Improving Fault Tolerance

Using Memory Redundancy and Hot-Plug Actions in Dell PowerEdge Servers

Features that enable redundancy across physical memory can enhance server reliability and help keep critical business applications available 24/7—particularly when combined with hot-plug capabilities designed into the Dell™ PowerEdge™ 6850 server. Allowing administrators to replace failing memory devices and add incremental memory upgrades while a server is running can help reduce system downtime and enhance scalability considerably. This article discusses memory redundancy and hot-plug features of PowerEdge 6850 servers.

RAID technology helps provide redundancy, fault tolerance, and high availability in enterprise disk drive subsystems. Different RAID levels, such as RAID-0, RAID-1, RAID-5, and RAID-10, can be configured to enable secondary storage benefits that suit specific long-term organizational requirements. Currently, the low cost of physical memory allows servers to support 32 GB to 64 GB of server RAM cost-effectively in dual in-line memory modules (DIMMs).

Unfortunately, the requirement for large memory capacities can increase the chance of memory errors simply because physical memory devices have the potential for failure. Thus, the more memory that resides in a system, the greater the potential for memory failure in that system. Safeguarding against potential memory failures and helping to ensure uninterrupted application availability, fault tolerance, redundancy, and hot-plug capabilities can be crucial for IT environments.

Hard and soft memory errors
Memory is an electronic storage device, and electronic storage devices have the potential to incorrectly return information—that is, data read from memory can be different from what was originally written to memory. Dynamic RAM (DRAM) stores 1s and 0s as charges on small capacitors residing on the DIMM, which must be continually refreshed to help ensure that the data is not lost. But even a small electrical disturbance near the memory cell can alter the charge in a capacitor, thus causing a memory error.

Typically, two types of error occur in a memory system. The first is called a repeatable, or hard, error.
Hard errors consistently return incorrect results and usually indicate that a piece of hardware is broken. For example, a bit may be stuck such that it always returns a 0 regardless of whether a 1 or a 0 is written to it. Hard errors are relatively easy to diagnose and fix because they are consistent and repeatable.

The second kind of error is called a transient, or soft, error. Soft errors occur when a bit reads back the wrong value once, but subsequently functions correctly. Soft errors are more common than hard errors—and are also more difficult to diagnose. They are not caused by circuit problems and, once corrected, do not reoccur.

**Memory error detection and correction**

Reliability in memory starts at the DIMM level. To help ensure a reliable memory system in servers, it is essential to provide protection from both hard and soft memory errors. Memory detection or correction protocols such as parity checking or error-correcting code (ECC) are designed to provide true protection from hard and soft memory errors.

Parity checking is one of the oldest and most basic forms of memory checking. It is a simple method of detecting single-bit errors in a memory system. Along with the eight bits of data stored in memory, parity checking uses one additional bit to determine the parity of the byte—odd or even. However, parity checking detects only odd-numbered single-bit errors, and does not enable administrators to locate and correct these errors. When a parity error is detected, the parity circuit generates what is called a nonmaskable interrupt (NMI), which is generally used to instruct the processor to halt immediately. The processor is halted to help ensure that the incorrect memory does not corrupt other data on the system.

ECC, on the other hand, not only detects both single-bit and multibit errors, but it also corrects single-bit errors. Each time data is stored in memory, ECC memory uses an algorithm to add a block of bits known as check bits. When this data is retrieved, the sum of the check bits (the checksum) is recomputed. The checksum of the written data is then compared with the checksum of the read data to determine whether any of the data bits are corrupted. If the checksums are identical, this indicates to the ECC memory that there is no error. If they are different, the data contains one or more errors. The ECC memory logic then isolates the errors and reports them to the system.

For a correctable single-bit error, the ECC memory logic corrects the error and outputs the corrected data without halting the system. For multibit errors—that is, errors involving two or more bits—ECC memory is capable of detecting but not correcting the errors. When multibit errors occur, ECC handles them by generating an NMI that instructs the system to halt.

As memory capacity increases, the number of soft errors will rise. Typically, a percentage of soft errors are multibit errors, which ECC cannot correct. As a result, administrators should expect the potential for failure in ECC systems to increase as memory is increased.

**Memory redundancy options in PowerEdge 6850 servers**

To help provide server fault tolerance, maximize memory capacity, and enhance reliability, Dell server hardware is designed with memory redundancy options that can help improve the performance and uptime of servers in memory error situations. The PowerEdge 6850 memory subsystem resides on up to four memory riser boards, or cards, supporting a system maximum of 64 GB when 4 GB DIMMs are installed. Each riser has four double data rate 2 (DDR2) slots arranged logically as two banks and one memory bridge (see Figure 1).

