Cisco UCS and HX M5 Memory Technical Overview - Memory RAS Features

Contents

Abstract

Modern servers, including Cisco UCS® M5 servers, provide increased memory capacities and run at higher bandwidths and lower voltages. These trends, along with higher application memory demands and advanced multicore processors, contribute to an increasing probability of experiencing memory errors.

Intel® Xeon® Scalable Processors (formerly "Skylake Servers") and 2nd Gen Xeon Scalable Processors (formerly "Cascade Lake Servers") implemented changes in Single Device Data Correction (SDDC). SDDC is a fundamental Intel RAS (Reliability, Availability, Serviceability) feature available on all Cisco® platforms. As a result of these architectural changes and memory DIMM errors, there is a difference in which errors will be corrected between the previous generation of processors and the Xeon Scalable Processor family generation. This may result in a higher rate of uncorrectable memory errors. **Cisco UCS M5 servers incorporate microcode updates and BIOS enhancements that improve management of memory faults by enabling additional RAS features**.

This paper describes the classification and handling of memory errors on Cisco UCS M5 servers with first- and second-generation Intel Xeon Scalable Processors. **Cisco recommends upgrading UCS firmware to enable Adaptive Double Device Data Correction (ADDDC) Sparing and Post Package Repair (PPR)**. These features provide the optimal balance of performance, memory capacity, and error resilience. For mission-critical applications that are unable to tolerate a memory failure, memory mirroring should be considered.

Software requirements and recommendations

Refer to Table 1 for required and recommended firmware versions. Before upgrading firmware the Memory RAS BIOS policy should be reviewed and set appropriately to avoid additional reboots after upgrading firmware. See "Handling memory errors," below, for details and recommendations on selecting a policy setting.

Table 1. Required and recommended firmware

Software product	Minimum server firmware supporting ADDDC Sparing	Minimum server firmware supporting ADDDC Sparing and PPR	Recommended server firmware

Cisco UCS M5 blade servers and integrated Cisco UCS M5 Rack Servers	3.2(3p)
	4.0(4i)
	4.1(1d)
	4.1(2a)
		4.1(1d)
		4.1(2a)
			4.1(3d) or later

Standalone Cisco UCS M5 Rack Servers	3.1(3k)
	4.0(4l)
	4.1(1g)
	4.1(2a)
		4.1(1d)
		4.1(2a)
			4.1(3c) or later

Overview of memory errors

Memory errors are among the most common types of errors on modern servers. Errors are encountered when an attempt is made to read a memory location and the value read does not match the value last written.

Memory errors can be soft or hard. Some errors are correctable, but multiple simultaneous soft or hard errors on a single memory access may be uncorrectable. Overall error rates will scale with total memory capacity in a system, both by individual DIMM capacity and total quantity of DIMMs. The impacts of soft and hard errors are mitigated through hardware and firmware features as described in this document. Soft and hard errors can be introduced post manufacturing by high energy particle strikes or normal device wear over time[1].

Soft errors

Errors caused by brief electrical disturbances within the DRAM or on an external interface are referred to as " soft" errors. Soft errors are often transient and do not always repeat. If the error was the result of a disturbance during the read operation, then retrying the read may yield correct data. If the error was caused by a disturbance that upset the contents of the memory, then rewriting the memory location will correct the error.

Soft error rates can be affected by temperature, altitude, and memory access patterns specific to particular workloads. Note that workloads do not correlate directly to applications. The same application that might trigger a correctable error on a Dual In-Line Memory Module (DIMM) in one server may not generate an error with a different data set. Memory-test algorithms are tuned to represent worst-case workload behavior, but previously undetected errors may still occur during runtime. Cisco reviews and revises test algorithms to improve fault detection.

Hard errors

Errors caused by persistent physical defects are traditionally referred to as " hard" errors. Hard errors may be caused by an assembly defect such as a solder bridge or cracked solder joint or may be the result of a defect in the memory chip. Rewriting the affected memory contents and retrying read access will not eliminate a hard error. The error will persist.

Correctable errors

If errors are detected and corrected, they are considered correctable. This can be accomplished by retrying the read or by calculating the correct memory contents using ECC data and writing the correct data back into memory. After an error is detected and corrected, the Cisco Integrated Management Controller (IMC) will log the event in the System Event Log.

Typically, correctable errors are the result of soft errors. If correctable errors persist within the same memory location over an extended period, it may indicate a potential hard error.

[1] Single-event Effects of Space and Atmospheric Radiation on Memory Components

Uncorrectable errors

An error is deemed uncorrectable when it exceeds the correction capability of the processor's ECC engine.

An uncorrectable error experienced during runtime results in a catastrophic processor crash or hang, which will cause a server outage. This requires a reboot of the affected server and a replacement of the component that is at the root of the error. Usually this is the memory module, but the root cause could also be tied to a processor, processor socket, or DIMM socket.

