CHAPTER 6

High Availability

This chapter describes how to identify and resolve problems related to High Availability.

This chapter 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 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. State and configuration remain constantly synchronized between two Virtual Supervisor Modules (VSMs) to provide seamless and statefu1 switchover in the event of a VSM failure.

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

- Virtual Ethernet Modules (VEMs) running within virtualization servers. These are represented as modules within the VSM.
- A remote management component, for example. 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

Single or Dual Supervisors

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

- Virtual Ethernet Modules (VEMs) running within virtualization servers (these are represented as modules within the VSM)
- A remote management component, for example. VMware vCenter Server.
- One or two Virtual Supervisor Modules (VSMs) running within Virtual Machines (VMs)

Single VSM Operation	Dual VSM Operation
• Stateless—Service restarts from the startup configuration • Stateful—Service resumes from previous state.	• 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. They are also synchronized and kept up to date with the active VSM in order to maintain the runtime context of “ready to run.” • On a switchover, the standby VSM takes over for the active VSM.

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 lets you 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, Release 4.0.*

Problems with High Availability

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 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 the control and management VLAN connectivity between VSM at the upstream and virtual switches.	If network problems exist: 1. From the vSphere client, shut down the VSM, which should be in standby mode. 2. From the 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 primary and secondary VSM are using the same image version using **show version** command.	If the active and 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.

Symptom	Possible Causes	Solution
The standby VSM reboots periodically.	The VSM has connectivity only through the management interface. • When a VSM is able to communicate through the management interface, but not through the control interface, the active VSM detects the situation and resets the standby VSM to prevent the two VSMs from being in HA mode and out of sync. • 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 secondary VSMs.
	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.

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: 1. From the vSphere client, shut down the VSM, which should be in standby mode. 2. From the vSphere client, bring up the standby VSM after network connectivity is restored.
	Different domain IDs in the two VSMs Check *domain* value 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 following a certain procedure. – 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 configuration, and power off the VSM. – Reconnect the isolated VSM and power it on.

Recovering VSMs in an HA Setup after Executing Write Erase

After entering the **write erase** command on a secondary VSM and bringing up this VSM to rejoin the primary VSM, the primary VSM resets and a cluster outage occurs. The **write erase** command clears the entire configuration except domain-id and system role.

This section contains the following topics to recover VSMs in an HA setup after executing the **write erase** command:

Recovering a Standalone VSM after Executing a Write Erase

You can recover a standalone VSM after executing the **write erase** command.

PROCEDURE

Step 1 Copy the running configuration to the TFTP server. Enter the following command:

```
copy running-configuration
```

Step 2 Erase all the configurations in the startup configuration. Enter the following command:

```
write erase
```

Step 3 Reload the VSM. Enter the following command:

```
reload
```

After the VSM is reloaded, it comes up as a fresh VSM.

Step 4 Configure the domain-id. Enter the following command:

```
domain id number
```

Step 5 Configure the role. Enter the following command:

```
role name role-name
```

Step 6 Set up the initial configuration.

Step 7 Copy the running configuration to the startup configuration from the TFTP server. Enter:

```
copy running-configuration
```

Note Ignore all the warnings and error messages.

All the modules are attached.

Step 8 Copy the running configuration to the startup configuration. Enter:

```
copy running-configuration start-configuration
```

Recovering an HA Setup VSMs after Executing a Write Erase

You can recover HA setup VSMs after executing the **write erase** command.

PROCEDURE

Step 1 Be sure HA is configured correctly between the primary and secondary VSMs.

Step 2 Copy the running configuration to the TFTP server. Enter the following command:

```
copy running-configuration
```

Step 3 Erase all the configurations in the startup configuration. Enter the following command:

```
write erase
```

Step 4 Reload the VSM. Enter the following command:

```
reload
```

After the VSM is reloaded, it comes up as a fresh VSM.

Step 5 Configure the domain-id on the primary VSM. Enter the following command:

domain id *number*

Step 6 Configure the role on the primary VSM. Enter the following command:

role name *role-name*

Step 7 Set up the initial configuration.

Step 8 Copy the running configuration from the TFTP server. Enter the following command:

copy running-configuration

Note Ignore all the warnings and error messages.

All the modules are attached.

Step 9 Copy the running configuration to the startup configuration. Enter the following command:

copy running-configuration

Step 10 Configure the domain ID on the secondary VSM. Enter the following command:

domain id *number*

Step 11 Configure the role on the secondary VSM. Enter the following command:

role name *role-name*

Step 12 When you are prompted to reload, enter **y**.

After the VSM is reloaded, it will start to synchronize as the standby in the HA setup.

Recovering an Individual Secondary VSM in an HA Setup after Executing a Write Erase

You can recover an individual secondary VSM in an HA setup after executing the **write erase** command.

PROCEDURE

Step 1 Be sure that HA is correctly configured between the primary and secondary VSMs.

Step 2 Perform a system switchover, if the secondary VSM is active.

Step 3 After the HA is configured, enter

write erase

All the configurations are erased in the startup-configuration.

Step 4 Reload only the secondary VSM. Enter:

reload module

Step 5 Copy the running configuration to the startup configuration in the primary active VSM. Enter:

copy running-configuration start-configuration

After the secondary VSM is reloaded, it comes up as a fresh VSM.

Step 6 On the secondary VSM, configure the domain-id. Enter the following command:

```
domain id number
```

Step 7 On the secondary VSM, configure the role. Enter the following command:

```
role name role-name
```

Step 8 When you are prompted to reload, enter **y**.

After the VSM is reloaded, it will start to synchronize as the standby in the HA setup.

Recovering an Individual Primary VSM in an HA Setup after Executing Write Erase

You can recover an individual primary VSM in an HA setup after executing the **write erase** command.

PROCEDURE

Step 1 Make sure that HA is configured successfully between primary and secondary VSMs.

Step 2 Perform a system switchover, if the primary VSM is active.

Step 3 After HA is configured, enter the following command:

```
write erase
```

All the configurations in the startup-configuration are erased.

Step 4 Reload only the primary VSM. Enter the following command:

```
reload module 1
```

Step 5 Copy the running configuration to the startup configuration in the secondary active VSM.

Step 6 Once Primary VSM reloaded, it will come up as fresh VSM.

Step 7 Disable the **Connect** option in the VC to disconnect the control and management communication from the primary VSM

Step 8 On the primary VSM, configure the domain-id. Enter the following command:

```
domain id number
```

Step 9 On the primary VSM, configure the role. Enter the following command:

```
role name role-name
```

Step 10 Skip the initial configuration.

Step 11 Enter the following command:

```
copy running-configuration start-configuration
```

Step 12 Power on the primary VSM in the VC.

After the primary VSM is powered on, it will start to synchronize as the standby in the HA setup.

High Availability Troubleshooting Commands

This section lists commands that can be used troubleshoot problems related to High Availability.

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

- **show cores**

```
n1000V# show cores
VDC No Module-num       Process-name      PID     Core-create-time
------ ----------       ------------      ---     ----------------
1      1                private-vlan      3207    Apr 28 13:29
```

- **show processes log [pid** *pid*]

```
n1000V# 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

n1000V# 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:

- **show system redundancy status**

```
N1000V# 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:

- **show system internal redundancy info**

```
n1000V# 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:

- **show system internal sysmgr state**

```
N1000V# 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
```

To reload a module, use the following command:

- **reload module**

  ```
  n1000V# reload module 2
  ```

  This command reloads the secondary VSM.

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

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

- **attach module**

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

  ```
  n1000V# attach module 2
  ```

  This command attaches to the console of the secondary VSM.