The PowerEdge 6850 offers three memory redundancy options, which are set in the BIOS: spare-bank memory, memory mirroring, and memory RAID. Organizations select these options when ordering the PowerEdge 6850. They may also opt not to use these three options. To disable these options, organizations should select the Redundancy Disabled option. Figure 2 shows BIOS options for memory redundancy modes.

**Spare-bank memory**

Within a riser board, memory can be set as a spare bank. Once sparing is enabled, when the error rate—that is, the rate of correctable errors—of a failing DIMM reaches a preset threshold (set by the administrator in the BIOS), the DIMM’s contents are copied to the spare bank. When the copy is in progress, all reads from the
SYSTEM ARCHITECTURE

Memory space mapped to the failing DIMM will be serviced by the spare bank, while all writes will be made to both the failing DIMM and the spare bank. Once the copy is complete, the failing DIMM is disabled and the spare bank services requests to that memory space. The BIOS logs an event in the system event log (SEL) indicating that a spare-bank switch occurred.

The sparing process is transparent to the OS. When sparing is enabled, only three-quarters of the total installed memory is available to the OS. Uncorrectable errors will not generate sparing failover events; instead, such errors will cause the system to halt.

Memory mirroring
Memory on one riser board can be mirrored to memory on another riser board if total memory is identical on both. Mirrored memory consists of redundant copies of the system memory. If a multibit error is detected when memory is accessed, the system will not crash because uncorrectable errors are considered to be correctable when memory mirroring is enabled. In that case, the chipset will automatically try to regenerate the data from the redundant mirrored copy of the memory, which resides on the redundant system memory. (If the retry limit exceeds a certain threshold, redundancy will be lost.) The BIOS logs an event in the SEL indicating whether memory mirroring is enabled and another event if redundancy is lost.

Memory mirroring can provide a high level of fault tolerance. The minimum configuration required for mirroring is two riser boards and two identical DIMMs on each board. Mirroring is transparent to the OS. When mirroring is enabled, only half of the total installed memory is available to the OS.

Memory RAID
The Dell PowerEdge 6850 disk drive subsystem can be configured as RAID-5 in server memory, in much the same way that RAID-5 is implemented in storage devices. Implementing RAID-5 requires that all four risers be present in the system and populated with equal amounts of memory. RAID is transparent to the OS, and when enabled, less than half of the total installed memory is available to the OS. Usable memory will be three times the total effective memory present in each riser. Behavior of correctable and uncorrectable errors for RAID is the same as it is for memory mirroring. In addition to correctable and uncorrectable memory SEL events, the BIOS logs an event in the SEL when memory RAID is disabled and logs another event in the SEL when memory RAID is enabled.

The Dell OpenManage™ systems management suite presents administrators with a graphical view of the system components. Figure 3 shows the memory redundancy options on a PowerEdge 6850.

Hot-plug functionality in PowerEdge 6850 servers
Dell PowerEdge 6850 servers offer two hot-plug functions designed to improve the uptime of servers: hot replace and hot add.

Hot-replace functionality
Hot-replace functionality refers to the capability to replace a failed DIMM on a removed riser board while the system continues to operate. PowerEdge 6850 servers provide hot-replace capability without requiring OS support. In fact, the hot-replace process is transparent to the OS.

A riser board can be hot replaced—which entails removing the board from the system, replacing failed DIMMs, and reinserting the board—only if a memory mirroring or memory RAID configuration is enabled as described in the preceding “Memory mirroring” and
“Memory RAID” sections in this article. Note: It is not necessary for DIMMs to fail before the riser board can be hot replaced; administrators can hot replace riser boards as needed.

When the chipset detects a DIMM failure on memory riser 1, all memory accesses are routed to the healthy memory riser 2 of the mirrored set. The system automatically turns off the green power LED and activates the amber attention LED on riser 1, and then initiates the hot-removal event. Prompted by the amber attention LED, an administrator pushes the attention switch, which switches on the blinking green power LED and switches off the amber attention LED. After BIOS completes setting status registers, the system turns off power to riser 1 (indicated by the green power LED switching off). At this point, the administrator can safely remove riser 1.

Hot insertion to a failed mirror or RAID set is initiated when an administrator inserts a riser board into the slot from which the board was previously removed. After seating the riser board, the administrator pushes the attention switch. The green power LED starts blinking and the chipset begins its recovery initialization. BIOS then performs memory initialization. Upon successful completion, re-silvering begins. Once this process is complete, risers 1 and 2 are once again redundant. The solid green power LED remains on and the mirror/RAID LED turns on.

