Managing Redundancy for Units and Processes

The unit server high availability and AVM protections architecture ensures continuous availability of Prime Network functionality by detecting and recovering from a wide range of hardware and software failures. The distributed design of the system enables the *impact radius* caused by a single fault to be confined. This prevents all types of faults from setting into motion the "domino" effect, which can lead to a crash of all the management services.

These topics describes how you can use Prime Network for unit redundancy and process protection:

- Overview of Unit and Process Protection, page 5-1
- What is the Impact of Unit or AVM Failures?, page 5-3
- Creating a New Unit Protection Group, page 5-9
- Switching to a Standby Unit (Disable Active Unit), page 5-10
- Changing Timeouts and Restarts for Unit and Process Protection, page 5-10
- Tracking Unit and Process Protection Events, page 5-11

For information on high availability for gateway servers, refer to the *Cisco Prime Network 4.3 Gateway High Availability Guide*.

Overview of Unit and Process Protection

The following topics explain the process (AVM) protection and unit high availability features provided by Prime Network. Most of these settings are enabled by default. When you create an AVM, AVM process protection is automatically enabled. When units are created during installation, you specify whether they will be standby or active units. Units are automatically added to the protection group default-pg.

Process (AVM) Protection

The *AVM protection* mechanism monitors AVM processes to make sure any failed AVMs are restarted. Protection is normally enabled by default and is controlled by way of the AVM Protection check box in the AVM properties dialog box.

All AVM processes within a unit are completely independent so that a failure of a single process does not affect other processes on the same machine. The exact number of processes on each unit depends on the capacity and computational power of the unit.

On the unit, a control process starts the watchdog protocol which continuously monitors all other processes on the unit. The watchdog protocol requires each AVM process to continuously handshake with the control process. A process that fails to handshake with the control process after a number of times is automatically canceled and reloaded.

The unit control process monitors AVM restarts and will escalate the issue, according to the system settings. For example, if a process has crashed more than *n* times within a given period, Prime Network will no longer attempt to restart it because it is suspected of having a recurring software problem.

The reload process is local to the unit, and thus very rapid, with a minimal amount of downtime. In many cases the process can use its previous cache information (temporary persistency used to improve performance), once the stuck process id detected, reloading the process takes only a few seconds with no data loss. This is the case for user-created AVMs that are hosting VNEs. However, for reserved AVMs that perform special function in Prime Network, some data loss will occur. All watchdog activity is logged and an alarm is generated and sent when the watchdog reloads a process.

Note An alarm persistency mechanism enables the system to clear alarms that relate to events that occurred while a VNE, an AVM, a unit, or the whole system was down, thus preserving system integrity. For more information about alarm persistency, see Changing Settings That Control VNE Data Saved After Restarts, page 12-37

Watchdog protocol parameters are configurable in the registry. See Changing Timeouts and Restarts for Unit and Process Protection, page 5-10.

Unit N+m High Availability

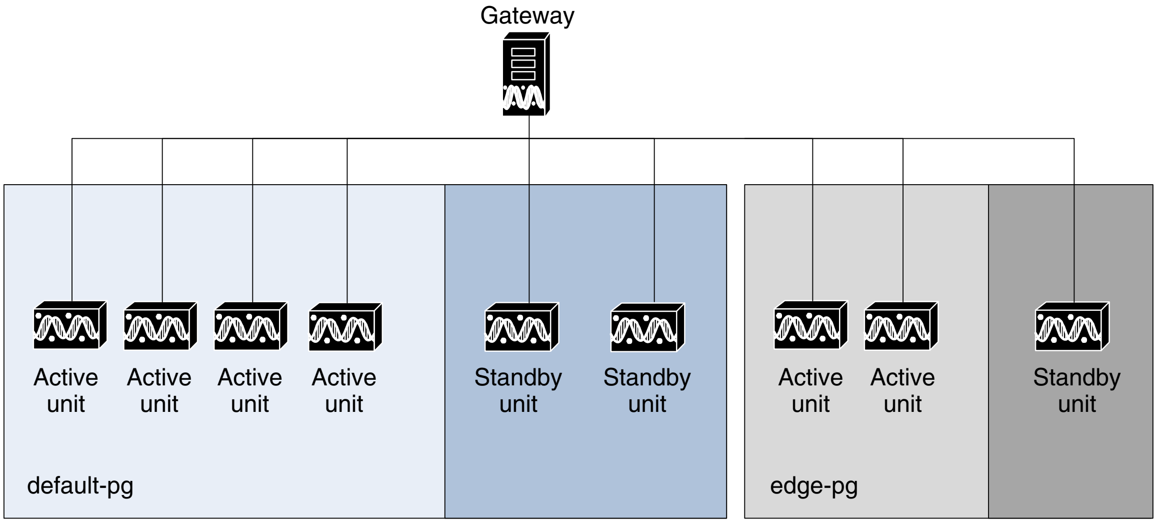
Unit availability is established in the gateway, running a *protection manager* process, which continuously monitors all the units in the network. Once the protection manager detects a unit that is malfunctioning, it automatically signals one of the standby servers in its cluster to load the configuration of the faulty unit (from the system registry), taking over all of its managed network elements. This design provides many possibilities for trading off protection and resources. These possibilities range from segmenting the network into clusters without any extra machines, to having a warm-swappable empty unit for each unit in the setup.