After experiencing an uncorrectable error and rebooting the server, Cisco UCS will automatically map out (disable) the affected DIMM. This allows the server to return to service while preventing a second failure from the same module.

If any memory errors are detected during system power on testing, they are deemed uncorrectable and the module will be mapped out. This is often an indication of a hard error and the module should be replaced.

Minimizing early production runtime errors

Some errors are most likely to occur earlier than expected in the DIMM lifecycle. As noted, errors can be caused by electrical disturbances or physical defects. Some electrical disturbances can degrade a device and introduce a physical fault (hard errors). Prior to shipment, Cisco performs comprehensive system-level testing on all servers. Analysis of field errors has shown a correlation between physical shipping and new errors. Based on this analysis, Cisco recommends running an Enhanced Memory Test (EMT)[2] in combination with memory diagnostics prior to placing servers into production. This minimizes the time interval and potential introduction of errors between the latest memory testing and execution of production workloads.

Likewise, after upgrading or swapping memory, Cisco recommends running an EMT in combination with memory diagnostics to help identify installation errors and minimize potential early failures. The most common errors observed during memory upgrades are attributed to DIMM-seating or installation issues.

See “Testing memory,” below, for details about EMT and additional details on other memory testing options available.

Handling memory errors

Cisco UCS server ECC capabilities

All Cisco UCS M5 servers use memory modules with ECC codes that can correct any error confined to a single x4 DRAM chip and detect any double-bit error in up to two devices.

Scrub protocol

Cisco UCS M5 servers utilize demand and patrol scrubbing to address correctable errors and decrease the chance of a multibit error. These features are enabled by default on all UCS M5 servers.

If a correctable error is detected during a read transaction, **Demand scrub** writes corrected data back into memory.

Patrol scrub proactively scans all memory every 24 hours. It uses **Demand scrub** to read memory locations and correct any detected errors. This allows for errors to be corrected proactively, reducing the potential for impact during a future read event.

[2] Named “Advanced Memory Test (AMT)” in 4.1(3c) and earlier firmware releases.

Advanced RAS policies

Fundamental ECC capabilities and scrub protocols were historically successful at handling and mitigating memory errors. As memory and processor technologies advance, RAS features must evolve to address new challenges. As a result, Cisco UCS M5 servers provide several advanced memory RAS policies to improve server resilience, provide additional memory redundancy options, and streamline maintenance.

Adaptive Double Device Data Correction (ADDDC Sparing)

ADDDC Sparing can correct two successive DRAM failures if they reside in the same region. This feature tracks correctable errors and dynamically maps out failing bits by spare-copying contents into a "buddy" cache line. This mechanism can mitigate correctable errors that, if left untreated, could become uncorrectable. This feature uses virtual lockstep (VLS) to assign cache line buddy pairs within the same memory channel at either DRAM bank level using bank VLS or DRAM device level using rank VLS. Platinum and Gold CPUs support both bank and rank VLS. Silver and Bronze CPUs support only bank VLS.

If errors persist after a sparing event, the process repeats as needed until all of the spare bits are consumed. Spare bits are obtained from buddy cache line pairs by reusing redundant ECC bits produced by the lockstep process. ADDDC Sparing does not require allocation or usage of spare main memory regions and does not reduce overall memory available to the operating system.

When an ADDDC Sparing event occurs, Cisco UCS Manager (UCSM) will generate fault F1706 to guide the administrator to the server health event and server firmware will generate an SEL event. See "Handling RAS events," below, for event details by firmware version.

When enabled, this feature will incur a marginal memory latency and bandwidth penalty. The following tables show the measured impact to memory-intensive benchmark tools and various workloads when ADDDC Sparing is enabled. As spare bank or rank regions are utilized, the potential performance impact increases. Results are dependent on the actual workloads.

Table 2. ADDDC performance penalty

Parameter	% Bandwidth change ('-' indicates drop)
Memory access: ALL reads	-4.5%
Memory access: 3:1 reads-writes	-1.3%
Memory access: 2:1 reads-writes	-1.2%
Memory access: 1:1 reads-writes	+1.7%
Memory access: Stream-triad like	-4.4%
Performance: SPECrate2017_fp_base	-3%
Performance: SPECrate2017_int_base	-2.3%

Software configuration

ADDDC Sparing is supported and enabled by default in all required and recommended firmware releases listed in Table 1.

Note: When Partial Memory Mirroring is enabled, the remaining non-mirrored memory is not protected by ADDDC or other RAS features beyond baseline ECC.

For UCSM-managed servers with the BIOS POLICY for RAS configuration set to " Platform Default," no changes are required for ADDDC Sparing to take effect.