The BIOS logs a memory removal event in the SEL when an administrator successfully removes a riser board, accompanied by a mirror or RAID redundancy lost SEL event (memory mirror redundancy requires two risers; memory RAID requires four risers). When an administrator successfully replaces a riser board, BIOS generates a SEL event indicating that memory mirroring or memory RAID redundancy has been regained as well as a memory-add event (see Figure 4). BIOS also generates a memory configuration error event if the current memory configuration does not support either adding or removing the riser board.

Hot-add functionality
Hot-add functionality allows administrators to increase memory size dynamically by adding a riser board to open slots. Note: A memory riser board should not be hot removed to add more memory. Only a previously empty slot may be used, and configuration rules must be followed.

A hot-add event is initiated when an administrator inserts a riser board into a previously empty slot. After seating the board, the administrator pushes the attention switch, the green power LED starts blinking, and the chipset starts the initialization process. BIOS then performs memory initialization, and the solid green power LED turns on.

Unlike hot-replace functionality, hot-add functionality requires the support of the OS. Upon successful completion of the hot add, control is transferred to the OS, which allocates memory and makes required changes to system properties. If an initialization step fails, power to the slot is turned off and all LEDs are turned off. Currently, only Microsoft® Windows Server™ 2003, Enterprise Edition, and Windows Server 2003, Datacenter Edition, support hot addition of system memory.

BIOS generates a memory-add event when an administrator successfully adds a memory card.

Notifying the OS about hot addition of memory
To notify the OS that the system is capable of supporting the hot addition of memory,1 Dell BIOS provides a static resource affinity table (SRAT).2 The SRAT also indicates to the OS the memory range that can potentially be hot added in the system.

To report added memory to the OS when an administrator adds a riser board, the BIOS instructs the chipset to generate a system control interrupt (SCI) after the added memory is

---

2 For more information about the SRAT, see the ACPI 2.0 specification at www.acpi.info.
SMBIOS provides an interface for motherboard and system vendors to present management information about their products in a standard format on Intel architecture systems.

The SMBIOS specification is intended to provide enough information for developers of management instrumentation to create generic routines that translate from SMBIOS format to the format used by their chosen management method.

SMBIOS contains information about total memory in the system and the DIMMs present in the system. When a hot-plug operation is initiated, fields in the SMBIOS table need to be updated at the end of the operation—otherwise, the management software using SMBIOS information will contain outdated information. BIOS POST code updates the SMBIOS table to ensure that the management software using SMBIOS has current information.

Synchronizing MSRs and MTRRs

Memory Type Range Register (MTRR) is a processor feature that allows the processor to optimize memory operations for different types of memory, such as RAM, ROM, and memory-mapped I/O. MTRRs configure an internal map of how physical-address ranges are mapped to various types of memory.

MTRRs and Model Specific Registers (MSRs) should be synchronized among all the processors in a system. Anytime MTRRs or MSRs are updated for any processor, all the processors in the system should be in sync; otherwise, the system may hang. Because the amount of memory in the system increases during hot-add operations, MTRRs should be updated so that the processor is aware that memory was added. At the end of a hot-plug operation, the BIOS updates the MTRRs and all of the processors’ MSRs and MTRRs will be synchronized.

The drive for IT reliability

To enhance reliability and enable 24/7 operations for servers supporting business-critical applications, administrators can combine server redundancy features—such as spare-bank memory, memory mirroring, and memory RAID—with hot-plug capabilities such as those available in the Dell PowerEdge 6850 server. As IT organizations are called upon to provide the memory fault tolerance necessary to minimize downtime, features that enhance redundancy and hot-plug memory capabilities are expected to become increasingly important tools in the administrative arsenal.

Chandra S. Mugunda is a senior development engineer in the Dell Instrumentation Software Group. Chandra has an M.S. in Computer Science from the Indian Institute of Technology, Roorkee, and a B.S. in Electrical Engineering from Andhra University in India.

Vijay Nijhawan is senior consulting engineer in the Dell Server BIOS Group. He is currently pursuing his M.S. in Engineering Management from The University of Texas at Austin. He has a bachelor’s degree in Computer Science from MD University in India.

Dennis Stultz is the lead BIOS engineer in the Dell Server BIOS Group for enterprise systems. He has a B.S. in Electrical Engineering from Mississippi State University.

Saurabh Gupta is a BIOS engineer in the Dell Server BIOS Group. He has an M.S. in Computer Engineering from Texas A&M University and a B.S. in Electrical Engineering from Gujarat University in India.

Harish Jayakumar is a test engineer in the Dell OpenManage software development and test organization. He has an M.S. in Computer Science from Arizona State University and a B.S. in Computer Science from the University of Madras in India.

FOR MORE INFORMATION

ACPI:
www.acpi.info

SMBIOS:
www.dmtf.org/standards/smbios