Whitepaper

Memory Errors and Dell EMC PowerEdge YX4X Server Memory RAS Features

Revision: **1.3**
Issue Date: **11/20/2020**

Introduction

Memory sub-system errors are some of the most common types of errors seen on modern computing systems. Understanding how memory errors occur and how to prevent or avoid them can be a complex subject – one that has challenged countless numbers of industry researchers and developers over the last 30 years. While Dell EMC PowerEdge servers are designed to provide industry leading Reliability, Availability, and Serviceability (RAS) on memory issues, we realize that many of our technically savvy customers may want to know more on what's happening 'under the hood' of their servers. This technical whitepaper is divided in four sections to help PowerEdge users to understand about the following memory error topics:

- Types of memory errors and how they may affect a server
- Dell EMC PowerEdge YX4X server memory RAS capabilities
- Configuring a PowerEdge server to achieve maximum memory up-time
- Recommended user actions when encountering memory errors

Important: The content covered in this whitepaper applies to Dell EMC PowerEdge YX4X servers with Intel Xeon SP processors. Customers with YX4X servers that utilize AMD EPYC or Intel Xeon E processors should refer to v1.0 of the RAS whitepaper.

The features described in this document assume the user is running the latest versions of Dell EMC PowerEdge server firmware, such as BIOS and iDRAC.

Revisions

Date	Description
January 3, 2020	• Initial release
July 10, 2020	• Removed content for platforms based on AMD EPYC and Xeon E processors • Added more information to primer on uncorrectable errors • Added clarification on PPR resources for genuine Dell DIMMs • Added MEM8000 SEL event to recommended user actions list • Added clarification to MEM9072 SEL event details and recommended user action • Added content specific to updates contained in BIOS 2.7.7
September 9, 2020	• Deprecated references to BIOS 2.7.7 and replaced with latest version, BIOS 2.8.2 • Clarifications to MCA Recovery and Memory Patrol Scrub sections • Clarifications to recommended actions after MEM0001 and MEM9072 SEL events • Clarifications and additional information on Fault Resilient Memory
November 20, 2020	• Added content specific to updates contained in BIOS 2.9.4 • Added VxRail platforms to list of applicable YX4X servers

Author

Name	Role
Jordan Chin 	Memory Technologist, Distinguished Member Technical Staff, Dell EMC

Acknowledgements

This paper had contributions from the following people:

Name	Role
Stuart Berke	CPU and Memory Technologist, VP, Fellow, Dell EMC
David Chalfant	BIOS Development, Technical Staff, Dell EMC

Huong Nguyen	BIOS Development, Technical Staff, Dell EMC
Ching-Lung Chao	BIOS Development, Technical Staff, Dell EMC
Fred Spreeuwers	IPS Engineering, Technical Staff, Dell EMC
Mark Dykstra	IPS Engineering, Senior Principal Engineer, Dell EMC
Rene Franco	Memory Systems Engineering, Senior Manager, Dell EMC
Mark Farley	Component Quality Engineering, Senior Principal Engineer, Dell EMC

A Primer on Memory Errors

To fully understand the memory RAS response capabilities of PowerEdge servers, it is first helpful to understand the various types of possible memory errors.

DRAM issues can be broadly classified into two categories described below:

- **Soft Errors**
  - Soft errors are transient in nature and may often be caused by electrical disturbances in the memory sub-system components. These disturbances could occur in any one of many locations within the memory subsystem including the processor memory controller, processor-internal buses, processor cache, processor socket or connector, motherboard bus traces, discrete memory buffer chips (if present), DIMM connectors, or individual DRAM components on DIMMs.
  - Soft errors may be caused by phenomena such as high-energy particle strikes in the memory subsystem or electrical noise in the circuits. Single or multiple bits can be affected, with single-bit errors and some multi-bit errors corrected using demand or patrol scrubbing.
- **Hard Errors**
  - Hard errors are persistent in nature and cannot be resolved over time, through system resets, or through system power-cycles. These types of errors could occur as a result of stuck-at faults (i.e. degradation of a single lane on a bus or a single memory cell in a DRAM component), due to failure of an entire device (for example connector, processor, memory buffer, or DRAM components), due to improper bus initialization, or memory power issues. Failures within a DRAM component may consist of entire device failure, bank region failure within a device, pin failure, column, or cell failure.
  - Hard errors may be caused by physical part damage, Electrostatic Discharge (ESD), electrical overcurrent conditions, over temperature conditions, or irregularities in processor or DRAM fabrication or module assembly.

The two categories of DRAM errors previously described can ultimately lead to two types of memory errors:

- **Correctable Errors (CEs)**
  - Correctable errors are errors that can be detected and corrected by the server platform. These are typically single-bit errors, though based on CPU and memory configuration, may also be some types of multi-bit errors (corrected by Advanced ECC). Correctable

errors can be caused by both soft and hard errors and will not disrupt operation of PowerEdge servers.
    - As DRAM based memory shrinks in geometry to grow in capacity, an increasing number of correctable errors are expected to occur as a natural part of uniform scaling. Additionally, due to various other DRAM scaling factors (e.g. decreasing cell capacitance) there is an expected increase in the number of error generating phenomenon such as Variable Retention Time (VRT) [1] and Random Telegraph Noise (RTN) [2].
    - Within the server industry, it is an increasingly accepted understanding shared by Dell that some correctable errors per DIMM is unavoidable and does not inherently warrant a memory module replacement. However, some server competitors will go as far as to say that an indefinite number of correctable errors are acceptable – a belief that is not shared by Dell Engineering. Instead, PowerEdge server firmware will intelligently monitor the health of memory and recommend self-healing action or module replacement based on a variety of factors including DIMM capacity, rates of correctable errors, and effectiveness of available self-healing. The intent behind Dell's proprietary predictive failure algorithms is to proactively identify DIMMs that are most likely to continue to degrade and potentially generate uncorrectable errors.