When the Prime Network software is installed, you can specify whether a unit is an *active* or *standby* unit. Using the GUI, you can designate a group of active units and a standby unit to be the members of a *protection group*, giving the group the name of your choice. A protection group can have multiple standby units, and you can define more than a single protection group. By default, all units are added to the default-pg protection group.

Even with unit redundancy, a unit switchover can result in the unavoidable loss of information. The impact depends on how long the unit is down and the functions the unit performed. See Impact of Unit Timeouts and Switchovers, page 5-8, for more information.

Figure 5-1 shows a protection group (cluster) of units controlled by a gateway with one unit configured as the standby for the protection group.

Figure 5-1 Prime Network Protection Groups—Example



In the example configuration, when the gateway determines that one of the units in the protection group has failed, it notifies the standby unit of the protection group to immediately load the configuration of the failed unit. The standby unit loads the configuration of the failed unit, including all AVMs and VNEs, and functions as the failed unit. We recommend that you have two standby units per cluster. In this case, if a unit fails, another standby unit is still available.

Because events are recorded in Prime Network Events, you can check for the specific problem and take action to bring the failed unit up again. When the failed unit becomes operational, you can decide whether to configure it as the new standby unit or to reinstate it to the protection group and configure another unit as the standby unit.

AVM 100 and Unit Server High Availability

You can configure AVM 100 to run on a unit instead of the gateway. If the unit is also configured with high availability, the AVM 100 on the standby unit will drop all events because it is not running. This is by design; it should not start until a switchover occurs.

The standby unit contains a port watchdog script that listens for events on the unit's Syslog and SNMP ports. The script prevents unnecessary ICMP unreachable messages being sent back to the network. If a switchover occurs, the standby unit and AVM 100 will start, and the watchdog script releases the ports.

When the original unit comes back up, the standby AVM 100 goes back down, and the watchdog script recommences listening on the standby unit's Syslog and SNMP ports.

What is the Impact of Unit or AVM Failures?

When a failure occurs in a unit or AVM, the length of time that the system is down depends on the type of failure, how long it takes to detect that the component is not working, and the length of the recovery period (during which the unit or AVM reloads and the system begins to function normally again).

These topics describe what you can expect to happen if there is a failure:

- Impact of AVM Process Failure, page 5-4
- Impact of Unit Timeouts and Switchovers, page 5-8

Impact of AVM Process Failure

These topics explain the impact of two very different failure scenarios—when individual AVMs are stopping and restarting, and when there is a catastrophic failure.

- Impact of AVM Timeouts and Restarts, page 5-4
- Impact of Catastrophic AVM Process Failure, page 5-6

Impact of AVM Timeouts and Restarts

Each AVM is constantly monitored by the management AVM (AVM 99) using a watchdog protocol pulse message sent to the AVM at preconfigured intervals. When the AVM fails to respond to the pulse message after a preconfigured number of attempts, the management AVM restarts the process.

