



Understanding Service-Level High Availability

This chapter contains the following sections:

- [Information About Cisco NX-OS Service Restarts, page 1](#)
- [Restartability Infrastructure, page 1](#)
- [Process Restartability, page 3](#)
- [Restarts on Standby Supervisor Services , page 5](#)
- [Restarts on Switching Module Services, page 5](#)
- [Troubleshooting Restarts, page 5](#)
- [MIBs, page 5](#)
- [RFCs, page 6](#)
- [Technical Assistance, page 6](#)

Information About Cisco NX-OS Service Restarts

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

Restartability Infrastructure

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

System Manager

The System Manager directs overall system function, service management, and system health monitoring, and enforces high-availability policies. The System Manager is responsible for launching, stopping, monitoring, and restarting services 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 or on modules.

Message and Transaction Service

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

HA Policies

Cisco NX-OS allows each service to have an associated set of internal HA policies that define how a failed service 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—Specifies the number of restart attempts to be performed by the System Manager. If the service has not restarted successfully after this number of attempts, the HA policy is considered to have failed, and the next HA policy is used. If no other HA policy exists, the default policy is applied, resulting in a supervisor switchover or restart.
- Minimum lifetime—Specifies the time that a service must run after a restart attempt in order to consider the restart attempt as successful. The minimum lifetime will be 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, which allows 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 sent by other services to the failed service will be available from the MTS when the service resumes.

Whether or not the new instance survives the stateful initialization depends on the cause of failure of the previous instance. If the service is unable to survive a few subsequent restart attempts, the restart is considered as failed. In this case, the System Manager 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 VSMS, 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 non-supervisor module services do not require a supervisor switchover.

On the VEMs, the DPA is restarted if it crashes. This situation causes the module to be removed and re-added 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 Trivial File Transfer Protocol (TFTP).

For information on collecting and using the generated information that relates to service failures, see the *Cisco Nexus 1000V Troubleshooting Guide*.

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>