- **Uncorrectable Errors (UCEs)**
    - Uncorrectable errors are multi-bit errors that could not be corrected by the server platform. These can be caused by any combination of soft or hard errors, but typically occur as a result of multiple hard errors.
    - Not all multi-bit errors are uncorrectable. CPUs that support Advanced ECC can correct some types of multi-bit errors, depending on the bit error pattern.
    - An uncorrectable error can be classified as being:
        - Detectable and consumed
        - Detectable and unconsumed
        - Silent and consumed
        - Silent and unconsumed
    - Consumed vs unconsumed refers to whether the data has been loaded into the CPU execution path. Unconsumed errors are typically found during Memory Patrol Scrub but may also be found during a CPU prefetch.
    - Detectable vs silent refers to whether the CPU's ECC scheme can detect the existence of the error. Silent errors are exceptionally rare and require the problematic cache line to meet a very specific bit error pattern to by-pass the CPU's ECC scheme.
        - Unless otherwise specified, references to uncorrectable errors in this whitepaper will refer to those classified as detectable.

	Detectable	Silent
Consumed	Poisoned upon detection then Machine Check Exception after consumption;	Depends on data usage. It can result in either incorrect application data or system service outage

	Outcome based on OS error containment	
Unconsumed	Poisoned upon detection; error waits to be consumed	Error waits to be consumed

A Primer on Dell EMC PowerEdge Server Memory RAS Capabilities

Previously discussed memory errors are mitigated through PowerEdge server memory RAS capabilities which entail fault avoidance, detection, and correction in hardware and software. These mitigating RAS features are all intended to improve system reliability and extend uptime in the event of memory errors.

> **FYI:** It is useful to understand the difference between x4 and x8 DIMMs. This refers to the width of the DRAM components on a memory module. x4 DIMMs utilize DRAM components that have a 4-bit data width and x8 DIMMs utilize components with an 8-bit data width.
>
> The common DIMM organizational notation is as follows: #RxN. Where # is the number of ranks and N is the width of the DRAM. Example – 2Rx4 means the DIMM has two ranks of x4 DRAM devices.

Single Error Correction - Double Error Detection (SEC-DED) ECC

SEC-DED Feature Support Table	
DIMMs Supported	x4 DIMMs: ☑
	x8 DIMMs: ☑

Single Error Correction - Double Error Detection ECC, or SEC-DED ECC, is the most basic form of error correcting code (ECC) available. All PowerEdge servers configured with ECC memory modules are capable of SEC-DED for each memory page access (64 data bits + 8 ECC bits). This means that any one bit among the 72-bits accessed from DRAM can be incorrect and PowerEdge server hardware will automatically correct it – regardless of cause.

Advanced ECC

Advanced ECC Feature Support Table	
DIMMs Supported	x4 DIMMs: ☑ (Use of x4 DIMMs May Provide DRAM Device Correction)
	x8 DIMMs: ☑ (Use of x8 DIMMs May Provide Nibble Correction)

Advanced ECC is a RAS feature that provides error correction on single-bit and multi-bit failures that are bound within 4-bits (nibble) of memory accesses. When used in conjunction with DIMMs based on x4 DRAM devices, Advanced ECC may provide error correction to an entire single DRAM device. This type of error correction that covers an entire DRAM device has been branded in various forms, most popularized as *Chipkill* and *Single Device Data Correction* (SDDC). Advanced ECC is a highly complex feature that is based on the concept of *Single Symbol Correcting – Double Symbol Detecting* (SSC-DSD) Reed-Solomon error correcting and detection code [3]. At a high level, SSC-DSD works by breaking up cache line accesses into 'code words' which in turn are made up of multi-bit symbols. The size of these symbols can vary depending upon the processor architecture. But regardless if the symbol size is 4-bits or 32-bits, as the SSC-DSD name implies, the coding is designed such that a single symbol may be corrected for various combinations of bit errors. In many cases, depending on the SSC-DSD implementation, all bits in a symbol could be corrected if they had errors. Studies have indicated that error correcting codes based on SSC-DSD may provide up to 42x better fault correction and avoidance than SEC-DED ECC alone [4].

An example SSC-DSD coding implementation is represented in the following figures where a 64-byte cache line is broken into four code words (Figure 1). Each code word is made up of eighteen 8-bit symbols and can be broken down into 128-bits of data and 16-bits of ECC (8-bits of CRC and 8-bits of parity). The data and ECC is arranged (as shown by the various colors) such that all bits from an entire symbol are located within a single x4 DRAM device. Depending on the SSC-DSD coding implementation, Advanced ECC may correct various combinations of multiple bits error patterns within a single symbol including the entire symbol itself (Figure 2). However, a pair of two single bit errors across two symbols would yield an uncorrectable error (Figure 3).