The management process also keeps a history of the number of times it has restarted the AVM. When it reaches the maximum number of preconfigured restart times, the management AVM stops restarting the AVM because this indicates a serious problem with the AVM. Each restart is logged as a System event (except when AVM 11 is restarted, because this AVM handles all persistency).

Failures on AVMs in the system are measured in a way similar to that used for catastrophic process failures (see Table 5-1), with the addition of the watchdog protocol overhead. This is measured by the pulse interval multiplied by the number of restart attempts.

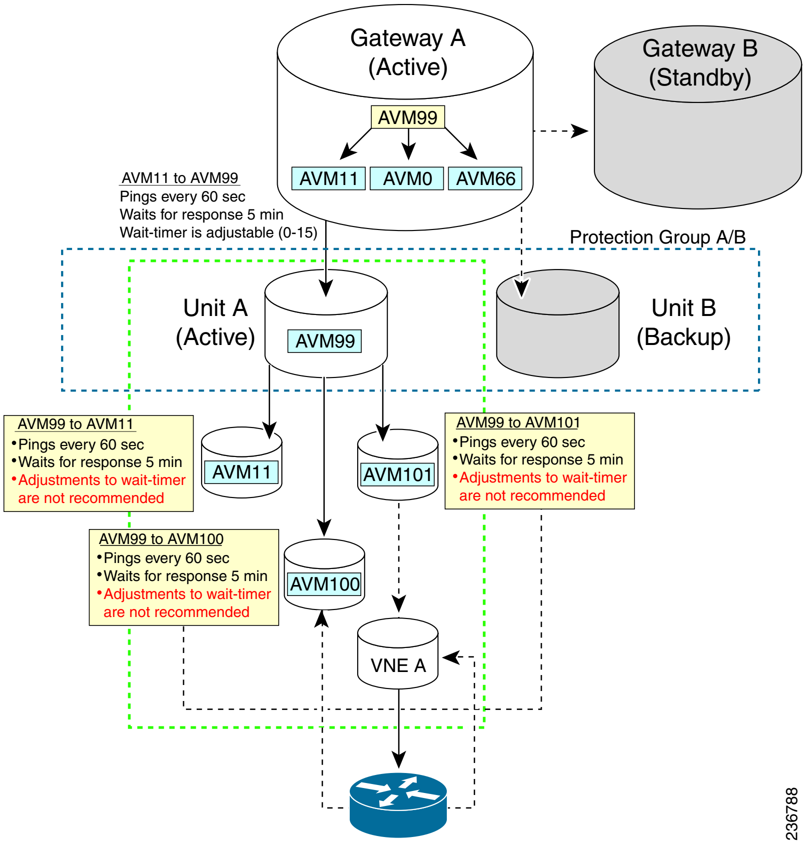
Keep the following in mind when evaluating an AVM failure:

- The maximum number of preconfigured restart times is five, after which the management process does not try to reload the AVM.
- It takes approximately one minute for the system to detect that an AVM (including AVM 100) is not working.
- The recovery period during which an AVM (including AVM 100) reloads and the system starts to function normally again is approximately five minutes, depending on the number of VNEs per AVM and the complexity of each.

Figure 5-2 provides a typical example of how unit server high availability timer parameters work while monitoring AVMs.

Note If you are using gateway server high availability, there is no overlapping between the processes that AVM 99 monitors that are illustrated in Figure 5-2, and the process that the gateway high availability software monitors.

Figure 5-2 Unit Server High Availability Parameter Timers and AVM Monitoring Example



Measuring Fault-Processing Down Time for AVMs

When a failure occurs on an AVM, the time during which ticket processing is down is measured as the sum of the following factors:

- The time it takes to determine that the AVM has failed.
- The time it takes for the AVM to reload, depending on the number of VNEs.
- The time it takes to pass syslogs or traps to the VNEs (in the case of AVM 100), or to pass events to the gateway (in the case of AVM 101-999).

Note For the first 30 minutes after AVM 99 (the management AVM) has started, there is no monitoring of the system to find unit server high availability issues. This allows the system enough time to get up and running.

Impact of Catastrophic AVM Process Failure

Each AVM has a log file which is constantly monitored by a Perl process for log messages about catastrophic failures, such as AVM processes running out of memory. When such a failure occurs, the Perl process restarts the AVM almost immediately, so the mean time to repair (MTTR) is based on the AVM loading life cycle.

