WHITE PAPER

September 1995

Prepared By
NetWare OS
Development

Compaq Computer
Corporation

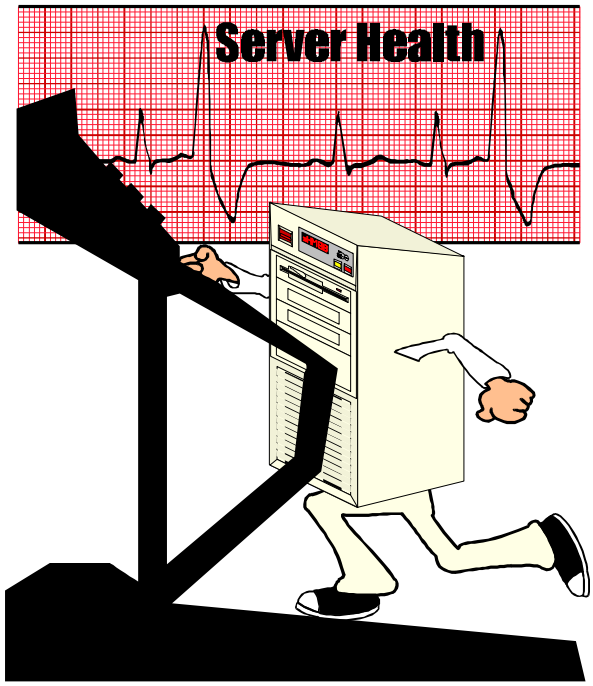
Diagnosing NetWare Server Failures

Server management is a set of Compaq features which record, diagnose and react to events that cause or may cause a server failure. The purpose of server management is to keep the server operating as continuously as possible by diagnosing the causes of failures, while protecting the hardware components. The Server Health driver, one component of server management, logs valuable information regarding server failures. A special feature of the Server Health driver is Automatic Server Recovery (ASR), which helps detect additional server failures and can reboot the server after a failure.

Server management features allow the unattended server to detect, diagnose and react to critical problems that usually require human intervention. This paper explains the features of the Server Health driver and how to use the Compaq error logs to better diagnose, correct and prevent future server failures.

Server management features are available on the Compaq ProSignia Family of Servers, Systempro/XL, and ProLiant Family of Servers.

CONTENTS



COMPAQ

NOTICE

The information in this publication is subject to change without notice.

COMPAQ COMPUTER CORPORATION SHALL NOT BE LIABLE FOR TECHNICAL OR EDITORIAL ERRORS OR OMISSIONS CONTAINED HEREIN, NOR FOR INCIDENTAL OR CONSEQUENTIAL DAMAGES RESULTING FROM THE FURNISHING, PERFORMANCE, OR USE OF THIS MATERIAL.

This publication does not constitute an endorsement of the product or products that were tested. The configuration or configurations tested or described may or may not be the only available solution. This test is not a determination of product quality or correctness, nor does it ensure compliance with any federal, state or local requirements. Compaq does not warrant products other than its own strictly as stated in Compaq product warranties.

Diagnosing NetWare Server Failures
First Edition (September 1995)
Document Number 396A/0995

WHY DO I NEED THE SERVER HEALTH DRIVER?

Server management in general and the Server Health driver in particular are useful whenever the reliable operation of a server is imperative. Although Compaq server failures are rare, the unexpected down time they cause makes detecting and reacting to server failures and other critical events essential. The Server Health driver can help reduce down time and maintain the integrity of your server in the following ways:

- Inspect the state of the server after a failure has occurred
- Automatically reboot the server after a failure has occurred
- Log corrected ECC memory errors
- Logging of critical hardware and software failures
- Gracefully down the server when overheating or loss of power is predicted

Server Failures

The most important feature of the Server Health driver is its ability to perform diagnostics on a server after a failure. The Server Health driver identifies the program or piece of hardware that instigated the problem, allowing you to take appropriate actions to prevent the failure from reoccurring.

