel failure	Supervisor reset (single) or switchover (dual)
Watchdog timeout	Supervisor reset (single) or switchover (dual)

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 the 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 executes the action specified by the service's HA policy, 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 will collect 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.

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, so 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.

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.

In a system with dual VSMSs, after a switchover, the active supervisor resets and comes back up as a standby supervisor.

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 overwhelmed 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.

Restarts on Switching Module Services

Service failures on nonsupervisor module services do not require a supervisor switchover.

On the VEMs, the Data Path Agent (DPA) is restarted if it crashes. This situation causes the module to be removed and readded on the VSM.

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 SNMP traps are enabled, the SNMP agent sends a trap when a service is restarted
- When a service failure occurs on a VSM, the event is logged. To view the log, use 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 entering 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 using a file transfer utility such as the Trivial File Transfer Protocol (TFTP).

MIBs

MIBs	MIBs Link
CISCO-PROCESS-MIB	To locate and download MIBs, go to the following URL: http://www.cisco.com/public/sw-center/netmgmt/cmtk/mibs.shtml

RFCs

No RFCs are supported by this feature.

Technical Assistance

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.

Go to the following URL: <http://www.cisco.com/cisco/web/support/index.html>



CHAPTER 4

Configuring System-Level High Availability

This chapter contains the following sections:

- [Information About System-Level High Availability, on page 19](#)
- [Information About VSM Restarts and Switchovers, on page 22](#)
- [Guidelines and Limitations, on page 23](#)
- [Configuring System-Level High Availability, on page 23](#)
- [Adding a Second VSM to a Standalone System, on page 27](#)
- [Replacing the Secondary VSM in a Dual VSM System, on page 29](#)
- [Replacing the Primary VSM in a Dual VSM System, on page 30](#)
- [Changing the Domain ID in a Dual VSM System, on page 31](#)
- [Changing the Domain ID in a Dual VSM System for VSMs Hosted on Cisco Nexus 1010, on page 32](#)
- [Disabling Domain ID Collision , on page 35](#)
- [Verifying the HA Status, on page 36](#)
- [Related Documents, on page 37](#)
- [Standards, on page 37](#)
- [MIBs, on page 37](#)
- [RFCs, on page 37](#)
- [Technical Assistance, on page 37](#)
- [Feature History for System-Level High Availability, on page 38](#)

Information About System-Level High Availability

Information About Single and Dual Supervisor Roles

The Cisco Nexus 1000V can be configured with a single Virtual Supervisor Module (VSM) or dual VSMs. The following table describes the HA supervisor roles for single and dual VSM operation.

Single VSM Operation	Dual VSM Operation
<ul style="list-style-type: none"> • Stateless—In case of failure, service restarts from the startup configuration. • Stateful—In case of failure, service resumes from previous state. 	<ul style="list-style-type: none"> • Redundancy is provided by one active VSM and one standby VSM. • The active VSM runs all the system applications and controls the system. • On the standby VSM, the applications are started and initialized in standby mode. The applications are synchronized and kept up to date with the active VSM in order to be ready to run. • On a switchover, the standby VSM takes over for the active VSM. • The control interface of the VSMs are used to pass heartbeats between the two VSMs. • The management interface is used to prevent split-brain scenarios.

HA Supervisor Roles

The redundancy role indicates not only whether the VSM interacts with other VSMs, but also the module number it occupies. The following table shows the available HA roles for VSMs.

role	Module Number	Description
Standalone	1	<ul style="list-style-type: none"> • This role does not interact with other VSMs. • You assign this role when there is only one VSM in the system. • This role is the default.
Primary	1	<ul style="list-style-type: none"> • This role coordinates the active/standby state with the secondary VSM. • This role takes precedence during bootup when negotiating active/standby mode. That is, if the secondary VSM does not have the active role at bootup, the primary VSM takes the active role. • You assign this role to the first VSM that you install in a dual VSM system.
Secondary	2	<ul style="list-style-type: none"> • This role coordinates the active/standby state with the primary VSM. • You assign this role to the second VSM that you install in a dual VSM system.

Dual Supervisor Active and Standby Redundancy States

Independent of its role, the redundancy state of a VSM can be one of the following described in this table.

Redundancy State	Description
Active	Controls the system and is visible to the outside world.
Standby	Synchronizes its configuration with that of the active VSM so that it is continuously ready to take over in case of a failure or manual switchover. You cannot use Telnet or Secure Shell (SSH) protocols to communicate with the standby VSM. Instead, you can use the attach module command from the active VSM to access the standby VSM console. Only a subset of the CLI commands are available from the standby VSM console.