Table 5-1 describes the impact on different AVMs when experiencing such a failure. For information on the Operations Reports AVM (AVM 44), refer to the *Cisco Prime Network 4.2. Operations Reports User Guide*.

Table 5-1 Catastrophic Process Failure Impact on AVMs

AVM Process	Results of AVM Failure	Average Time To Repair Failed AVM	Degree of Impact to System if AVM Fails
AVM 0 (High availability/ switch)	Loss of messages to and from the machine.	1 minute to reach bootstrap.	High. Messages are constantly being sent and received in the system.
AVM 11 (Gateway)	Loss of persistence information for faults (except for the I persistency information handled by AVM 25 and AVM 100). No user authentication will be performed on gateway connections, and GUI clients will lose gateway connectivity.	6-10 minutes to reach bootstrap.	High. AVM 11 handles Oracle communication and various gateway functions such as alarm processing.
AVM 25 (Event persistence)	Loss of persistence information and new tickets for actionable events that are processed while AVM 25 is down. When it comes up, new events that correlate to "lost" events will be persisted but will *not* be associated with a ticket until the integrity process identifies the broken chains (due to lost events) and creates new tickets.	1 minute to reach bootstrap.	High. Network events are constantly processed in a live, scaled system.
AVM 35 (Service discovery)	Network services displayed on maps (such as Ethernet service and MPLS-TP) are not updated to reflect network changes.	Depends on network size. While AVM 35 only needs 1 minute to reach bootstrap, time to repair depends on network size (to redisplay already-discovered services, detect changes that occurred when AVM was down). Could be 30 minutes to 10 hours.	Low for small networks, higher for larger networks because network services display would be updated after a discovery resync process is finished.

Table 5-1 Catastrophic Process Failure Impact on AVMs (continued)

AVM Process	Results of AVM Failure	Average Time To Repair Failed AVM	Degree of Impact to System if AVM Fails
AVM 41 (Compliance Audit)	Compliance Audit functionality would not work. Compliance Policy and Policy Profile Page would not show Policies and Profiles respectively.	1 minute	Low, because only Compliance Audit functionality would be impacted.
AVM 44 (Operations Reports)	Operations Reports inventory data would not be in sync with the network.	Depends on network size. While AVM 44 only requires 1 minute to reach bootstrap, time to repair depends on network size (to discover new devices, resync with Operations Reports database). Resync will begin within 1 minute after AVM 44 is restarted.	Depends on how frequently reports are used in the deployment.
AVM 45 (Infobright database)	Operations Reports Infobright databases at local and remote sites would be out of sync.	Depends on how much data needs to be resynchronized between the local and remote sites.	High, because there would be a lapse in redundancy between the local and remote sites.
AVM 76 (Job scheduler)	No jobs can be added, executed, or removed.	1 minute to reach bootstrap.	Depends on job types.
AVM 77 (Change and Configuration Management)	Loss of device configuration changes. Configuration changes will not be backed up to the archive during down time.	10 minutes for DM server startup and bundle deployment, plus time to fetch all configurations for managed devices.	High (if using Change and Configuration Management), depending on network size and frequency of change notifications.
AVM 78 (VNE topology)	Topology links between VNEs on different units will not discovered.	1 minute to reach bootstrap.	Low; there may be some missing topology links.
AVM 83 (TFTP server for Change and Configuration Management)	Change and configuration management TFTP operations will fail. (Operations using secure protocol or FTP will not be affected.)	5 minutes.	High (if using Change and Configuration Management); Change and Configuration Management device properties would fail.
AVM 84 (Reports)	Loss of reports.When AVM 84 is down running reports will fail.	1 minute.	Low; reports would need to be rerun.
AVM 99 (Management)	Loss of registry notifications on changes made to golden source registry.	1 minute to reach bootstrap.	Low, because registry modifications are made only when the VNE is first loaded into the system. Modifications are rarely made while the system is up and running. For the first 30 minutes after AVM 99 has started, there is no system monitoring for unit server high availability. This allows the system enough time to get up and running

Table 5-1 Catastrophic Process Failure Impact on AVMs (continued)