Automatic Server Recovery (ASR)

There are two reasons to enable the Automatic Server Recovery (ASR) option:

- ASR detects failures not detected by NetWare, including server lock-ups
- ASR reboots a server after any failure
- ASR protects server from hostile environmental conditions

ASR detects server failures by polling a hardware timer in Compaq servers. This hardware timer allows ASR to detect server failures unnoticed by NetWare. When the timer reaches a specified time-out value, ASR will reboot the server, allowing normal operation to be restored. Quick ASR, one of Compaq's newest ASR features, further reduces down time by rebooting the server immediately after NetWare detects a failure, instead of waiting for the ASR time-out. This greatly reduces down time should a server failure occur at any time the server is unattended.

ASR also protects your server against hostile environmental conditions such as overheating. Whenever such a condition occurs, the server will shutdown. This shutdown prevents damage to hardware components. If you are using the Server Health driver and have the ASR and Thermal Shutdown option enabled in the Compaq System Configuration Utility, pending I/O requests will complete, users are informed, and data stored in cache will be flushed to the hard drive before the server is downed.

PROBLEMS ADDRESSED BY THE SERVER HEALTH DRIVER

The following sections describe the problems addressed by the Server Health driver--server failures, correctable memory errors, hostile thermal conditions, and power loss--and detail how the Server Health driver reacts to them.

Handling of Server Failures

The Server Health driver divides the causes of server failures into five basic categories:

- Software Exception ABENDs
- Processor Exception ABENDs
- Non-Maskable Interrupt ABENDs (NMIs)
- Lockups (ASR must be enabled for detection.)
- Debugger Entry (ASR must be enabled for detection.)

Software exceptions, processor exceptions, and NMIs are detected by NetWare and are called Abnormal Ends (ABENDs). When the OS, the processor or the hardware detect a catastrophic failure, NetWare produces an ABEND and halts the invalid action. Halting the system prevents data corruption. If the Server Health driver is loaded when the failure occurs, NetWare reports the ABEND to the driver. The Server Health driver can then add supplemental information to the NetWare ABEND message. The ABEND message is then displayed on the monitor and added to the critical error log. The following sections provide more detail on each of these categories and how to service each problem.

Software Exceptions and Processor Exceptions

Both software exceptions and processor exceptions are typically caused by software errors. Processor exceptions are errors caused by software, but discovered by the processor. Processor exceptions are identifiable by the phrase "Processor Exception" in the ABEND message and in the critical error log.

The critical error log includes additional information not found in the ABEND message, including the name of the program running at the time of the ABEND, its version number and its offset at the time of the ABEND. If your server experiences a processor exception or software exception, you should upgrade to the newest version of the program identified in the critical error log and apply any relevant patches to the NetWare operating system. If there are no newer software versions or applicable patches available, contact the software vendor and provide them with the information found in the critical error log. This information will help the software vendor analyze the problem quickly.

The following is an example of a critical error log message for a software exception:

```
ABEND: Invalid Resource Tag Passed to kernel
Running Process: Init Process
Code executing in module BROKE.DSK v2.00g at offset +027Fh
```

The example above shows that the process running at the time of the ABEND was BROKE.DSK. The initialization failed for this driver. Upgrade this device driver and apply any OS patches.

The following is an example of a critical error log message for a processor exception:

```
ABEND: Page Fault Processor Exception
Running Process: Interrupt Service Routine
Code executing in module BADNLM.NLM v1.00 at offset +0A29h
```

The example above shows that the process running at the time of the ABEND was an interrupt service routine begun by the software module BADNLM.NLM. The information in the last section of this paper shows an error list and what action to take. In this case, you should upgrade any related software or contact your software vendor.

Processor exceptions may also be caused by faulty hardware. If you have eliminated the software as the cause of the problem, the hardware controlled by the software that was logged in the critical error log may be the problem. For example, if a device driver is listed in the error message and you are certain the driver is not the problem, the controller associated with that driver may be at fault. Contact the hardware vendor for more information or help with diagnosing the problem.

Non-Maskable Interrupts (NMIs)

NMIs are always caused by a hardware failure, such as memory or an expansion board, and are identifiable by the phrase "NMI" in the ABEND message in the critical error log.