Dual Supervisor Synchronization

The active and standby VSMs are in the operationally HA state and can automatically synchronize when the internal state of one supervisor module is Active with HA Standby and the internal state of the other supervisor module is HA Standby.

If the output of the **show system redundancy** command indicates that the operational redundancy mode of the active VSM is None, the active and standby VSMs are not yet synchronized.

This example shows the VSM internal state of dual supervisors as observed in the output of the **show system redundancy status** command:

```
switch# show system redundancy status
Redundancy role
-----
      administrative:  standalone
      operational:    standalone

Redundancy mode
-----
      administrative:  HA
      operational:    None

This supervisor (sup-1)
-----
      Redundancy state:  Active
      Supervisor state:  Active
      Internal state:    Active with no standby

Other supervisor (sup-2)
-----
      Redundancy state:  Not present
switch#
```

Information About VSM Restarts and Switchovers

Restarts on Standalone VSMs

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 restart with no prior state information.

Restarts on Dual VSMs

When a VSM fails in a system with dual supervisors, the system performs a switchover rather than a system restart in order to maintain a stateful operation. In some cases, a switchover might not be possible at the time of the failure. For example, if the standby VSM is not in a stable standby state, a restart rather than a switchover is performed.

Switchovers on Dual VSMs

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

Switchover Characteristics

A switchover occurs when the active supervisor fails (for example, if repeated failures occur in an essential service or if the system that is hosting the VSM fails).

A user-triggered switchover could occur (for example, if you need to perform maintenance tasks on the system hosting the active VSM).

An HA switchover has the following characteristics:

- It is stateful (nondisruptive) because the control traffic is not affected.
- It does not disrupt data traffic because the VEMs are not affected.

Automatic Switchovers

When a stable standby VSM detects that the active VSM has failed, it initiates a switchover and transitions to active. When a switchover begins, another switchover cannot be started until a stable standby VSM is available.

If a standby VSM that is not stable detects that the active VSM has failed, then, instead of initiating a switchover, it tries to restart the system.

Manual Switchovers

Before you can initiate a manual switchover from the active to the standby VSM, the standby VSM must be stable.

Once you have verified that the standby VSM is stable, you can manually initiate a switchover.

Once a switchover process begins, another switchover process cannot be started until a stable standby VSM is available.

Guidelines and Limitations

- Although primary and secondary VSMs can reside in the same host, to improve redundancy, install them in separate hosts and, if possible, connect the VSMs to different upstream switches.
- The console for the standby VSM is available through the vSphere client or by entering the **module attach x** command, but configuration is not allowed and many commands are restricted. Enter this command at the console of the active VSM.
- You cannot use Telnet or Secure Shell (SSH) protocols to communicate with the standby VSM because the management interface IP is unconfigured until the VSM becomes active.
- The active and standby VSMs must be on the same management subnet.

Configuring System-Level High Availability

Changing the VSM Role

The Cisco Nexus 1000V VSM software installation provides an opportunity for you to designate the role for each VSM. You can use this procedure to change that initial configuration.



Caution

Changing the role of a VSM can result in a conflict between the VSM pair. If a primary and secondary VSM see each other as active at the same time, the system resolves this problem by resetting the primary VSM.

Use this procedure to change the role of a VSM to one of the following after it is already in service:

- Standalone
- Primary
- Secondary

Before you begin

- Log in to the CLI in EXEC mode.
- If you are changing a standalone VSM to a secondary VSM, be sure to first isolate it from the other VSM in the pair to prevent any interaction with the primary VSM during the change. Power the VM off from the vSphere Client before reconnecting it as standby.

For an example on how to change the port groups and port profiles assigned to the VSM interfaces in the vSphere Client, see the *Cisco Nexus 1000V Installation and Upgrade Guide*.

You must understand the following information:

- The possible HA roles are standalone, primary, and secondary.

- The possible HA redundancy states are active and standby.
- To activate a change from primary to secondary VSM, you must reload the VSM by doing one of the following:
 - Enter the **reload** command.
 - Power the VM off and then on from the vSphere Client.
- A change from a standalone to a primary VSM takes effect immediately.