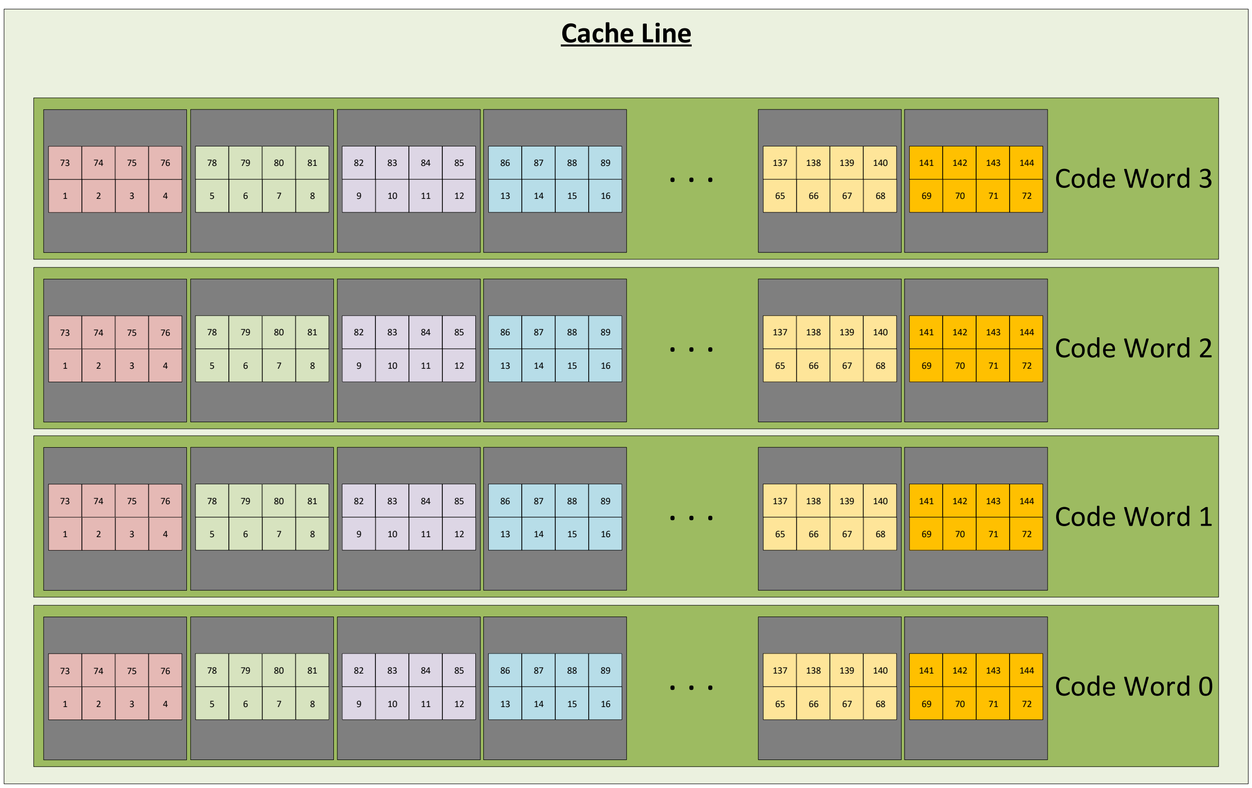

Figure 1 - Example Implementation of 8-Bit Symbol Size Advanced ECC

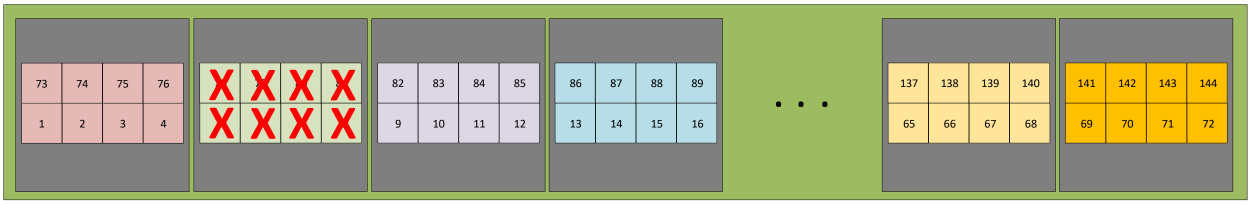

Figure 2 - Advanced ECC can correct multi-bit errors in a single symbol…

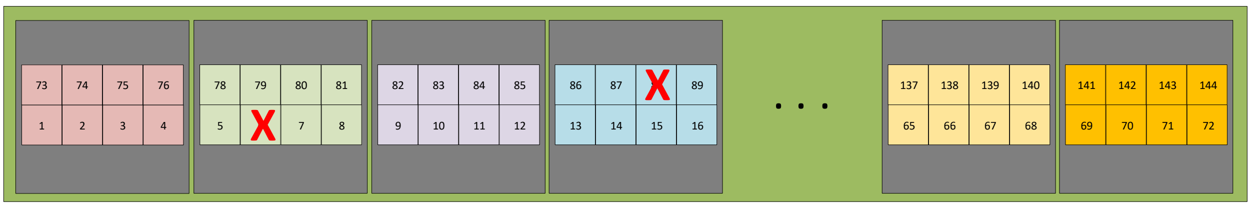

Figure 3 - But Advanced ECC cannot correct errors in multiple symbols

As described earlier, SSC-DSD implementations will vary based on CPU platform architecture and generation. This results in different error correction coverage and memory configuration requirements to enable Advanced ECC. The current SSC-DSD coding implementation in PowerEdge YX4X servers with Intel Xeon SP processors will provide data correction on most of the possible error patterns within a single symbol. In PowerEdge YX4X servers, Advanced ECC is now enabled by default as part of Independent Mode on all Intel Xeon SP based platforms.

It was earlier mentioned that Advanced ECC, when used in conjunction with x4 DIMMs, may provide error correction to a single DRAM device including failure of the entire device itself. This is achieved by the processor organizing memory accesses such that a given DRAM device is only contributing data to a single symbol – and through SSC-DSD, any one symbol can be fully redundant. On the other hand, DIMMs with x8 DRAM devices will straddle two symbols and may only provide partial device correction at the nibble level. An uncorrectable error will occur should both symbols within a x8 DRAM device experience an error or the entire device fail.