NetWare reports most NMIs as "parity errors generated by the system board" because it is unable to diagnose hardware failures. When NetWare reports an NMI, the Server Health driver uses special hardware components in Compaq servers to identify the specific component which caused the failure. The Server Health driver then replaces the NetWare ABEND message with a detailed description of the hardware failure and logs the information in the critical error log.

You should run diagnostics on the hardware item identified in the critical error log and replace it, if necessary. On some Compaq systems, if a section of bad memory caused the error, that memory location is automatically mapped out by the System ROM when the server is rebooted (either manually or by ASR) to prevent future errors. Replacement in this case is not immediately necessary. You can replace any bad memory at the next scheduled maintenance.

The following is an example of a critical error log message for an NMI ABEND:

```
ABEND: NMI (Expansion bus master timeout: Slot 4)
```

The example NMI ABEND shows a problem with the board in slot 4. You should run Compaq Diagnostics on this board to try and determine what the problem might be. You may need to replace the board.

Lockups and Debugger Entry

Both lockups and debugger entry are caused by software errors not identifiable by NetWare. Lockups, identifiable when the server hangs or ceases to respond, are typically caused when software waits for operations which will never complete. The software never returns control to the operating system so the operating system cannot execute. These events are typically hardware interrupts or software counters.

Debugger entry, identifiable by the NetWare internal debugger screen on the server monitor, is caused by an illegal operation in software. Typically, developmental test code which instructs the server to enter the debugger was not removed by the software developer, or an illegal memory location was accessed, which forced NetWare to enter the debugger.

Lockups and debugger entry are the most difficult types of server failures to diagnose and correct because NetWare cannot detect them and does not produce ABEND messages. Debugger entry requires knowledge of the NetWare debugger in order to perform diagnostics, and a lockup prevents the administrator from performing any diagnostics. The administrator is left with almost no indication of why the failure occurred or what corrective action to take. ASR solves this problem by detecting the failure and allowing the Server Health driver to pinpoint the cause of the failure.

If a lockup or debugger entry occurs while ASR is enabled, the ASR timer will time-out, informing the Server Health driver that a failure has occurred. The Server Health driver identifies the program that was executing, its version number and offset at the time of the failure and records this information in the critical error log. ASR then reboots the server.

If your server experiences a lockup or debugger entry, upgrade to the newest version of the program identified in the critical error log and apply any relevant NetWare patches. If there are no newer software versions or applicable patches, contact the software vendor and provide them with the information found in the critical error log. This information will help the software vendor analyze the problem quickly.

The following is an example of a critical error log message for a lockup:

```
Code executing in module PADLOCK.LAN v4.83
at offset +3A1Fh when ASR NMI occurred
```

The following is an example of a critical error log message for debugger entry:

```
INT3 Debug Exception caused by code executing in module JUMPWEED.NLM
at offset +1724h when ASR NMI occurred
```

In both of the above examples, the critical error log identifies the software that cause the problem, PADLOCK.LAN and JUMPWEED.NLM, respectively. You should upgrade to the newest version of these programs. In both cases, ASR rebooted the server so that normal operation could continue.

Correctable Memory Errors

When the server detects a correctable memory error (for systems that have Error Correcting Code or Advanced Error Correcting Code memory systems), the server corrects the error and continues normal operation. The Server Health driver logs the event in the corrected memory error log and will report the condition to the server's console. If a memory location constantly reports correctable memory errors, the Server Health driver will stop reporting the errors to the corrected memory log to minimize the performance impact on the server. The driver will display a message to this effect on the server console. Memory errors will continue to be corrected.

If the Server Health driver reports that a memory location consistently needs correcting, schedule down time and replace the faulty memory as soon as possible to avoid a critical memory error.

Hostile Thermal Conditions and Power Loss

Hostile thermal conditions and power loss are caused by environmental problems, fan failure, or power supply failure. Since these problems prevent normal server operation, but do not cause data corruption, graceful server shutdown is the best course of action. In power loss situations, additional hardware is required to supply the power needed to perform a graceful shutdown.

