High Availability

This chapter describes how to identify and resolve problems related to high availability, and includes the following sections:

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 if a failure occurs:

- 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 statefu1 switchover if a VSM failure occurs.

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

- Virtual Ethernet Modules (VEMs) running within virtualization servers. These VEMs are represented as modules within the VSM.
- A remote management component, such as VMware vCenter Server.
- One or two VSMs running within virtual machines (VMs).

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 state and configuration remain constantly synchronized between the two VSMs to provide a statefu1 switchover if the active VSM fails.

Network-Level High Availability

The Cisco Nexus 1000V HA at the network level includes port channels and 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.

For additional information about port channels and LACP, see the *Cisco Nexus 1000V Layer 2 Switching Configuration Guide*.

Problems with High Availability

Symptom	Possible Causes	Solution
The active VSM does not see the standby VSM.	MAC addresses mismatch. • Check that the peer VSM MAC addresses that are learned by the active VSM by using the **show system redundancy status** command.	Confirm that the standby VSM MAC addresses are correctly learned by the active VSM. 1. Compare the standby VSM MAC addresses with the output MAC addresses by using the **show system redundancy status** command on the active VSM. 2. If the compared MAC addresses are different, use the **peer-sup mac-addresses clear** command to clear the stale MAC addresses that are learned by the active VSM.

Symptom	Possible Causes	Solution
The active VSM does not see the standby VSM.	Roles are not configured properly. • Check the role of the two VSMs by using the **show system redundancy status** command.	1. Confirm that the roles are the primary and secondary role, respectively. 2. If needed, use the **system redundancy role** command to correct the situation. 3. Save the configuration if roles are changed.
	Network connectivity problems. • Check that the control and management VLAN connectivity between the VSM at the upstream and virtual switches.	If network problems exist, do the following: 1. From vSphere Client, shut down the VSM, which should be in standby mode. 2. From vSphere Client, bring up the standby VSM after network connectivity is restored.
The active VSM does not complete synchronization with the standby VSM.	Version mismatch between VSMs. • Check that the primary and secondary VSMs are using the same image version by using the **show version** command.	If the active and the standby VSM software versions differ, reinstall the secondary VSM with the same version used in the primary.
	Fatal errors during gsync process. • Check the gsyncctrl log using the **show system internal log sysmgr gsyncctrl** command and look for fatal errors.	Reload the standby VSM using the **reload module** *module-number* command, where *module-number* is the module number for the standby VSM.
	• The VSM has connectivity only through the management interface. • Check the output of the **show system internal redundancy info** command and verify if the *degraded_mode* flag is set to *true*.	Check control VLAN connectivity between the primary and the secondary VSMs.

Symptom	Possible Causes	Solution
The standby VSM reboots periodically.	The VSM has connectivity only through the management interface. • Check the output of the **show system internal redundancy info** command and verify that the *degraded_mode* flag is set to true.	Check the control VLAN connectivity between the primary and the secondary VSMs.
	The VSMs have different versions. Enter the **debug system internal sysmgr all** command and look for the active_verctrl entry that indicates a version mismatch, as the following output shows: `2009 May 5 08:34:15.721920 sysmgr: active_verctrl: Stdby running diff version- force download the standby sup.`	Isolate the standby VSM and boot it. Use the **show version** command to check the software version in both VSMs. Install the image matching the active VSM on the standby.
Active-Active detected and resolved	When control and management connectivity between the active and the standby goes down for 6 seconds, the standby VSM transitions to the active state. Upon restoration of control and management connectivity, both VSMs detect an active-active condition.	1. Once the system detects active-active VSMs, one VSM is automatically reloaded based on various parameters such as VEMs attached, vCenter connectivity, last configuration time, and last active time. 2. To see any configuration changes that are performed on the rebooted VSM during the active-active condition, enter the **show system internal active-active remote accounting logs** CLI command on the active VSM.
VSM Role Collision	If another VSM is configured/provisioned with the same role (primary or secondary) in the system, the new VSM collides with the existing VSM. The **show system redundancy info** command displays the MAC addresses of the VSM(s) that collide with the working VSM.	If the problems exist, do the following: 1. Enter the **show system redundancy status** command on the VSM console. 2. Identify the VSM(s) that owns the MAC addresses that are displayed in the output of the **show system redundancy status** command. 3. Move the identified VSM(s) out of the system to stop role collision.