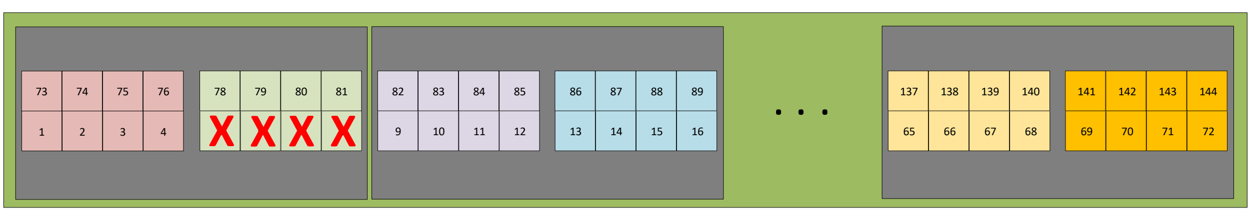

Figure 4 – Advanced ECC with x8 DIMMs allows a nibble per DRAM to be protected but not the entire device

Adaptive Double Device Data Correction (ADDDC)

ADDDC Feature Support Table	
DIMMs Supported	x4 DIMMs: ☑
	x8 DIMMs: ☒
Memory Configuration Required	• Two or more memory ranks per memory channel

Adaptive Double Device Data Correction (ADDDC) is an Intel platform-specific technology that allows for two DRAM devices to sequentially fail before loss of fault-avoidance. ADDDC is only supported with x4 DIMM populations and requires a memory configuration of two or more memory ranks channel (two DIMMs per channel or a single DIMM with multiple ranks).

ADDDC works by having the BIOS track the number of correctable errors per DRAM bank. If the number approaches a threshold deemed unsafe by BIOS, then ADDDC is activated and the failing DRAM bank is dynamically mapped out while a 'buddy' bank is mapped in to take its place. The DIMM continues to operate with SDDC coverage. At this point, memory performance will be impacted as the memory controller must do two reads for every read to the mapped-out cache-lines.

> **FYI:** ADDDC will only provide fault coverage for sequential DRAM failures over time. Two parallel DRAM failures within the same memory access still result in a service outage. Additionally, ADDDC only applies to correctable errors and only helps to protect an uncorrectable error from occurring by reducing the chance that correctable errors become uncorrectable.

Memory Patrol Scrub

Memory Patrol Scrub Feature Support Table	
DIMMs Supported	x4 DIMMs: ☑
	x8 DIMMs: ☑

Memory Patrol Scrub is a Dell memory RAS feature designed to decrease the probability of a user encountering a multi-bit error by removing the accumulation of soft errors in DRAM. This in turn reduces the chance of encountering an uncorrectable error (depending on other RAS capabilities enabled and where the multi-bit error occurs). Memory patrol scrub works by having the CPU memory controller periodically scan through DRAM and correct any correctable errors that it encounters.

In addition to scrubbing for correctable errors, patrol scrub can also detect latent uncorrectable errors in memory. These UCEs are referred to as unconsumed uncorrectable errors – or uncorrectable errors detected in a non-execution path of the CPU. Detection of these unconsumed UCEs are logged in the System Event Log as a critical event, MEM9072: "The system memory has faced uncorrectable multi-bit memory errors in the non-execution path of a memory device at the location <location>." If the server is running BIOS version 2.8.2 or higher, it is highly recommended to perform a cold reboot as soon as possible. This will prevent the system from consuming the UCE and allow the server BIOS to perform self-healing at the affected memory location.

Memory patrol scrubbing is enabled by default and configured to perform in the background every 24 hours. Memory patrol scrub can be disabled or set to run at an accelerated schedule (every four hours) in the BIOS setup under the power management menu. Memory patrol scrub may have an impact on system performance for some workloads while it is running.

> **FYI:** *Demand Scrub* occurs when the memory controller encounters a correctable error during a regular run-time read transaction and writes back corrected data. The usefulness of Patrol Scrub is highlighted in scenarios where memory access patterns are highly focused in some areas and thus the other areas are not getting the benefits of Demand Scrub.

Memory Page Retire (MPR)

Memory Page Retire Feature Support Table	
DIMMs Supported	x4 DIMMs: ☑ x8 DIMMs: ☑

Memory Page Retire (MPR) is a feature implemented by PowerEdge server BIOS that instructs operating systems to stop using memory page locations (4 KB in size) that BIOS has deemed as potentially unhealthy – essentially removing it from the operating system's memory pool. This feature is also known as Operating System Memory Page Retirement or Page Off-lining. BIOS makes the determination of a potentially unhealthy memory page based on a proprietary PowerEdge server algorithm that takes into account correctable error patterns and error rates at a given memory page location.

Studies into memory page retirement (aka off-lining) have found that MPR can reduce memory error rates by as much as 94% [5]. This feature is automatically enabled in BIOS and will be activated provided the operating system supports memory page retirement. Most modern operating systems support the capability of receiving such memory page retirement requests. To know more about compatibility of OSs with receiving such request, contact your OS vendor.

Memory Rank Sparing

Memory Rank Sparing Feature Support Table	
DIMMs Supported	x4 DIMMs: ☑ x8 DIMMs: ☑
Memory Configuration Required	• Single Rank Sparing – Two or more memory ranks per memory channel • Multi Rank Sparing – Three or more memory ranks per memory channel

Memory Rank Sparing is a memory RAS feature available on Intel platforms that will reserve one or more memory ranks per channel as spares for failover. When the PowerEdge server memory health monitor has determined that one of the ranks in a memory channel has degraded, it will trigger rank sparing failover. The failover process consists of checking the health of the spare rank(s) through patrol scrubbing then seamlessly copy the contents of the degraded rank to the spare rank(s). Memory rank sparing is disabled by default and can be enabled in BIOS setup if required.