Caution Temperature Detection

The Server Health driver monitors the thermal conditions inside the server. Should the temperature approach dangerous levels, the Server Health driver logs the event in the critical error log and, if ASR is configured, can gracefully down the server to protect it from overheating. When the temperature cools, the server is rebooted. This prevents damage to the hardware and decreases the likelihood of more severe problems in the future.

Fan Failure

Should a fan in the server fail, the temperature of the hardware components (such as the processor) may reach dangerous levels, even though the surrounding temperature may remain at acceptable levels. When the Server Health driver detects a fan failure, it logs the event in the critical error log and, if ASR is enabled, will gracefully down the server to protect it from

overheating. This prevents damage to the hardware, and decreases the likelihood of more severe problems in the future. The failed fan should be replaced, even if it still appears functional.

Loss of Commercial Power

The Uninterruptible Power Supply (UPS) detects loss of commercial power and provides battery power to the server after the loss. When the UPS battery is near depletion, the Server Health driver logs the event in the critical error log, and the UPS can gracefully down the server to preserve user data. When commercial power returns, the server is rebooted. A Compaq UPS and UPS driver are required to enable this feature.

Redundant Power Supply Failure

Failure of the primary power supply in a redundant power supply configuration will cause the secondary power supply to begin supplying power. The failure of the primary power supply is logged in the critical error log by the Server Health driver. Operations will continue normally, but the failed power supply should be replaced to restore fault tolerance in case of future failures.

Failing System Battery

The system battery is necessary for the storage of the system's non-volatile RAM, which contains the server's configuration information. The system battery also maintains the server's real-time clock. Should battery failure be imminent, the Server Health driver logs the event in the critical error log. Operations will continue normally, but the battery should be replaced within seven days to preserve the configuration data.

FLOW OF EVENTS FOR SERVER FAILURES

As mentioned earlier, you can configure the Server Health driver to automatically reboot the server by enabling ASR. If ASR is enabled, it will reboot the server when it reaches a time-out after a server failure. If Quick ASR is used, it will reboot the server immediately after the ABEND message is logged.

Note: If the Server Health driver is loaded when an ABEND occurs, it will log the event information, whether or not ASR is enabled. However, if ASR is disabled and a lockup or debugger entry occurs, the failure cannot be detected, and therefore no information will be logged.

Use the following charts to understand the events that will take place with and without ASR enabled.

Figure 1 shows the flow of events when a NetWare server failure occurs without the Server Health driver loaded:

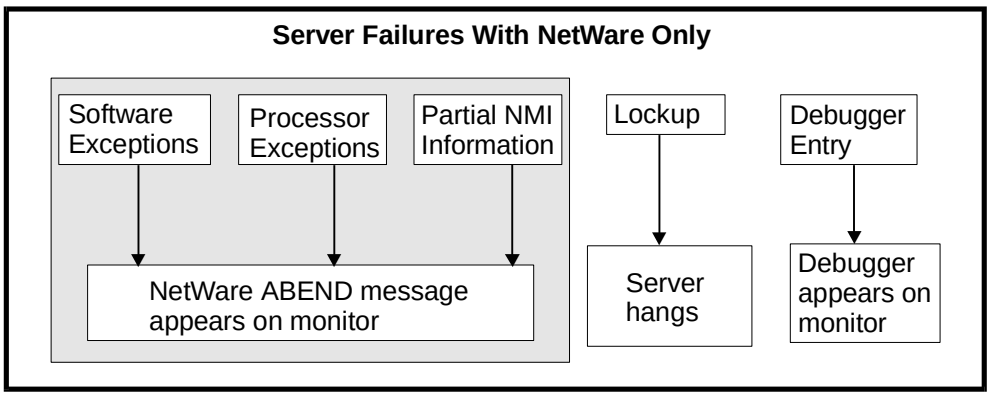


Figure 1: Server Failures with NetWare only

Figure 2 shows the flow of events when a server failure occurs with the Server Health driver loaded, but ASR not enabled:

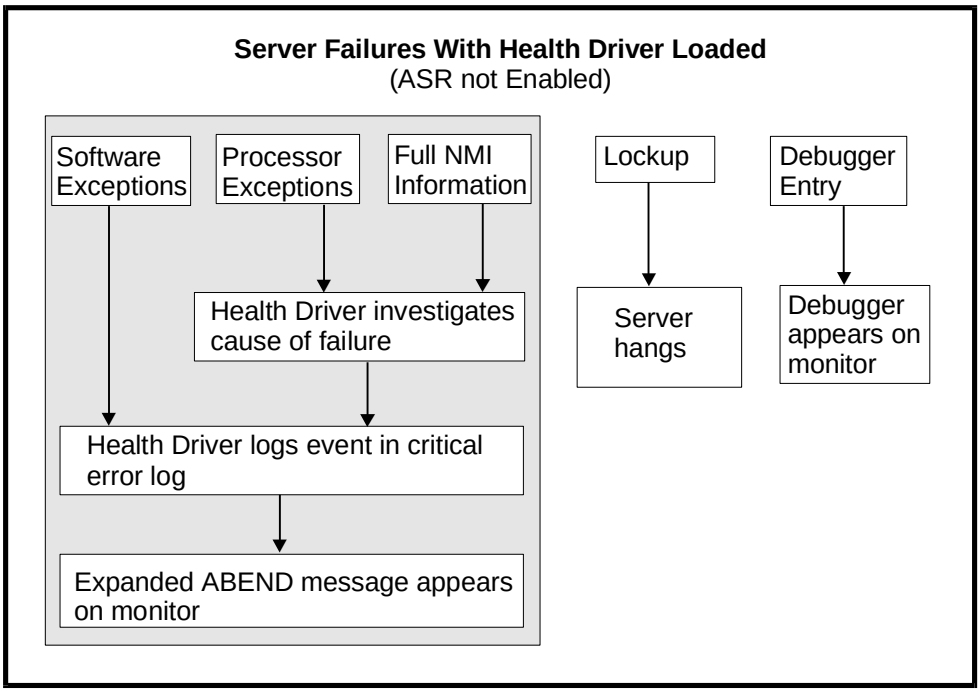


Figure 2: Server Failures with Server Health driver

Figure 3 shows the flow of events when a server failure occurs with the Server Health driver loaded and ASR enabled. If Quick ASR is used, it immediately reboots the server after software exceptions, processor exceptions, and NMIs, instead of waiting for a time-out to reboot the server.