AVM Process	Results of AVM Failure	Average Time To Repair Failed AVM	Degree of Impact to System if AVM Fails
AVM 100 (Event Collector)	Loss of traps and syslogs from devices, including raw event persistency.	1 minute to reach bootstrap, plus time for all the VNEs to register again for traps and syslogs. Normally a matter of minutes.	High, because raw events from devices are constantly received in a live, scaled system. Only devices registered to the failed AVM 100 are affected. No events will be handled during downtime. See AVM 100 and Unit Server High Availability, page 5-3. (Raw event persistency is recovered before events are forwarded to the VNEs.)
AVM 101-999 (User-defined AVMs)	Loss of management to a section of devices managed by the AVM; alarm state inconsistencies (user will have to clear tickets).	1 minute to reach bootstrap, plus time to load the VNEs (depending on number, type, services, etc.).	High (but only for a period of one minute), because no raw events sent to the VNEs can be processed when the AVM is down.

Impact of Unit Timeouts and Switchovers

The Prime Network gateway constantly monitors units by sending a watchdog protocol pulse message to the unit management AVM at preconfigured intervals. If the unit management AVM fails to respond to the pulse message after a preconfigured number of retries, the gateway loads the standby unit to replace it.

The impact of such a failure on the system is that the unresponsive unit does not manage the devices for a period of time. This unmanaged period of time is measured by the pulse interval multiplied by the number of retry times, plus the unit load time.

Note Unit load time depends on the configuration of the unit—the hardware, the number of VNEs, the types of VNEs, and the services running on the VNEs. All of these factors impact the load time required for the VNEs to complete their modeling, as described in Table 5-1.

(On the other hand, if the problematic unit has not completely failed and continues to operate *after* the switchover, you may see duplicate events in the Oracle database. In this case you should stop the original problematic unit using **networkctl stop**.)

Measuring Ticket-Processing Down Time for Units

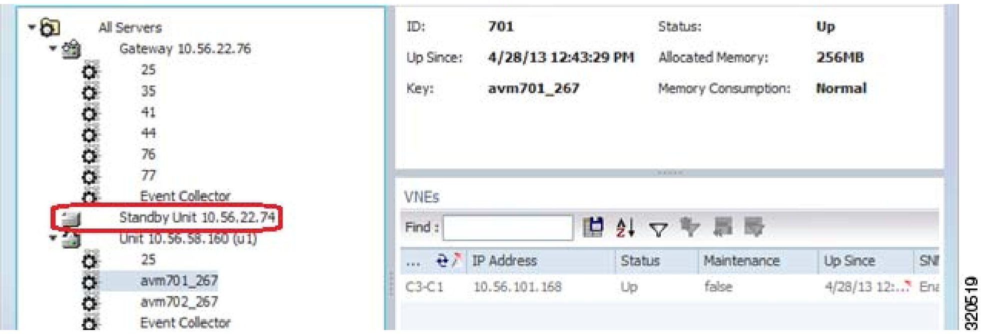
When a failure occurs on a unit, the time during which ticket processing is down is measured as the sum of the following factors:

- The time it takes to determine that the unit has failed (depending on the ping interval).
- The time it takes for the unit to reload, depending on the number of AVMs and VNEs in the unit.
- The time it takes to pass correlated events to the gateway (a minimum of five minutes to obtain device history, plus a variable time depending on the number of VNEs per AVM).

Creating a New Unit Protection Group

New units are added to Prime Network during installation as described in the *Cisco Prime Network 4.3.1 Installation Guide*. During installation, you can specify whether the unit is a standby unit, and which protection group the unit should belong to (the default is default-pg). If you create a standby unit, it is displayed in the Administration GUI client as shown in Figure 5-3. If you look at the standby unit's properties, its status will be standby.

Figure 5-3 Standby Unit in Administration GUI Client



ID:	701	Status:	Up
Up Since:	4/28/13 12:43:29 PM	Allocated Memory:	256MB
Key:	avm701_267	Memory Consumption:	Normal

...	IP Address	Status	Maintenance	Up Since	SN
C3-C1	10.56.101.168	Up	false	4/28/13 12:...	Ena

Before You Begin

Keep the following guidelines in mind when configuring protection groups:

- Design protection groups according to geography.
- Add an additional standby unit to heavily-loaded groups.
- Do not assign active and standby units to more than one group.
- Units in a group must have the same operating system. If any unit has a database connection, all other units must also have a connection.