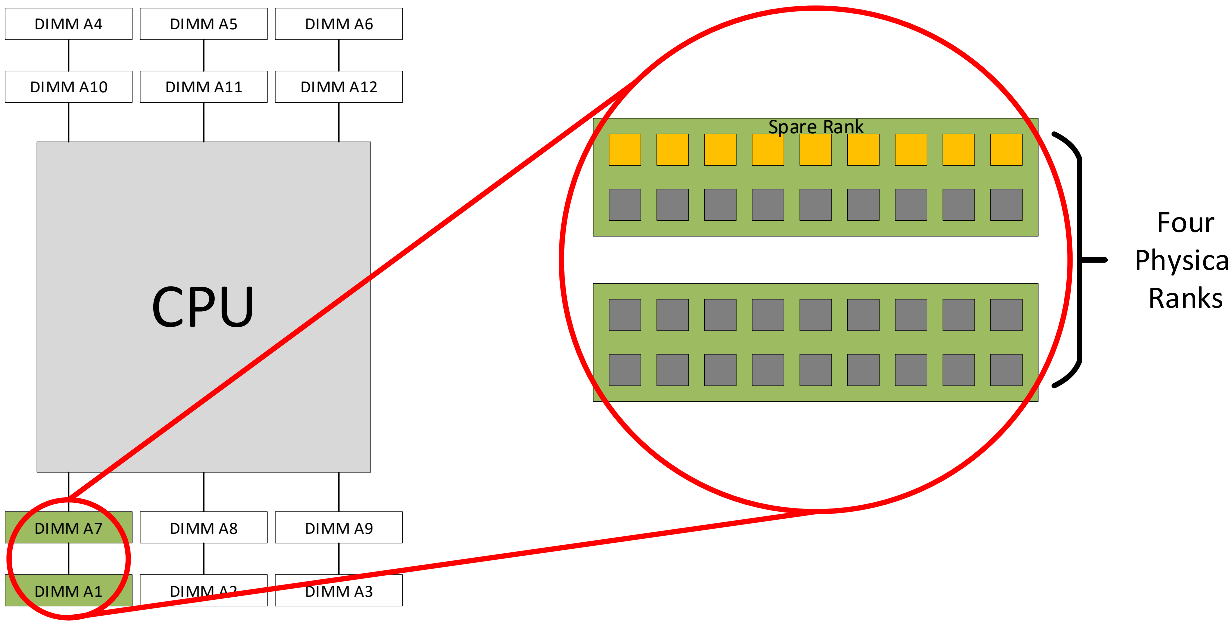

Figure 5 - Example of two 16GB (2Rx8) RDIMMs with one rank held as spare

In order to support single rank sparing, a system must be populated with at least two memory ranks per memory channel. The memory capacity reduction due to rank sparing is based on the memory configuration (number of ranks per channel and size of ranks). After one failover event has occurred, no further failover can occur in single rank sparing mode.

- Rank per channel capacity reduction impact (assuming same sized ranks) = 1 per number of ranks
    - E.g. 4 ranks = 25% reduction
- Largest rank in channel is always held as spare
    - E.g. One 32 GB RDIMM (2Rx4) and one 16 GB RDIMM (2Rx8) installed = two 16 GB ranks and two 8 GB ranks. One of the 16 GB ranks will be held as spare, resulting in a 33% capacity reduction.

In order to support multi rank sparing, a system must be populated with at least three memory ranks per memory channel. Like single rank sparing, the memory capacity reduction due to rank sparing is based on the memory configuration (number of ranks per channel and size of ranks). Up to two failover events may occur in multi rank sparing mode.

- Rank per channel capacity reduction impact (assuming same sized ranks) = 2 per number of ranks
    - E.g. 4 ranks = 50% reduction
- Largest rank in channel is always held as spare

- E.g. One 32 GB RDIMM (2Rx4) and one 16 GB RDIMM (2Rx8) installed = two 16 GB ranks and two 8 GB ranks. Both 16 GB ranks will be held as spares, resulting in a 66% capacity reduction.

Memory Mirroring

Memory Mirroring Feature Support Table	
DIMMs Supported	x4 DIMMs: ☑
	x8 DIMMs: ☑
Memory Configuration Required	• All identical DIMMs • Memory channels must be populated as either all one DIMM per channel or two DIMMs per channel

<u>Memory Mirroring</u> is a memory RAS feature available on Intel platforms that provides the highest level of protection against memory errors – including uncorrectable errors – at the cost of a 50% memory capacity reduction. This feature essentially works like RAID1 in the storage domain, where a redundant copy of memory contents is stored in a separate memory location. If an uncorrectable error is detected during a memory access at one location, then the mirrored content is retrieved instead. There is no impact to overall memory performance when full mirroring is enabled except during heavy write traffic.

In order to configure for memory mirroring, all memory modules must be identical in size, speed, density, and technology (RDIMM vs LRDIMM, etc). Additionally, memory channels must be populated with all one DIMM or all two DIMMs (for example, 24 DIMM systems should have 12 DIMMs or 24 DIMMs installed). Memory mirroring is disabled by default and must be enabled through the BIOS setup menu.

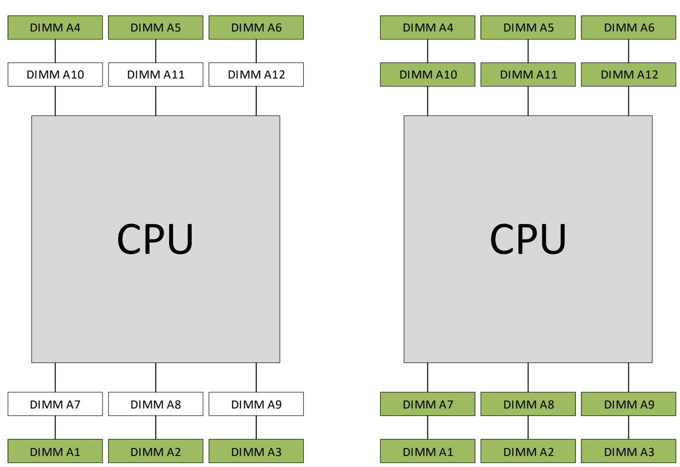

Figure 6 - 6 DIMM/CPU and 12 DIMM/CPU population rules using identical modules

> **Important:** Consult your PowerEdge server installation and service manual for complete memory population guidelines to properly enable Memory Mirroring.