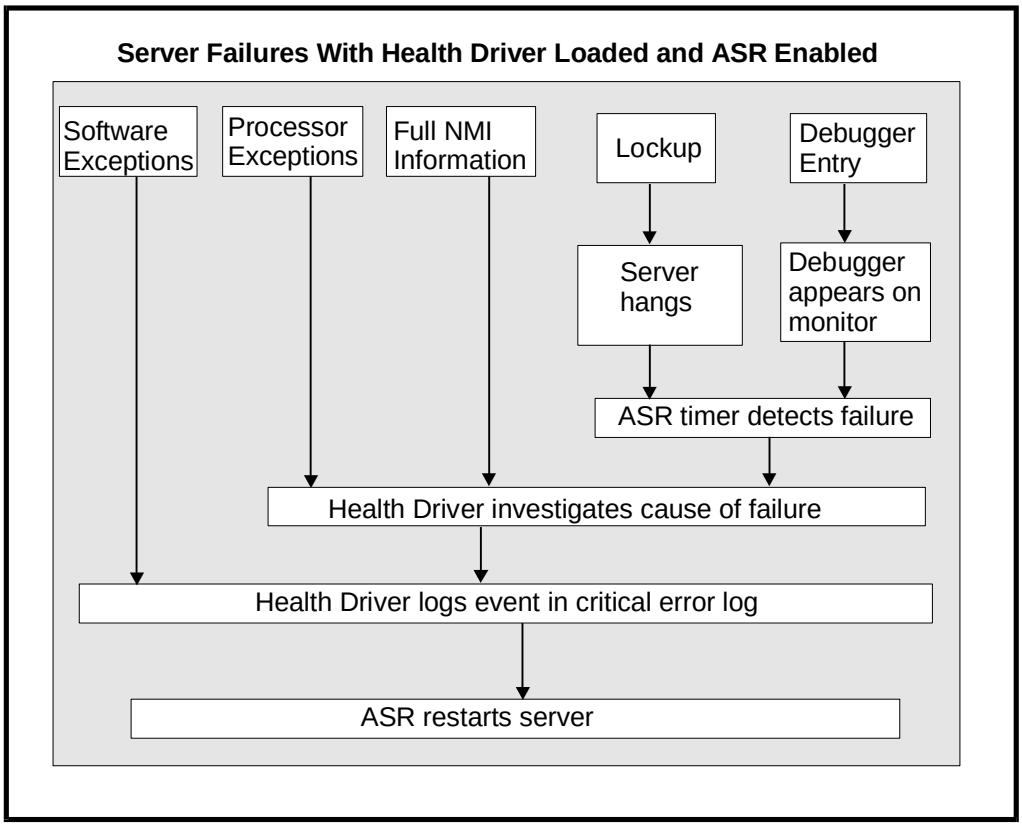


Figure 3: Server Failures with ASR

INSTALLATION AND CONFIGURATION

The following sections describe how to install and load the Server Health driver and how to configure your system to take full advantage of its features. To install the Server Health driver, CPQHLTH, perform the following steps:

1. Log into a client workstation as SUPERVISOR.
2. Insert the NPFC diskette in the A drive.
3. Execute the following MAP command to ensure you have a path to the server:

   ```
   MAP G:=SYS:SYSTEM
   ```

4. To support the Compaq Server Health Logging features, copy the Compaq Server Health driver for NetWare to the server:

   ```
   COPY A:\HEALTH\*.* G:
   ```

5. Log out from the server and return to the server console to complete the installation.

6. Add the following command to AUTOEXEC.NCF so the driver will be loaded when the NetWare server is restarted in the future:

   ```
   LOAD CPQHLTH
   ```

   If you wish to use the Quick ASR feature, add the -Q parameter to this command:

   ```
   LOAD CPQHLTH -Q
   ```

7. Configure ASR, graceful server shutdown, and pager number using the Compaq System Configuration Utility or Compaq Insight Manager, as explained in the following sections.

8. To activate the Server Health driver enter the same load command you added to AUTOEXEC.NCF, or restart the server.

Configuring ASR

By default, the ASR option is disabled. To enable ASR, use Compaq Insight Manager or the Compaq System Configuration Utility to enable the Software Error Recovery feature. ASR can be configured to reboot the operating system or to boot the Compaq Utilities. Booting the Compaq Utilities allows remote asynchronous (modem) server management or network based server management (refer to your server documentation for supported remote maintenance features). For environmental recovery features (thermal conditions and power loss), the server will always reboot the operating system.

Configure the Software Error Recovery Timeout value using the Compaq System Configuration Utility. The Software Error Recovery Timeout value determines how long (in minutes) ASR will wait before it determines a server failure has occurred and the server is rebooted. The default is 10 minutes.

For NetWare 3.12 and above, add the following statement to STARTUP.NCF to allow the server to complete the boot process after an ASR reset:

```
SET AUTO TTS BLACKOUT FLAG = ON
```

Pager Number

Use Compaq Insight Manager or the Compaq System Configuration Utility to configure your server to send an alert to a defined pager number when it is rebooted after any critical event, such as a server failure or shutdown. You must have a properly configured modem connected to the server to use this option.

Configuring Graceful Shutdown

In general, you will want to configure the Server Health driver to gracefully shutdown your server to prevent overheating or in the event of power loss. Graceful shutdown is the default for all critical events, except for those events that do not allow it.

