CHAPTER 16

Unit Server High Availability and AVM Protection

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:

For information on high availability for gateway servers, see Using Veritas Gateway Server High Availability, page 17-1 and Using RHCS/ADG Gateway Server High Availability, page 18-1.

Overview of Unit Server High Availability

High availability of the server backbone is achieved at several complementary levels. For example:

- NEBS-3 compliant carrier-class server hardware.
- Watchdog within each unit, responsible for monitoring and, if necessary, automatically reloading failed processes.
- N+m warm standby protection for unit groups.

See the following topics for more information:

Watchdog Protocol for AVM Protection

The *watchdog protocol* monitors the AVM processes to make sure any AVMs that have failed are restarted. This is called *AVM protection* and the GUI, the watchdog protocol is controlled by the AVM Protection check box. Each unit executes several processes: one control process and several AVM processes that execute VNEs. Each process within the unit is completely independent. The isolation

concept is tailored throughout the design 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.

The control process executes a watchdog protocol, which continuously monitors all other processes on the unit. This 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 cancelled and reloaded.

The dynamic design of the control process implements runtime adaptation and escalation. The escalation procedure moves the AVM to suspended mode; that is, the process is suspended. An example of an escalation procedure is to stop reloading a process that has crashed more than *n* times within a given period, 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 Chapter 26, "VNE Persistency Mechanism."

All watchdog protocol parameters, such as *pulse interval* and *retry times*, are configurable in the registry. The higher these parameter values are, the longer the AVM or unit failure lasts, but this increases the certainty that a failure has actually occurred. Configuring these parameters with lower values may shorten the AVM or unit recovery, but might result in a "false positive" which could unnecessarily restart an AVM or revert to a standby unit when the AVM is just busy or the unit is processing a heavy load of data. For information on these registry settings, see High Availability Registry Settings, page 16-12.

Unit N+m High Availability

The clustered N+m unit server high availability mechanism uses the Prime Network fabric is designed to handle the failure of a unit. Such failures include hardware failures, operating system failures, power failures, and network failures, which disconnect a unit from the Prime Network fabric.

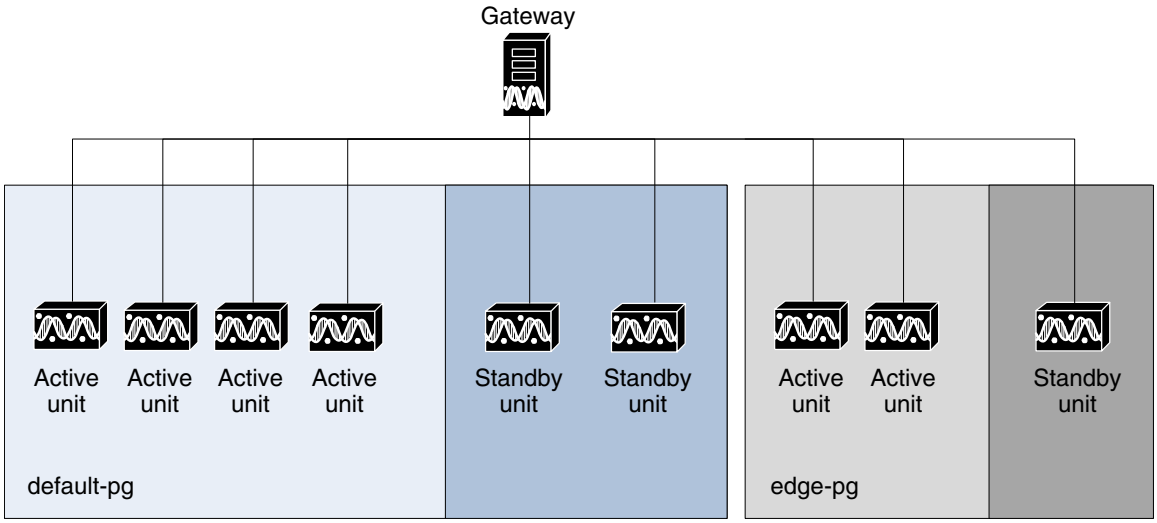
Unit availability is established in the gateway, running a *protection manager* process, which continuously monitors all the units in the network. Once the protection manager detects a unit that is malfunctioning, it automatically signals one of the standby servers in its cluster to load the configuration 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 a unit is configured, it can be designated as being 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.

A unit switchover results in the unavoidable loss of information. The impact depends on how long the unit is down, and the functions the unit performed.

Figure 16-1 shows a protection group (cluster) of units controlled by a gateway with one unit configured as the standby for the protection group.

Figure 16-1 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.