Symptom	Possible Causes	Solution
Both VSMs are in active mode.	Network connectivity problems. • Check for control and management VLAN connectivity between the VSM at the upstream and virtual switches. • When the VSM cannot communicate through any of these two interfaces, they will both try to become active.	If network problems exist, do the following: 1. From vSphere Client, shut down the VSM, which should be in standby mode. 2. From vSphere Client, bring up the standby VSM after network connectivity is restored.
	Different domain IDs in the two VSMs Check the *domain* value by using **show system internal redundancy info** command.	If needed, update the domain ID and save it to the startup configuration. • Upgrading the domain ID in a dual VSM system must be done as follows: – Isolate the VSM with the incorrect domain ID so that it cannot communicate with the other VSM. – Change the domain ID in the isolated VSM, save the configuration, and power off the VSM. – Reconnect the isolated VSM and power it on.

High Availability Troubleshooting Commands

This section lists commands that can be used troubleshoot problems related to high availability.

Command	Description
attach module	See Example 6-9attach module Command, page 6-10
reload module	See Example 6-8reload module Command, page 6-10
show cores	Use to list process logs and cores. See Example 6-1show cores Command, page 6-6
show processes [pid pid]	See Example 6-2show processes log [pid pid] Command, page 6-6
show system internal active-active	See Example 6-7show system internal active-active remote accounting logs Command, page 6-10

Command	Description
show system internal redundancy info	See Example 6-4show system internal redundancy info Command, page 6-7
show system internal sysmgr state	See Example 6-5show system internal sysmgr state Command, page 6-8
show system redundancy status	See Example 6-3show system redundancy status Command, page 6-6
show system redundancy status	See Example 6-6show system redundancy status Command, page 6-9

To list process logs and cores, use the following commands:

Example 6-1 show cores Command

```
switch# show cores
VDC No Module-num       Process-name      PID     Core-create-time
------ ----------       ------------      ---     ----------------
1      1                private-vlan      3207    Apr 28 13:29
```

Example 6-2 show processes log [pid pid] Command

```
switch# show processes log
VDC Process          PID     Normal-exit  Stack  Core   Log-create-time
--- ---------------  ------  -----------  -----  -----  ---------------
  1 private-vlan     3207             N      Y      N  Tue Apr 28 13:29:48 2009

switch# show processes log pid 3207
====================================================
Service: private-vlan
Description: Private VLAN

Started at Wed Apr 22 18:41:25 2009 (235489 us)
Stopped at Tue Apr 28 13:29:48 2009 (309243 us)
Uptime: 5 days 18 hours 48 minutes 23 seconds

Start type: SRV_OPTION_RESTART_STATELESS (23)
Death reason: SYSMGR_DEATH_REASON_FAILURE_SIGNAL (2) <-- Reason for the process abort
Last heartbeat 46.88 secs ago
System image name: nexus-1000v-mzg.4.0.4.SV1.1.bin
System image version: 4.0(4)SV1(1) S25

PID: 3207
Exit code: signal 6 (core dumped) <-- Indicates that a cores for the process was
generated.

CWD: /var/sysmgr/work
...
```

To check redundancy status, use the following commands:

Example 6-3 show system redundancy status Command

```
switch# show system redundancy status
Redundancy role
---------------
      administrative:   primary <-- Configured redundancy role
```

```
        operational:   primary <-- Current operational redundancy role

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

This supervisor (sup-1)
-----------------------
    Redundancy state:   Active <-- Redundancy state of this VSM
    Supervisor state:   Active
      Internal state:   Active with HA standby

Other supervisor (sup-2)
------------------------
    Redundancy state:   Standby <-- Redundancy state of the other VSM
    Supervisor state:   HA standby
      Internal state:   HA standby <-- The standby VSM is in HA mode and in sync
```

To check the system internal redundancy status, use the following command:

Example 6-4 show system internal redundancy info Command

```
switch# show system internal redundancy info
My CP:
  slot: 0
  domain: 184 <-- Domain id used by this VSM
  role:   primary <-- Redundancy role of this VSM
  status: RDN_ST_AC <-- Indicates redundancy state (RDN_ST) of the this VSM is Active (AC)
  state:  RDN_DRV_ST_AC_SB
  intr:   enabled
  power_off_reqs: 0
  reset_reqs:     0
Other CP:
  slot: 1
  status: RDN_ST_SB <-- Indicates redundancy state (RDN_ST) of the other VSM is Standby
(SB)
  active: true
  ver_rcvd: true
  degraded_mode: false <-- When true, it indicates that communication through the control
interface is faulty
Redun Device 0: <-- This device maps to the control interface
  name: ha0
  pdev: ad7b6c60
  alarm: false
  mac: 00:50:56:b7:4b:59
  tx_set_ver_req_pkts:   11590
  tx_set_ver_rsp_pkts:   4
  tx_heartbeat_req_pkts: 442571
  tx_heartbeat_rsp_pkts: 6
  rx_set_ver_req_pkts:   4
  rx_set_ver_rsp_pkts:   1
  rx_heartbeat_req_pkts: 6
  rx_heartbeat_rsp_pkts: 442546 <-- Counter should be increasing, as this indicates that
communication between VSM is working properly.
  rx_drops_wrong_domain: 0
  rx_drops_wrong_slot:   0
  rx_drops_short_pkt:    0
  rx_drops_queue_full:   0
  rx_drops_inactive_cp:  0
  rx_drops_bad_src:      0
  rx_drops_not_ready:    0
  rx_unknown_pkts:       0
```

```
Redun Device 1: <-- This device maps to the mgmt interface
  name: ha1
  pdev: ad7b6860
  alarm: true
  mac: ff:ff:ff:ff:ff:ff
  tx_set_ver_req_pkts:   11589
  tx_set_ver_rsp_pkts:   0
  tx_heartbeat_req_pkts: 12
  tx_heartbeat_rsp_pkts: 0
  rx_set_ver_req_pkts:   0
  rx_set_ver_rsp_pkts:   0
  rx_heartbeat_req_pkts: 0
  rx_heartbeat_rsp_pkts: 0 <-- When communication between VSM through the control
interface is interrupted but continues through the mgmt interface, the
rx_heartbeat_rsp_pkts will increase.
  rx_drops_wrong_domain: 0
  rx_drops_wrong_slot:   0
  rx_drops_short_pkt:    0
  rx_drops_queue_full:   0
  rx_drops_inactive_cp:  0
  rx_drops_bad_src:      0
  rx_drops_not_ready:    0
  rx_unknown_pkts:       0
```

To check the system internal sysmgr state, use the following command:

Example 6-5 show system internal sysmgr state Command

```
switch# show system internal sysmgr state

The master System Manager has PID 1988 and UUID 0x1.
Last time System Manager was gracefully shutdown.
The state is SRV_STATE_MASTER_ACTIVE_HOTSTDBY entered at time Tue Apr 28 13:09:13 2009.

The '-b' option (disable heartbeat) is currently disabled.

The '-n' (don't use rlimit) option is currently disabled.

Hap-reset is currently enabled.

Watchdog checking is currently disabled.

Watchdog kgdb setting is currently enabled.


        Debugging info:

The trace mask is 0x00000000, the syslog priority enabled is 3.
The '-d' option is currently disabled.
The statistics generation is currently enabled.


        HA info:


slotid = 1    supid = 0
cardstate = SYSMGR_CARDSTATE_ACTIVE .
cardstate = SYSMGR_CARDSTATE_ACTIVE (hot switchover is configured enabled).
Configured to use the real platform manager.
Configured to use the real redundancy driver.
Redundancy register: this_sup = RDN_ST_AC, other_sup = RDN_ST_SB.
EOBC device name: eth0.
Remote addresses:  MTS - 0x00000201/3      IP - 127.1.1.2
```

```
MSYNC done.
Remote MSYNC not done.
Module online notification received.
Local super-state is: SYSMGR_SUPERSTATE_STABLE
Standby super-state is: SYSMGR_SUPERSTATE_STABLE
Swover Reason : SYSMGR_SUP_REMOVED_SWOVER <-- Reason for the last switchover
Total number of Switchovers: 0 <-- Number of switchovers
                              >> Duration of the switchover would be listed, if any.

        Statistics:

Message count:          0
Total latency:          0               Max latency:            0
Total exec:             0               Max exec:               0
```

When a role collision is detected, a warning is highlighted in the CLI output. Use the following command to display the CLI output:

Example 6-6 show system redundancy status Command

```
switch# 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.
```

If no collisions are detected, the highlighted output is not displayed.

Use the following command to display the accounting logs that are stored on a remote VSM.

Example 6-7 show system internal active-active remote accounting logs Command

```
switch# show system internal active-active remote accounting logs
```

To reload a module, use the following command:

Example 6-8 reload module Command

```
switch# reload module 2
```

This command reloads the secondary VSM.

Note Entering the **reload** command without specifying a module will reload the whole system.

To attach to the standby VSM console, use the following command.

Example 6-9 attach module Command

The standby VSM console is not accessible externally, but can be accessed from the active VSM through the **attach module** *module-number* command.

```
switch# attach module 2
```

This command attaches to the console of the secondary VSM.