Use Compaq Insight Manager or the Compaq System Configuration Utility to configure graceful shutdown. Enable the Thermal Shutdown option to gracefully down the server to prevent critical overheating. To gracefully down the server in the event of a power loss, enable the UPS Shutdown option. (**Note**: You must have a Compaq UPS and appropriate drivers loaded to enable UPS Shutdown.)

Graceful shutdown is not configurable in the following cases:

- Failure of a processor fan in servers with multiple processors, such as the ProLiant 2000, ProLiant 4000, or ProLiant 4500, always causes an immediate graceful server shutdown, even if Thermal Protection is disabled. The Thermal Protection option will still control shutdown for failure of other system fans.
- Complete graceful shutdown may not be possible if the UPS battery is not charged or if the UPS is not given enough time to shutdown the server.

VIEWING THE ERROR LOGS

There are two server health logs: the critical error log and the corrected memory error log. Server failures, hostile environmental conditions, and miscellaneous hardware failures are logged in the critical error log. Corrected memory errors, which do not halt server functions, are logged in the corrected error log. To view the error logs when the server is active, use Compaq Insight Manager. To view the error logs when the server is down, use the Compaq Inspect Utility (part of the Compaq Diagnostics Utilities) or Quick Diagnostics (press F9 during system reboot).

You should view the error logs as part of regular maintenance and whenever you suspect that an error may have occurred. Use Compaq Insight Manager to confirm if a critical error has occurred. When you view server information, the Recovery Icon in Insight Manager will show a degraded condition (yellow border) if a critical error has occurred. View the critical error log and the ASR status windows to view the type of error that has occurred.

The critical error log contains information on the most recent errors. Each error message includes a date/time stamp, a description of the failure, and the source of the error. Hardware error entries identify the faulty hardware component. Software error entries identify the running process, its version number, and offset at the time of the error in the "Last Failure Message."

Note: The extra information for software errors is stored for only the most recent failure. Be sure to copy down the information for future reference.

The corrected memory error log also contains information on the most recent corrected memory errors. Each error message contains a date/time stamp and the location of the memory error to allow identification of the memory module which caused the error.

ERROR LOG MESSAGES

The following table lists error messages found in the critical error log, a brief description of the error, and the corresponding action or actions you should take to correct the problem.

Log Message	Action to Take
Abnormal Program Termination (ABEND): The operating system has encountered an abnormal situation which has caused a system failure.	Check ABEND message to determine cause. Update any software or device drivers related to the ABEND. Check for OS patches or fixes related to the ABEND.
AC Failure - Manual UPS Shutdown: The server was manually shut down while the server was running on UPS battery power.	Check commercial power. Check power connection between the UPS and the wall outlet.
AC Line Failure: The server is running on UPS battery power due to a commercial power failure.	Check commercial power. Check power connection between the UPS and the wall outlet.
ASR - Reset Occurred: System ROM detected an ASR system reset	Look for ABEND messages or other Critical Errors.
ASR reset limit reached: The maximum number of system resets due to the ASR timer expiration has been reached, resulting in the discontinuation of boot-up.	ASR has reset the system multiple times. Look for ABEND message or repeated Critical Errors.
ASR Test Event: The System Configuration Utility generated a test alert.	No action required.
ASR Timeout NMI: The operating system received notice of an impending Automatic Server Recovery timer expiration.	Use information in the Critical Error Log to identify which software module locked up the system. Upgrade any applicable software. Apply related OS patches. Contact the software manufacturer and give the name, version, and code offset recorded in the Critical Error Log. If the software is current, the fault may reside in hardware controlled by the specified software. If the module is a device driver, test or replace the hardware related to the specified software module.