Procedure

	Command or Action	Purpose
Step 1	switch# system redundancy role { <i>standalone</i> <i>primary</i> <i>secondary</i> }	Designates the HA role of the VSM.
Step 2	(Optional) switch# show system redundancy status	Displays the current redundancy status for the VSMs.
Step 3	(Optional) switch# copy running-config startup-config	Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration.

Example

This example shows how to change the VSM role:

```
switch# system redundancy role standalone
switch# show system redundancy status
Redundancy role
-----
      administrative:  standalone
      operational:    standalone

Redundancy mode
-----
      administrative:  HA
      operational:    None

This supervisor (sup-1)
-----
      Redundancy state: Active
      Supervisor state: Active
      Internal state:Active with no standby

Other supervisor (sup-2)
-----
      Redundancy state:  Not present
switch#
```

Configuring a Switchover

Guidelines and Limitations for Configuring a Switchover

- When you manually initiate a switchover, system messages are generated that indicate the presence of two VSMs and identify which one is becoming active.
- A switchover can only be performed when both VSMs are functioning.

Verifying that a System is Ready for a Switchover

Use one of the following commands to verify the configuration:

Command	Purpose
show system redundancy status	<p>Displays the current redundancy status for the VSM(s).</p> <p>If the output indicates the following, you can proceed with a system switchover:</p> <ul style="list-style-type: none"> • The presence of an active VSM • The presence of a standby VSM in the HA standby redundancy state
show module	<p>Displays information about all available VEMs and VSMs in the system.</p> <p>If the output indicates the following, you can proceed with a system switchover:</p> <ul style="list-style-type: none"> • The presence of an active VSM • The presence of a standby VSM in the HA standby redundancy state

Manually Switching the Active VSM to Standby

Be sure you know the following about manually switching the active VSM to a standby VSM:

- A switchover can be performed only when two VSMs are functioning in the switch.
- If the standby VSM is not in a stable state (ha-standby), you cannot initiate a manual switchover and will see the following error message:

```
Failed to switchover (standby not ready to takeover in vdc 1)
```

- If a switchover does not complete successfully within 28 seconds, the supervisors reset.

Before you begin

- Log in to the active VSM CLI in EXEC mode.

- Complete the steps in [Verifying that a System is Ready for a Switchover, on page 25](#) and verify that the system is ready for a switchover.

Procedure

	Command or Action	Purpose
Step 1	switch# system switchover	On the active VSM, initiates a manual switchover to the standby VSM. Once you enter this command, you cannot start another switchover process on the same system until a stable standby VSM is available. Before proceeding, wait until the switchover completes and the standby supervisor becomes active.
Step 2	(Optional) switch# show running-config diff	Verifies the difference between the running and startup configurations. Any unsaved running configuration in an active VSM is also unsaved in the VSM that becomes active after switchover. Save that configuration in the startup if needed.
Step 3	(Optional) switch# copy running-config startup-config	Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration.

Example

This example shows how to switch an active VSM to the standby VSM and displays the output that appears on the standby VSM as it becomes the active VSM:

```
switch# system switchover
-----
2009 Mar 31 04:21:56 n1000v %% VDC-1 %% %SYSMGR-2-HASWITCHOVER_PRE_START:
This supervisor is becoming active (pre-start phase).
2009 Mar 31 04:21:56 n1000v %% VDC-1 %% %SYSMGR-2-HASWITCHOVER_START:
This supervisor is becoming active.
2009 Mar 31 04:21:57 n1000v %% VDC-1 %% %SYSMGR-2-SWITCHOVER_OVER: Switchover completed.
2009 Mar 31 04:22:03 n1000v %% VDC-1 %% %PLATFORM-2-MOD_REMOVE: Module 1 removed (Serial
number )
```

This example shows how to display the difference between the running and startup configurations:

```
switch# show running-config diff
*** Startup-config
--- Running-config
*****
*** 1,38 ****
version 4.0(4)SV1(1)
role feature-group name new
role name testrole
username admin password 5 $1$S7HvKc5G$aguYqH10dPtTBJAhEPwys1 role network-admin
```



```
telnet server enable
ip domain-lookup
```

Configuring the VSM-to-VSM Heartbeat Interval

If the communication network between two VSMs in an HA pair experiences interruptions longer than the 15-second default, you can change the VSM-to-VSM heartbeat interval so that it is less sensitive to falsely detecting active VSM failures.