Fault Resilient Memory (FRM)

Fault Resilient Memory Feature Support Table	
CPUs Supported	Xeon Platinum and Gold SP Families Only
DIMMs Supported	x4 DIMMs: ☑
	x8 DIMMs: ☑
Memory Configuration Required	• Memory channels must be populated as either all one DIMM per channel or two DIMMs per channel

> **FYI:** Dell has published a separate technical whitepaper specifically for Fault Resilient Memory.

Fault Resilient Memory (FRM) is a memory RAS feature that leverages partial memory mirroring to create a fault resilient memory region specifically for critical memory functions. This will prevent any uncorrectable errors in this memory region from generating a kernel panic or termination of virtual machines or applications. A comparative advantage of FRM versus full memory mirroring is that the memory capacity overhead can be configured based on need. The maximum memory redundancy overhead for FRM is 25% compared to 50% with full memory mirroring.

When used with hypervisors such as VMware vSphere ESXi version 5.5 or later, the hypervisor will ensure that only kernel memory and critical virtual machines utilize the mirrored memory region. In BIOS Setup, a user may select to enable this feature with 12.5% or 25% memory redundancy. Separately, the hypervisor must also be configured to utilize the mirrored memory region. Reference the Dell FRM technical whitepaper for additional guidance.

When used with Linux operating systems that support address range or partial memory mirroring, the OS will ensure that kernel memory utilizes the mirrored memory region. Since only the kernel will be using the redundant memory space, the redundancy footprint for this usage model can big significantly less than the hypervisor FRM configuration. In some cases, the redundancy footprint can be as little as 4GB of total memory. To configure this, a user must enable FRM in BIOS Setup selecting 12.5% or 25% memory redundancy. Then, the user must configure the operating system to request a new mirrored memory size. The OS request will override the static memory mirroring size specified in BIOS Setup. For more information about support for this feature, see the production documentation of the respective operating systems.

> **Important:** Use of FRM is only officially supported by Dell with VMware vSphere ESXi versions 5.5 or later (when used in conjunction with their Reliable Memory feature). Use of FRM with Linux is not officially supported by Dell.

Memory channels must be populated with all one DIMM or all two DIMMs (for example, 24 DIMM systems should have 12 DIMMs or 24 DIMMs installed). Fault Resilient Memory is disabled by default and must be enabled through the BIOS setup menu.

> **Important:** Consult your PowerEdge server installation and service manual for complete memory population guidelines to properly enable Fault Resilient Memory.

Memory Self-Healing

Memory Self-Healing Feature Support Table	
DIMMs Supported	x4 DIMMs: ☑
	x8 DIMMs: ☑

The Dell EMC PowerEdge server memory self-healing capability has two key components: post package repair and memory retraining.

Post Package Repair (PPR) is an industry-standard capability, defined by JEDEC, where a memory module is capable of swapping out degraded rows of memory with spare ones being held in reserve. While JEDEC requires that all DDR4 memory be built with at least one spare row per DRAM bank group, Dell requires all memory suppliers manufacture genuine Dell DIMMs with a significantly larger number of available spare rows. This is done to ensure that PowerEdge servers have a robust self-healing memory ecosystem.

> **FYI:** Genuine Dell DIMMs are any memory modules procured directly from Dell Technologies at time of system purchase, through aftermarket sales (e.g. dell.com), or through Dell services as a replacement.
>
> Any genuine Dell RDIMM or LRDIMM rated for speeds of 2666 MT/s or higher will support the extended PPR resource requirements (extra spare rows) by Dell Technologies.

When the server platform determines that a DRAM row has one or more faulty cells, it can instruct the DRAM to electrically swap out the old row and replace it with a new one. This happens through electrical fusing and is a permanent process. Additionally, the PPR process can only occur at the beginning of a boot process – before memory training and test can occur. Similar to Memory Page Retirement, deeming which DRAM require Post Package Repair is determined by a proprietary Dell algorithm that takes into account correctable error rates and error patterns. As of BIOS versions 2.8.2 or higher, self-healing through PPR on detection of uncorrectable errors is now also supported.

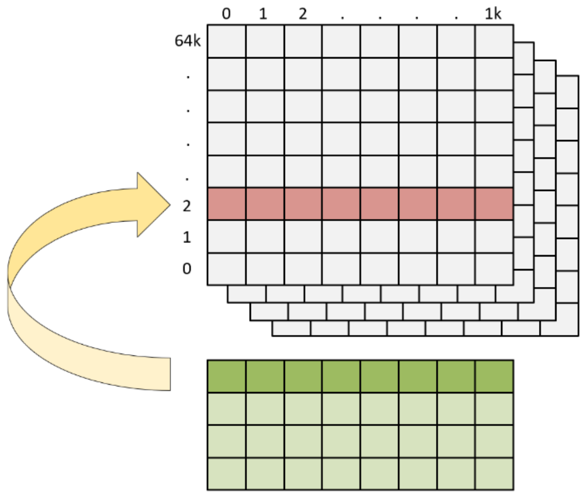

Figure 7 - PPR for a row in a bank group of a 4Gb x4 device

PPR is always available on PowerEdge server platforms that support it and if deemed necessary by BIOS will automatically execute after a system cold reboot. For PPR to successfully execute, it is recommended that users do not swap or replace DIMMs between boots when receiving memory error event messages, unless instructed to do so by Dell technical support personnel.

In addition to PPR, the PowerEdge server memory self-healing process also includes memory re-training. Memory training is the process by which the CPU initializes, calibrates, and tunes the link between itself and the memory modules. While performing full memory training can help to ensure that the memory bus operates at the highest level of signaling integrity, it is also a time-consuming process that directly impacts server boot times. Therefore, PowerEdge servers only perform this step when necessary, such as during the memory self-healing process.