To create or edit a protection group:

Step 1 Create the new protection group.

a. Choose **Global Settings > Protection Groups**.

b. Open the New Protection Group dialog box by right-clicking **Protection Groups**, then choose **New Protection Group**. For an existing group, right-click the group and choose **Properties**.

c. Enter a name and description, or edit the description.

d. Click **OK**. The content area displays details of the new protection group and all currently defined protection groups in the Protection Groups table.

Step 2 Add units to the new protection group. Units should not belong to multiple protection groups.

a. Right-click the unit and select Properties.

b. In the Protection Group drop-down list, select the new protection group and click **OK**.

Switching to a Standby Unit (Disable Active Unit)

Prime Network will automatically switch over to a standby unit occurs when the gateway discovers that one of the active units has failed. Such failures include hardware failures, operating system failures, power failures, and network failures, which disconnect a unit from the Prime Network fabric. If the protection group has more than one standby unit, Prime Network randomly selects the standby unit.

Note If the problematic unit has not completely failed and continues to operate *after* the switchover, you may see duplicate events in the Oracle database. In this case you should stop the original problematic unit using **networkctl stop**.

You can also perform a manual switchover to a standby unit (for example, if you have to shut down the unit for maintenance).

When a switchover occurs, Prime Network automatically transfers all data from the failed unit to a standby unit in the same protection group. The original unit is removed from the standby setup and is no longer displayed in Prime Network Administration.

Note When a unit switches to its standby, all VNEs on the unit that were in maintenance mode will be moved to the VNE Down state.

To manually switch to a standby unit:

Step 1 Expand the All Servers branch and select the required unit.

Step 2 Right-click the required unit, then choose **Switch**. A confirmation message is displayed.

Step 3 Click **Yes**. The standby unit becomes the active unit and is displayed in the All Servers branch. The original unit is removed from the setup and can be safely shut down. It is no longer displayed in the All Servers branch in the navigation tree.

Changing Timeouts and Restarts for Unit and Process Protection

The AVM process and unit protection functions are controlled by settings in the registry. The registry entries and default values are provided in Table 5-2.

Caution Increasing these values allows AVM or unit failures to last longer, but it also increases the certainty that a failure has actually occurred. However, decreasing these values can result in a "false positive." In other words, the shorter allowable AVM or unit failure period can result in unnecessary AVM restarts or unit switchovers when an AVM or unit is simply busy processing a large amount of data.

Table 5-2 Registry Settings for Unit Server High Availability and AVM Watchdog Protocol

Registry Entry	Description	Default Value
agent_defaults/delay	Grace period (in milliseconds) during which events are not raised. This defines the amount of time during which the system does not perform high availability operations of any kind on the configured target (either the AVM or the unit). The grace period begins at system startup. The only exception is when the configured target responds for the first time with a ping; at that point the grace period is over.	1800000 (30 minutes)
agent_defaults/timeout	AVMs recovery period (in milliseconds). This period includes device polling and inventory buildup. (End-to-end services, such as RCA and topology, can take longer before they become available.)	300000 (5 minutes)
haservice/timeout	Units recovery period (in milliseconds).	300000 (5 minutes)
agent_defaults/maxTimeoutReloadTime	Threshold for permitted AVM retries (in milliseconds). When exceeded, the AVM is suspended.	1800000 (180 minutes)
agent_defaults/maxTimeoutReloadTries	Maximum number of AVM retries. When exceeded, the AVM is suspended.	5

Tracking Unit and Process Protection Events

The following table provides ways you can get historical information on unit high availability and AVM process protection events. You can tailor your search or reports by specifying keywords (switchover, high availability, watchdog protocol, and so forth).

For historical events related to:	See:
AVM process protection Unit high availability Protection groups and unit switchovers	AVM and other appropriate log files (see Log Files Reference, page C-3) The following reports, which you can launch from the main menu by choosing **Reports > Run Report > Events Reports > Detailed Non-Network Events**: • Detailed System Events • Detailed Security Events