Procedure

	Command or Action	Purpose
Step 1	switch# configure terminal	Enters global configuration mode.
Step 2	switch(config)# system inter-sup-heartbeat time <i>time-interval</i>	Configures the VSM-to-VSM heartbeat with a time interval from 6 to 30 seconds. The default is 15 seconds.
Step 3	switch(config)# show running-config grep heartbeat	Displays the heartbeat interval setting in the running configuration.
Step 4	(Optional) switch(config)# copy running-config startup-config	Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration.

Example

The following example shows how to set the VSM-to-VSM interval to 10 seconds and verify this setting:

```
switch# configure terminal
switch(config)# system inter-sup-heartbeat time 10
switch(config)# show running-config | grep heartbeat
system inter-sup-heartbeat time 10
switch(config)# copy running-config startup-config
```

Adding a Second VSM to a Standalone System

Adding a Second VSM to a Standalone System

The following list is designed to guide you through the process of adding a second VSM to a standalone system.

1. Change the standalone VSM to a primary VSM. For more information, see [Changing the Standalone VSM to a Primary VSM, on page 28](#).
2. Install the second VSM. For more information, see http://www.cisco.com/en/US/products/ps9902/prod_installation_guides_list.html

- Verify the change to the dual VSM system. For more information, see [Verifying the Change to a Dual VSM System, on page 29](#).

Changing the Standalone VSM to a Primary VSM

You can change the role of a VSM from standalone in a single VSM system to primary in a dual VSM system. A change from a standalone to a primary VSM takes effect immediately.

Before you begin

Log in to the CLI in EXEC mode.

Procedure

	Command or Action	Purpose
Step 1	switch# system redundancy role primary	Changes the standalone VSM to a primary VSM. The role change occurs immediately.
Step 2	switch# peer-sup mac-address clear	Clears old peer VSM MAC addresses, if any.
Step 3	(Optional) switch# show system redundancy status	It also displays peer VSM MAC addresses as not learned.
Step 4	(Optional) switch# copy running-config startup-config	Saves the change persistently through reboots and restarts by copying the running configuration to the startup configuration.