Recommendations for Configuring High Availability

Keep the following guidelines in mind when configuring protection groups:

Note Units in a protection group must have the same operating system.

- Protection groups should be designed according to geography.
- For heavily loaded protection groups, add an additional standby unit.
- Units (active and standby) should not be assigned to more than one protection group.

Estimating 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).

Three types of failure can occur, as described in these topics:

- Impact of Catastrophic AVM Process Failure, page 16-4
- Impact of AVM Timeouts and Restarts, page 16-6
- Impact of Unit Timeouts and Switchovers, page 16-8

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 16-1 describes the impact on different AVMs when experiencing such a failure.

Table 16-1 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.

Table 16-1 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 25 (Event persistence)	Loss of persistence information and new tickets for actionable network 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 opens 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.	1 minute to reach bootstrap, plus several minutes to redisplay already discovered services, plus time required to detect changes that occurred when the AVM was down (30 minutes to 10 hours, depending on number, type, services, etc.).	Low: Network services display would be updated after a discovery resynch process is finished.
AVM 66 (Workflow engine)	Running workflows would abort and scheduled workflows would not run. Templates would not be deployed.	1 minute to reach bootstrap, but with large number of workflows in the system, this may increase.	Low, because AVM 66 should sustain a large number of executed workflows (per the system limitations). Templates would need to be redeployed and aborted workflows would need to be rerun. Check the Provisioning events in Prime Network Events to verify what ran prior to failure, and then issue an rollback (no automatic rollback is done).
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);because configuration change notifications can happen all the time.
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.

Table 16-1 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 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
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 16-3. (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 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 16-1), 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:

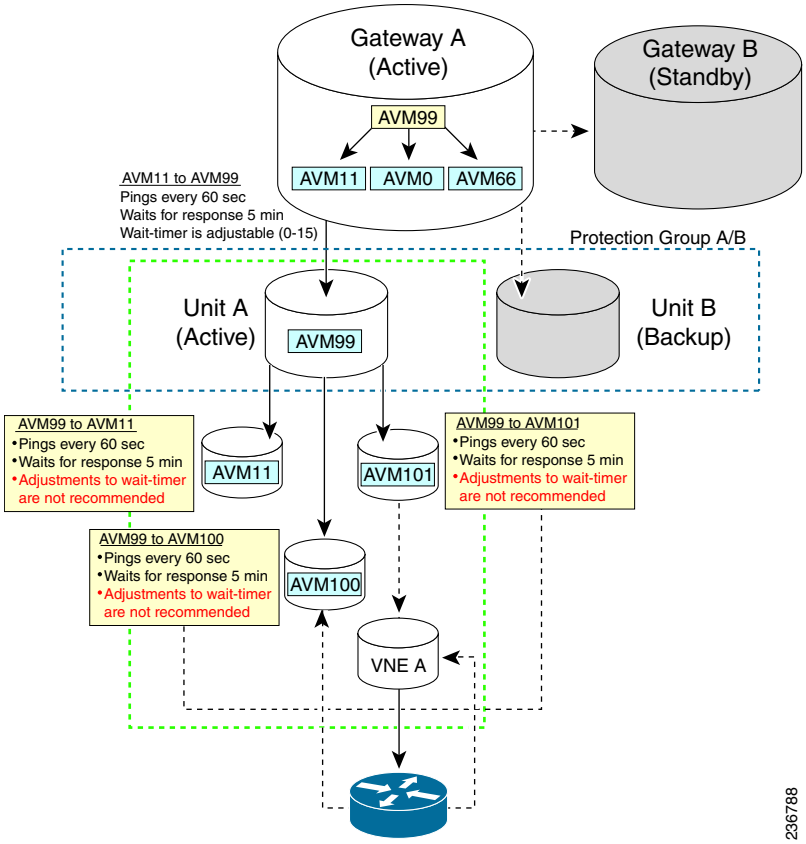
- The maximum number of preconfigured restart times is five, after which the management process does not try to reload the AVM.
- It takes approximately one minute for the system to detect that an AVM (including AVM 100) is not working.

- 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 16-2 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, note that there is no overlapping between the processes that AVM 99 monitors that are illustrated in Figure 16-2, and the process that Veritas Cluster Manager monitors (the ANA Gateway Veritas agent). For an illustration, see Figure 17-7 on page 17-10.

Figure 16-2 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 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 16-1.

(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 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 Unit Protection Groups and Designating Standby Units

New units are added to Prime Network using the installation scripts described in the *Cisco Prime Network 3.8 Installation Guide*. By default, units are added to a protection group named default-pg. Each protection group, or cluster of units, should have at least two standby unit servers.

Note Units in a protection group must have the same operating system.