Machine Check Architecture Recovery

MCA Recovery Feature Support Table	
DIMMs Supported	x4 DIMMs: ☑
	x8 DIMMs: ☑

Machine Check Architecture Recovery, or MCA Recovery, is an advanced RAS feature which when used in conjunction with operating systems that support it, can prevent some uncorrectable memory errors from generating an unexpected system-wide outage event. MCA Recovery is not a memory-specific RAS feature. Its capabilities extend to various forms of CPU data consumption including data from I/O. The scope of MCA Recovery discussed here will be limited to data consumption from system memory.

Essentially, MCA Recovery is a CPU capability that allows BIOS to signal consumption of an uncorrectable error to the operating system through a Machine Check Exception. This allows the OS an opportunity to potentially perform memory error containment. The outcome depends entirely on whether the impacted memory is associated with kernel space or user/application/VM space:

- If the impacted data was in kernel memory, then the OS will kernel panic.

- If the impacted data was in user/application/VM memory, then the OS will terminate the associated process or VM without impacting the rest of the system.
- If the impacted data was in user/application/VM memory but the OS had a redundant copy of the data, then the associated process or VM will recover.

Consult your operating system documentation on error containment for more information on OS behaviors.

Other Memory RAS Capabilities on PowerEdge servers

- Memory Map Out – If critical failures (such as uncorrectable errors) are detected in the memory training and test phase of POST, PowerEdge servers will automatically map out the affected DIMMs from the system memory pool. This prevents the faulty DIMM from incurring potential service outages. The affected DIMM will not be mapped back into the memory pool until there is a memory configuration change (such as a DIMM replacement).

Achieving Maximum Memory Up Time

Based on the memory RAS features discussed in the previous section, the following is a summary of how users can configure their systems to achieve maximum memory up time:

- Configure server using genuine Dell DIMMs
    - Benefit: Memory modules are fully validated and assured by Dell; additional self-healing (PPR) resources above and beyond industry standards
- Configure server with x4 DRAM based DIMMs
    - Benefit: Single DRAM Device Correction and ADDDC
- Configure server to operate in the following redundancy modes (in descending order of protection):
    - Best – Configure server to operate in Memory Mirroring Mode
        - Benefit: RAID1 level memory protection, significantly reduced probability of UCEs
        - Downside: 50% memory capacity reduction
    - Better – Configure server to operate in Fault Resilient Memory mode
        - Benefit: Significantly reduced probability of UCEs in critical portions of memory used by operating systems; low memory capacity reduction overhead (depending on the system settings)
        - Downside: Up to 25% memory capacity reduction, only officially supported VMware vSphere 5.5 or higher only
    - Good – Configure server to operate in Rank Sparing Mode
        - Benefit: Run-time elimination of memory ranks that are operating in a degraded state due to a large number of correctable errors
        - Downside: Varying amount of memory capacity reduction depending on memory configuration
- Configure server to run memory patrol scrub in ‘Extended Mode’

    - Benefit: Patrol scrub will run every four hours (instead of 24); increased frequency will reduce the accumulation of errors in areas of memory with low utilization and thus not being corrected by demand scrub

It is also recommended that users keep their PowerEdge server firmware up to date, especially server BIOS. This is because even after products ship, PowerEdge server development continuously works to improve its RAS algorithms and behaviors for an optimal customer experience. Users will also benefit from keeping BIOS up to date by receiving regular maintenance releases to their platform memory reference code.

FYI: *Memory Reference Code* (MRC) is a BIOS code that performs memory training, configuration, and link optimization.

As an example, since the version 1.0 publication of this whitepaper, several new features have been introduced in the latest versions of PowerEdge BIOS. See "What's New in BIOS 2.8.2" for additional details.

Recommended User Actions When Encountering Memory Errors

Reminder: The list of common memory errors and recommended response actions detailed below are for customers running PowerEdge BIOS versions 2.8.2 or higher. Customers with earlier BIOS versions should refer to v1.0 of the RAS whitepaper.

The following is a list of the most common memory errors (as reported in the system event log) and the recommended user response actions:

- **MEM0001** – This is an indication that the system has consumed an uncorrectable memory error at the specified DIMM location in the event message. Based on the OS error containment (MCA Recovery) process, the server may see one of three possible outcomes:
    1. Kernel panic
    2. Application or VM termination
    3. Application or VM recovery
    - **Recommended Response Action:** Perform a cold reboot of the server if it has not done so automatically. PowerEdge server BIOS will perform self-healing at the affected DIMM location (note that BIOS may initiate more reboots during this process). Do not remove or swap the DIMM at the specified location in the event message. Look for the system to report MEM0804 or MEM0805 for next steps. Repeat a cold reboot if neither of these events are reported. Contact Dell technical support if neither of these two events are reported after a second attempt.
- **MEM0802** – This is an indication that the system is encountering correctable errors at the specified DIMM location and would benefit from Dell memory self-healing.
    - **Recommended Response Action:** Perform a cold reboot of the server at the earliest convenience. PowerEdge server BIOS will perform self-healing at the affected DIMM

location (note that BIOS may initiate more reboots during this process). Do not remove or swap the DIMM at the specified location in the event message.

- **MEM0804** – This is an indication that the system has successfully performed memory-self healing at the specified DIMM location in the event message.
    - **Recommended Response Action:** No response required. DIMM is operating nominally.
- **MEM0805** – This is an indication that the system is unable to successfully perform memory self-healing at the specified DIMM location in the event message.
    - **Recommended Response Action:** Contact Dell technical support.
- **MEM8000** – This is an indication that the system is encountering correctable errors at a rate and pattern that may impact system performance. To preserve maximum performance, correctable error logging has been temporarily disabled at the specified DIMM location. See "What's New in BIOS 2.8.2" for additional background and improvements regarding this event.
    - **Recommended Response Action:** Reseat the DIMM identified by the SEL message. Do not swap or move the identified DIMM to a new location. The system may also attempt to initiate the self-healing process. If the problem persists, contact Dell technical support.
- **MEM9072 –** This is an indication that the system has detected an uncorrectable memory error in a non-execution code path (during Patrol Scrub or CPU prefetch). Without user action, the corrupted data can still be later consumed and may generate a service outage.
    - **Recommended Response:** Perform a cold reboot of the server as soon as possible. PowerEdge server BIOS will perform self-healing at the affected DIMM location (note that BIOS may initiate more reboots during this process). Do not remove or swap the DIMM at the specified location in the event message. Look for the system to report MEM0804 or MEM0805 for next steps. Contact Dell technical support if neither of these two events are reported.