Battery Failing: Failure of the battery supporting the system's non-volatile RAM is imminent.	Replace the system non-volatile RAM battery within 7 days.
Diagnostic Test Error: Compaq diagnostic test routines detected a failure.	Refer to the ERROR MESSAGE section of your documentation for further information on this diagnostics error.
Error detected on boot-up (Error #): The system detected an error during the Power-On Self-Test.	Refer to the server's documentation for listings of POST error codes.
Fan Failure (Fan #): A system fan has failed.	Replace specified fan in the system, even if the specified fan appears functional (spinning). If the specified fan is still spinning, you can temporarily disable Thermal Protection using the System Configuration Utility until the fan can be replaced. This will allow the system to boot the OS.
Critical Temperature (Zone #):	Ensure that all system fans are functional. Ensure airflow to all system vents is not obstructed. Check room temperature, make sure air conditioning is not turned off at night.
NMI - CPU local error (CPU #): The CPU subsystem detected a fatal error.	Contact your service provider. The specified CPU chip or CPU board requires replacement.
NMI - Expansion board error: A board on the expansion bus indicated an error condition, resulting in a system failure. **NMI - Expansion Bus Slave Timeout:** A board on the expansion bus delayed a bus cycle beyond the maximum time, resulting in a system failure.	Run Diagnostics on the expansion board specified. Contact your service provider. The specified expansion board requires replacement.
NMI - Expansion bus arbitration error: Memory refresh cycles were delayed, potentially leading to data loss. The error results in a system failure.	Run Diagnostics on the system board and any expansion boards in the system. Contact your service provider.
NMI - Failsafe timer expiration (CPU #) **NMI - Software generated interrupt (CPU #):** Software was unable to reset the system fail-safe timer, or software indicated a system error, resulting in a system failure.	Use ABEND information to determine what software was running. Apply any OS patches. Upgrade any device drivers or other server software.

NMI - Processor cache parity error (CPU #): A data error occurred within the processor cache, resulting in a system failure.	If possible, disable the cache on the specified processor. Replace cache module, processor, or processor board.
NMI - Processor internal error # (CPU #) **NMI - Processor address error # (CPU #)** **NMI - Processor parity error (CPU #)** The processor detected an internal parity error.	Contact your service provider. The specified processor requires replacement.
NMI - Uncorrectable memory error: A data error occurred in the system memory which the system was unable to correct, resulting in a system failure.	Replace memory module specified. If a specific module is not listed, replace the entire bank of memory.
NMI -Expansion bus master timeout (Slot #): A bus master type expansion board in the indicated slot refused to release the bus after its maximum time, resulting in a system failure.	Run Diagnostics on the specified EISA board Contract your service provider. The EISA board specified requires replacement.
PCI Bus Parity Error (Bus #, Device #): A parity error was detected on the PCI bus.	Run Diagnostics on the PCI expansion board specified (may be an embedded device, such as a SCSI chip.) Contact your service provider. The PCI board specified requires replacement. If problem persists, check other boards on the same PCI bus.
Processor exception (CPU #, Exception Nr.#): The indicated processor exception occurred.	Check ABEND information to determine the source of the exception. Under NetWare, processor exceptions are generally caused by software faults. Upgrade any related software or contact your software vendor.
Processor failure (CPU #): The processor in the indicated slot failed the power-on diagnostic, resulting in a processor shutdown.	Replace specified processor or processor board.
Redundant Power Supply Failure: One of the redundant power supplies failed.	Replace specified power supply.
Server Manager failure (Error #): An error occurred in the system interface with the COMPAQ 32-Bit Server Manager/R board.	Refer to Server Manager/R documentation for error code. Run Diagnostics on the Server Manager/R board.

UPS Auto Shutdown - Battery timeout: The system was shut down automatically after commercial power failed and the user-specified timeout had expired.	Check commercial power. Check power connection between the UPS and the wall outlet.
UPS Shutdown - Battery Depleted: The system was shut down automatically after commercial power failed and the UPS battery power was low.	Check commercial power. Check power connection between the UPS and the wall outlet. Ensure that the UPS battery is good.

SUMMARY

Loading the Server Health driver can shorten unexpected down time, identify the cause of a server failure, help preserve hardware components, and detect a need for preventative maintenance. The ability of the Server Health driver to identify the hardware component causing an NMI or the process running at the time of a server failure makes diagnosing the cause of an ABEND much easier. The ASR feature of the Server Health driver can also improve server availability by detecting server failures that NetWare cannot detect. It can also automatically reboot the server after a failure without user intervention.