Example

This example shows how to display the current system redundancy status for the VSM:

```
switch# system redundancy role primary
switch# peer-sup mac-address clear
switch# show system redundancy status
Redundancy role
-----
      administrative:  primary
      operational:    primary

Redundancy mode
-----
      administrative:  HA
      operational:    None

This supervisor (sup-1)
-----
      Redundancy state:  Active
      Supervisor state:  Active
      Internal state:    Active with no standby

Other supervisor (sup-2)
-----
      Redundancy state:  Not present
```

```
Peer Sup Mac Addresses Learnt
-----
Control Interface:  Not Learnt
Mgmt Interface:    Not Learn

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

Verifying the Change to a Dual VSM System

Use one of the following commands to verify the configuration:

Command	Purpose
show system redundancy status	Displays the current redundancy status for VSMs in the system.
show module	Displays information about all available VSMs and VEMs in the system.

Replacing the Secondary VSM in a Dual VSM System



Note Equipment Outage—This procedure requires that you power down and reinstall a VSM. During this time, your system will operate with a single VSM.

If have your VSM HA Pair in 5.2(1)SV3(1.1) with Domain Id > 1023 [Setup created Via upgrade] , then kindly change your domain ID to less than 1024 before replacing Primary or Secondary VSM in an Dual VSM system.

Procedure

- Step 1** Ensure that the primary VSM is active (see the output of the **show system redundancy status** command).
- Note** Do a system switchover if necessary.
- Step 2** Power off the secondary VSM.
- Step 3** Log into the CLI in EXEC mode on the active primary VSM. Enter **peer-sup mac-addresses-clear** command.
- Step 4** Verify that the "Peer Sup Mac Addresses Learnt" section in the **show system redundancy status** command displays "Not Learnt."
- Step 5** Install the new VSM as a secondary, with the same domain ID as the existing VSM, using the procedure in the "Installing and Configuring the VSM VM" section in the *Cisco Nexus 1000V Installation and Upgrade Guide*.

After the new VSM is added to the system, both VSMs learn the peer VSM MAC addresses and the new VSM synchronizes with the existing VSM.

Replacing the Primary VSM in a Dual VSM System

You can replace an active/primary VSM in a dual VSM system.



Note Equipment Outage—This procedure requires that you power down and reinstall a VSM. During this time, your system will operate with a single VSM.

If have your VSM HA Pair in 5.2(1)SV3(1.1) with Domain Id > 1023 [Setup created Via upgrade] , then kindly change your domain ID to less than 1024 before replacing Primary or Secondary VSM in an Dual VSM system.

Before you begin

- Log in to the CLI in EXEC mode.
- Power off the primary VSM.
- Configure the port groups so that the new primary VSM cannot communicate with the secondary VSM or any of the VEMs during the setup. VSMs with a primary or secondary redundancy role have built-in mechanisms for detecting and resolving the conflict between two VSMs in the active state. To avoid these mechanisms during the configuration of the new primary VSM, you must isolate the new primary VSM from the secondary VSM.

Procedure

-
- Step 1** Ensure that the secondary VSM is active.
- Note** Do a system switchover if necessary.
- Step 2** Log into CLI in EXEC mode on the active secondary VSM. Enter **peer-sup mac-addresses-clear** command.
- Step 3** Verify that the "Peer Sup Mac Addresses Learnt" section in the **show system redundancy status** command displays "Not Learnt."
- Step 4** On the vSphere Client, change the port group configuration for the new primary VSM VM to prevent communication with the secondary VSM and the VEMs during the setup.
- For an example on how to change the port groups and port profiles assigned to the VSM interfaces in the vSphere Client, see the *Cisco Nexus 1000V Installation and Upgrade Guide*.
- Step 5** Install the new VSM, change the HA role from standalone to primary and set the same domain ID as the existing VSM, using the **Installing the Cisco Nexus 1000V Software Using ISO or OVA Files** section in the *Cisco Nexus 1000V Installation and Upgrade Guide*.
- Step 6** Save the configuration.
- Step 7** Power off the VM.
- Step 8** On the vSphere Client, change the port group configuration for the new primary VSM to permit communication with the secondary VSM and the VEMs.
- Step 9** Power up the new primary VSM. Both VSMs will learn the peer VSM's MAC addresses.

The new primary VSM starts and automatically synchronizes all configuration data with the secondary VSM, which is currently the active VSM. Because the existing VSM is active, the new primary VSM becomes the standby VSM and receives all configuration data from the existing active VSM.

Changing the Domain ID in a Dual VSM System

Before you begin

- Have access to the console of both the active and standby VSM.
- Isolate the standby VSM from the active VSM to avoid the built-in mechanisms that detect and resolve conflict between two VSMs with a primary or secondary redundancy role. This procedure has a step for isolating the VSMs.



Note Equipment Outage—This procedure requires that you power down a VSM. During this time, your system will operate with a single VSM.

Procedure

Step 1 On the vSphere Client for the standby VSM, do one of the following to isolate the VSMs and prevent their communication while completing this procedure:

- Change the port group configuration for the interfaces using port groups that prevent the VSMs from communicating with each other.
- Unmark the “Connected” option for the interfaces.

The standby VSM becomes active but cannot communicate with the other active VSM or the VEM

Step 2 At the console of the standby VSM, change the domain ID and save the configuration.

Example:

This example shows how to change the domain ID and save the configuration:

```
switch# configure terminal
switch(config)# svs-domain
switch(config-svs-domain)# domain id 100
Successfully cleared old Peer VSM's MAC Addresses
=====
IMPORTANT NOTE: If this VSM is replacing a Standby VSM which was in HA
pair then, please execute "peer mac-addresses clear" CLI on Active VSM
to clear old Peer VSM's MAC Addresses on Active VSM as well.
=====
switch(config-svs-domain)# copy running-config startup-config
```

The domain ID is changed on the standby VSM and the VEM connected to it

Step 3 Power down the standby VSM.

Step 4 At the console of the active VSM, change the domain ID and save the configuration.

Example:

```
switch# configure terminal
switch(config)# svcs-domain
switch(config-svs-domain)# domain id 100
Successfully cleared old Peer VSM's MAC Addresses
=====
IMPORTANT NOTE: If this VSM is replacing a Standby VSM which was in HA
pair then, please execute "peer mac-addresses clear" CLI on Active VSM
to clear old Peer VSM's MAC Addresses on Active VSM as well.
=====

switch(config-svs-domain)# copy running-config startup-config
```

The domain ID is changed on the active VSM and the VEM that is connected to it.

Step 5 On the vSphere Client for the standby VSM, do one of the following to permit communication with the active VSM:

- Change the port group configuration for the interfaces.
- Make sure that the "Connect at power on" option is marked for the interfaces.

When the standby VSM is powered up, it will be able to communicate with the active VSM.

Step 6 Power up the standby VSM.

Both VSMs are now using the new domain ID and will synchronize.

Changing the Domain ID in a Dual VSM System for VSMs Hosted on Cisco Nexus 1010

Before you begin

- Have access to the CLI of the active VSM (primary VSB) and standby VSM (secondary VSB).
- Have access to the Cisco Nexus 1010 CLI.

Procedure

Step 1 On the Cisco Nexus 1010 CLI, complete the following tasks:

a) Power down the secondary VSB.

Example:

```
switch# configure terminal
switch(config)# virtual-service-blade VSM1
switch(config-vsbs-config)# shutdown secondary
```

b) Check the status of the secondary VSB and verify that the status is **VSB POWERED OFF**.

Example:

```
switch(config-vsbs-config)# show virtual-service-blade summary
```

```
-----
Name Role State Nexus1010-Module
-----
VSM1 PRIMARY VSB POWERED ON Nexus1010-PRIMARY
VSM1 SECONDARY VSB POWERED OFF Nexus1010-SECONDARY
```

- c) Delete the secondary VSB and verify that the status is **VSB NOT PRESENT**.

Example:

```
switch(config-vsbs-config)# no enable secondary
switch(config-vsbs-config)# show virtual-service-blade summary
```

```
-----
Name Role State Nexus1010-Module
-----
VSM1 PRIMARY VSB POWERED ON Nexus1010-PRIMARY
VSM1 SECONDARY VSB NOT PRESENT Nexus1010-SECONDARY
```

Step 2 On the primary VSB, complete the following tasks:

- a) Verify that the secondary VSB is not connected.
- b) Verify that the operational status of primary VSB is **Connected**.

Example:

```
switch# show svcs connections
connection VC:
ip address: 192.168.0.1
protocol: vmware-vim https
certificate: default
datacenter name: Hamilton-DC
DVS uuid: ac 36 07 50 42 88 e9 ab-03 fe 4f dd d1 30 cc 5c
config status: Enabled
operational status: Connected
```

- c) Change the domain ID and save the configuration.

Example:

```
switch# configure terminal
switch(config)# svcs-domain
switch(config-svs-domain)# domain id 100
Successfully cleared old Peer VSM's MAC Addresses
=====
IMPORTANT NOTE: If this VSM is replacing a Standby VSM which was in HA
pair then, please execute "peer mac-addresses clear" CLI on Active VSM
to clear old Peer VSM's MAC Addresses on Active VSM as well.
=====
switch(config-svs-domain)# copy running-config startup-config
```

- d) Verify the new domain ID.

Example:

```
switch(config)# show svcs domain
SVS domain config:
Domain id: 100
L2/L3 Control mode: L2
```

```
L3 control interface: NA
Status: Config push to VC successful.
```

- e) Verify that the domain ID is updated on VEMs by running the following command on the VEM modules:

Example:

```
switch# vemcmd show card
ard UUID type 2: 58f8afd7-e1e3-3c51-85e2-6e6f2819a7b8
Card name: sfish-srvr-1
Switch name: n1000v
Switch alias: DvsPortset-0
Switch uuid: 56 e0 36 50 91 1c 32 7a-e9 9f 31 59 88 0c 7f 76
Card domain: 100
Card slot: 4
VEM Control (Control VLAN) MAC: 00:02:3d:14:00:03
VEM Packet (Inband) MAC: 00:02:3d:24:00:03
VEM Control Agent (DPA) MAC: 00:02:3d:44:00:03
VEM SPAN MAC: 00:02:3d:34:00:03
Management IP address: 172.23.232.102
Max physical ports: 32
Max virtual ports: 216
Card control VLAN: 3002
Card packet VLAN: 3003
    Processors: 4
    Processor Cores: 4
    Processor Sockets: 2
    Physical Memory: 4290351104
```

Step 3 On the Cisco Nexus 1010 CLI, complete the following tasks:

- a) Deploy a secondary VSB.

Example:

```
switch# configure terminal
switch(config)# virtual-service-blade VSM1
switch(config-vsb-config)# enable secondary
Enter vsb image: [dcos_vsm.iso]
Enter domain id[1-4095]: 100
Management IP version [V4/V6]: [V4]
Enter Management IP address: 10.78.109.67
Enter Management subnet mask length: 27
IPv4 address of the default gateway: 10.78.109.65
Enter HostName: switch
Enter the password for 'admin': xz35vb1zx
```

- b) Check the status of the secondary VSB and verify that the status is **VSB POWERED ON**.

Example:

```
switch(config-vsb-config)# show virtual-service-blade summary

-----
Name Role State Nexus1010-Module
-----
VSM1 PRIMARY VSB POWERED ON Nexus1010-PRIMARY
VSM1 SECONDARY VSB POWERED ON Nexus1010-SECONDARY
```

Step 4 On the primary VSB, verify that the HA pair is formed.

Example:

```
switch# show system redundancy status
Redundancy role
```



```

-----
administrative: primary
operational: primary
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

```

Disabling Domain ID Collision

Procedure

- Step 1** Log in to the primary VSM console in EXEC mode.
- Step 2** Disable the domain ID collision detection.
- Example:**
- ```
switch# peer-sup mac-addresses check disable
```
- Step 3** Verify that the check is disabled.
- Example:**
- ```
switch# show peer-sup mac-addresses details
Peer MAC Address Check = Disabled
Peer HA0 MAC Address = 00:50:56:b5:3a:99
Peer HA1 MAC Address = 00:50:56:b5:5e:05
```
- Step 4** Log in to the secondary VSM console in EXEC mode
- Step 5** Disable the domain id collision detection.
- Example:**
- ```
switch# peer-sup mac-addresses check disable
```
- Step 6** Verify that the check is disabled.
- Example:**
- ```
switch# show peer-sup mac-addresses details
Peer MAC Address Check = Disabled
Peer HA0 MAC Address = 00:50:56:b5:3a:99
Peer HA1 MAC Address = 00:50:56:b5:5e:05
```

Verifying the HA Status

Use one of the following commands to verify the configuration:

Command	Purpose
<code>show system redundancy status</code>	Displays the HA status of the system.
<code>show module</code>	Displays information about all available VSMS and VEMs in the system.
<code>show processes</code>	<p>Displays the state of all processes and the start count of the process.</p> <p>The states and types are described as follows:</p> <ul style="list-style-type: none"> • State: R (runnable), S (sleeping), Z (defunct) • Type: U (unknown), O (non sysmgr), VL (vdc-local), VG (vdc-global), VU (vdc-unaware), NR (not running), ER (terminated)

Starting with Release 5.2(1)SV3(1.1), the VSM drops the HA packets when the source MAC address is not known. If the peer VSM's MAC addresses are not learned correctly (such as when a standby VSM is replaced without following a correct procedure), the VSM is not formed.

The `show system redundancy status` command displays a note in this case output. After clearing the old MAC addresses, HA should be formed if the problem was due to incorrect MAC addresses.

```
switch# show system redundancy status

Redundancy role
-----
      administrative:  primary
      operational:    primary

Redundancy mode
-----
      administrative:  HA
      operational:    None

This supervisor (sup-1)
-----
      Redundancy state:  Active
      Supervisor state:  Active
      Internal state:   Active with no standby

Other supervisor (sup-2)
-----
      Redundancy state:  Not present

Peer Sup Mac Addresses Learnt
-----
      Control Interface: 00:50:56:91:44:c8
      Mgmt Interface:   00:50:56:91:1f:6f
```

```
HA Packet Drops Due to Domain id Collision
```

```
-----
Control Interface: 36
Mgmt Interface: 51
-----
```

```
IMPORTANT NOTE: Please compare Peer Sup MAC addresses learnt above
with the actual Peer Sup's MAC addresses. If they are not same, execute
"peer-sup mac-addresses clear" on this VSM to form HA again
-----
```

```
switch#
```

Related Documents

Related Topic	Document Title
Software upgrades	<i>Cisco Nexus 1000V Installation and Upgrade Guide</i>
Cisco Nexus 1000V commands	<i>Cisco Nexus 1000V Command Reference</i>

Standards

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-PROCESS-MIB	To locate and download MIBs, go to the following URL: http://www.cisco.com/public/sw-center/netmgmt/cmtk/mibs.shtml

RFCs

No RFCs are supported by this feature.

Technical Assistance

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.

Go to the following URL: <http://www.cisco.com/cisco/web/support/index.html>

Feature History for System-Level High Availability

This table includes only the updates for those releases that have resulted in additions or changes to the feature.

Feature Name	Releases	Feature Information
Enhancements for Domain ID Collision	5.2(1)SV3(1.1)	This feature was introduced.
System - Level High Availability	4.0(4)SV1(1)	This feature was introduced.