These topics explain how to create new protection groups and work with standby units:

- Creating a Protection Group and Adding Units to the New Group, page 16-9
- Configuring Standby Units, page 16-9
- Switching to a Standby Unit, page 16-10

Creating a Protection Group and Adding Units to the New Group

By default, all units in the Prime Network fabric belong to one group (or cluster), the default-pg protection group. You can create additional groups as your network grows. You should have at least two standby units for each cluster.

Note Units in a protection group must have the same operating system.

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.

a. Right-click the unit and select Properties.

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

Configuring Standby Units

Prime Network Administration enables you to configure standby units and assign standby units to protection groups.

To configure a standby unit:

Step 1 If you are changing an active unit into a standby unit, first disconnect the active unit to stop the gateway-unit communication. Right-click the unit and select **Disconnect**.

Step 2 Verify the protection group settings for the unit by right-clicking the unit and selecting **Properties**. Make sure the correct protection group is chosen from the drop-down list.

Step 3 Change the unit properties so that the unit becomes a standby for the protection group.

a. Right-click the unit in the All Server branch and select **Properties**.

b. Check the Enable Unit Protection check box and click **OK**.

Note Standby units are not displayed in the All Servers branch in the navigation tree.

Switching to a Standby Unit

Prime Network Administration enables you to switch to a standby unit either manually or automatically.

- Automatic switchover 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. For more information on automatic switchover, see Unit N+m High Availability, page 16-2.

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

- Manually switching to a standby unit is useful if you must temporarily 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 Prime Network Administration window.

Note In the event of unit failover, the Prime Network gateway randomly selects a redundant unit when more than one standby unit is available.

Managing the Watchdog Protocol (AVM Protection)

The following topics describe how to define AVMs for units and enable or disable protection (the watchdog protocol) on the AVM:

- Enabling AVM Protection (Watchdog Protocol) on AVMs, page 16-10
- Viewing and Changing AVM Protection (Watchdog Protocol) Settings, page 16-11

Enabling AVM Protection (Watchdog Protocol) on AVMs

Every AVM in the Prime Network fabric is, by default, managed by the watchdog protocol. Prime Network Administration enables you to define AVMs for units and enable or disable the watchdog protocol on each AVM.

To define an AVM:

- The unit must be installed.
- The unit must be connected to the transport network.
- The following default AVMs must be running:
  - AVM 0—The switch AVM.
  - AVM 99—The management AVM.
  - AVM 100—The trap management AVM (one instance must be running either on the gateway server or one of the units).
- The new AVM must have a unique identifier within the unit.

Note For detailed information on defining AVMs, see Viewing AVM Properties, page 4-6.

To enable AVM protection on an AVM:

Step 1 Open the New AVM dialog box by right-clicking the required unit (or gateway), then choose **New AVM**.

Step 2 Define the properties of the AVM. For more information, see Viewing AVM Properties, page 4-6.

Step 3 Check the **Enable AVM Protection** check box to enable the watchdog protocol.

Note We strongly recommended that you do not uncheck the Enable AVM Protection check box.

Step 4 Click **OK**. The new AVM, with the watchdog protocol enabled, is added to the selected unit and is displayed in the content area.

Adding the new AVM creates the registry information for the new AVM in the specified unit. The AVM can now host VNEs.

Viewing and Changing AVM Protection (Watchdog Protocol) Settings

Note For detailed information on defining and editing AVMs, see Chapter 4, "Basic AVM and VNE Administration Tasks."

To view and edit AVM settings:

Step 1 Open the AVM Properties dialog box by right- clicking the required AVM, then choose **Properties**.

Step 2 Edit the details of the AVM, as required.

Note We strongly recommended that you do not uncheck the Enable AVM Protection check box.

Step 3 Click **OK**. The new properties for the AVM are displayed in the content area.

High Availability Registry Settings

The high availability and AVM watchdog protocol functions are controlled by settings in the registry. The registry entries and default values are provided in Table 16-2.

Note All changes to the registry should only be carried out with the support of Cisco. For details, contact your Cisco account representative.

Table 16-2 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. The grace period begins at system startup. It 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). There is one exception: When the configured target responds for the first time with a ping, the grace period is over.	1800000 (30 minutes)
agent_defaults/timeout	Timeout (in milliseconds) for AVMs. This is the initial recovery period. 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	Timeout (in milliseconds) for units.	300000 (5 minutes)
agent_defaults/maxTimeoutReloadTime	Threshold (in milliseconds) for AVM reload retries. When exceeded, the AVM is suspended.	1800000 (180 minutes)
agent_defaults/maxTimeoutReloadTries	Maximum number of retries for AVM reloads. When exceeded, the AVM is suspended.	5