Applicable Platforms

The following platforms are considered PowerEdge YX4X servers with Xeon SP processors and are therefore covered by this document:

> **Important:** Subsequent to the publication of this document, Dell may continue to add products to its YX4X server lineup. If a product is not listed below, please consult with a Dell sales or support representative to confirm the server generation.
>
> Additionally, PowerEdge leveraged products such as some Precision workstations, OEM, DSS, and ESI products may also be covered by this document. Please consult with a Dell sales or support representative to confirm.
>
> **Note:** The platforms listed below with an asterisks (*) will utilize a different BIOS version numbering schema than mentioned in this document. Please consult with a Dell sales or support representative to determine the appropriate comparable BIOS firmware version.

- PowerEdge T440*
- PowerEdge T640
- PowerEdge C4140
- PowerEdge C6420
- PowerEdge XR2*
- PowerEdge R440*
- PowerEdge R540*
- PowerEdge R640
- PowerEdge R740
- PowerEdge R740xd
- PowerEdge R740xd2
- PowerEdge R840
- PowerEdge R940
- PowerEdge R940xa
- PowerEdge FC640
- PowerEdge M640
- PowerEdge MX740c
- PowerEdge MX840c
- PowerEdge XE2420*
- PowerEdge XE7420*
- PowerEdge XE7440*

The following VxRail platforms are leveraged from PowerEdge YX4X servers with Xeon SP processors and are therefore are also covered by this document:

- VxRail E560
- VxRail E560F
- VxRail E560N
- VxRail P570
- VxRail P570F
- VxRail V570
- VxRail V570F
- VxRail S570
- VxRail G560
- VxRail G560F
- VxRail P580N
- VxRail D560*
- VxRail D560F*

PowerEdge servers with Xeon E and AMD EPYC processors are not covered in this whitepaper. Customers with these servers should continue to refer to [v1.0 of the RAS whitepaper].

What's New in BIOS 2.8.2

- **Self-Healing on Uncorrectable Errors** – Prior to this update, PowerEdge server BIOS was capable of performing self-healing only whenever its health monitoring algorithms deemed it necessary. With this PowerEdge server BIOS release, if the CPU detects an uncorrectable error, the server will automatically schedule self-healing to occur on the next cold reboot of the server. Look for MEM0804 (success) or MEM0805 (failure) in the System Event Log (SEL) for an indication of next steps.
- **Updates to Self-Healing Algorithm** – Prior to this update, PowerEdge server BIOS would rigorously test the target memory location prior to performing self-healing. This would ensure that self-healing is run only after 100% confirmation of a hard error. However, because of the abundance of self-healing resources found in genuine Dell DIMMs and the clear benefit self-healing has to customers, Dell modified the algorithm to extend self-healing to both 'confirmed' fail locations and 'questionable' fail locations.
- **MCA Recovery across all Xeon SP SKUs** – Prior to this update, only PowerEdge servers utilizing Advanced RAS Intel processors could perform MCA Recovery. With this BIOS update, PowerEdge servers utilizing any Intel Xeon SP processor will have capability to perform the MCA Recovery feature.
- **Improvements to the MEM8000 Event** – When a PowerEdge server logs the MEM8000 message in the System Event Log, it is intended to imply that BIOS has detected a rate of correctable errors that may impact performance. Starting from BIOS 2.0.3, Dell made an improvement to the CE rate detection scheme associated with the MEM8000 event. However, the improvement resulted in an uptick in MEM8000 events that were not substantiated by results from memory component failure analysis. With this PowerEdge server BIOS update, the following changes have been made:
    - Signaling of the MEM8000 event has been modified
    - PowerEdge server BIOS may schedule self-healing to occur on the identified DIMM

  Dell Technologies continues to investigate opportunities to improve the detection scheme related to an MEM8000 event.

What's New in BIOS 2.9.4

- **Optimization to Predictive Failure Algorithms** – Dell Technologies maintains its position that DRAMs experiencing certain types of correctable errors may later experience uncorrectable errors at the same memory locations. In BIOS 2.9.4, Dell Technologies has optimized its predictive failure algorithms to do a better job of focusing on these types of correctable errors. This means benign correctable errors are less likely to report false service warnings (MEM0802) and invoke performance impacting RAS features (such as ADDDC). Correctable errors that may evolve to uncorrectable errors are still detected and latched for mitigation by the system's RAS features.

Legal Notices

References

[1] P. Restle, J. Park and B. Lloyd, "DRAM Variable Retention Time," IEEE, 1992.

[2] K. Aadithya, A. Demir, S. Venugopalan and J. Roychowdhury, "Accurate Prediction of Random Telegraph Noise Effects in SRAMs and DRAMs," IEEE, 2013.

[3] "Reed–Solomon error correction," Wikipedia, [Online]. Available: https://en.wikipedia.org/wiki/Reed%E2%80%93Solomon_error_correction.

[4] V. Sridharan and D. Liberty, "A study of DRAM failures in the field," IEEE, 2012.

[5] A. Hwang, S. Ioan and B. Schroeder, "Cosmic rays don't strike twice: understanding the nature of DRAM errors and the implications for system design," ACM, 2012.