For UCSM-managed servers with the BIOS POLICY for RAS configuration NOT set to " Platform Default," then the policy must be changed to ADDDC Sparing (or Platform Default) to take advantage of ADDDC.

- Consult the chapter titled “Server-related policies, RAS memory BIOS settings” in the Cisco UCS Manager Server Management Guide for the installed version of UCSM.
  - Example: https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/ucs-manager/GUI-User-Guides/Server-Mgmt/4-1/b_Cisco_UCS_Manager_Server_Mgmt_Guide_4_1/4-1-trial_chapter_01100.html#reference_08F63A1A00F1470387C604B11CA9A5BF

For standalone (non-UCSM-managed) servers, no changes are required for ADDDC Sparing to take effect.

- Consult the chapter titled “Managing the server, configuring BIOS settings” in the Cisco UCS C-Series Integrated Management Controller GUI Configuration Guide.
  - Example: https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/c/sw/gui/config/guide/4_1/b_Cisco_UCS_C-series_GUI_Configuration_Guide_41/b_Cisco_UCS_C-series_GUI_Configuration_Guide_41_chapter_0100.html#task_xs1_d4z_j1b

Memory mirroring

Memory mirroring utilizes a portion of system memory to store a duplicate copy of memory contents. Dynamic failover (without reboot) to the mirrored memory is transparent to the operating system (OS) and application, providing protection against uncorrectable errors that would result in an outage.

Full memory mirroring provides fully redundant system memory. Half of the physical memory capacity is used to store duplicate (mirrored) copies of memory contents. This reduces the memory available to the OS by half. This RAS configuration should be considered on systems that are not able to withstand any individual server outages due to memory errors.

Table 3. Full memory mirroring performance change

Parameter	Mirroring performance penalty (ratio-enabled/disabled)
Memory access: ALL reads	1.00
Memory access: 3:1 reads-writes	0.78
Memory access: 2:1 reads-writes	0.68
Memory access: 1:1 reads-writes	0.61
Memory access: STREAM-triad-like	0.66
Performance: SPECrate2017_fp_base	0.88
Performance: SPECrate2017_int_base	0.95

Partial memory address mirroring allows users to define portions of the memory to be mirrored. This memory range can be determined through interaction between the OS and the server, allowing for critical code to operate more reliably. This feature, introduced in Cisco UCS Manager Release 4.1, establishes a middle ground by increasing reliability without sacrificing a significant amount of memory. Partial mirroring requires support from the host OS.[3]

Note: **When Partial Memory Mirroring is enabled, the remaining non-mirrored memory is not protected by ADDDC or other RAS features beyond baseline ECC.**

Software configuration

Full and partial mirror mode is supported in all required and recommended firmware releases.

- For UCSM-managed servers, consult the chapter titled “Server-related policies, partial memory mirroring” in the Cisco UCS Manager Server Management Guide for the installed version of UCSM.
  - Example: https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/ucs-manager/GUI-User-Guides/Server-Mgmt/4-1/b_Cisco_UCS_Manager_Server_Mgmt_Guide_4_1/4-1-trial_chapter_01100.html#id_126266
- For standalone Cisco UCS C-Series Rack Servers, consult the chapter titled “Managing the server, configuring BIOS settings” in the Cisco UCS C-Series Integrated Management Controller GUI Configuration Guide.
  - Example: https://www.cisco.com/c/en/us/td/docs/unified_computing/ucs/c/sw/gui/config/guide/4_1/b_Cisco_UCS_C-series_GUI_Configuration_Guide_41/b_Cisco_UCS_C-series_GUI_Configuration_Guide_41_chapter_0100.html#task_xs1_d4z_j1b

[3] VMware KB article 2146595 on Configuring Reliable Memory

Maximum Performance

The Maximum Performance RAS setting disables all firmware RAS protection beyond baseline ECC. This provides the least impact on performance, but also offers the least protection from uncorrected errors.

Post Package Repair (PPR)

Post Package Repair (PPR) can permanently repair failing memory regions within a DIMM by leveraging redundant DRAM rows. This permanent in-field repair allows for rapid recovery from hard errors without needing to replace the DIMM. To perform a repair, the system must experience an ADDDC event and go through at least one reboot cycle. This repair activity does not affect performance or the total memory available to the OS.

Cisco performs regular failure analysis on DIMMs to understand specific underlying causes and drive continuous quality improvements. Based on the analysis of 273 field-returned DDR4 DIMMs, PPR will repair 70 percent of DIMM faults that were detected by ADDDC. PPR will not initiate on DIMMs that have suffered uncorrectable ECC errors.

Software configuration

See Table 1 for required and recommended firmware releases with PPR support. For PPR to take effect, ensure that ADDDC Sparing is enabled. No other configuration is required. The Platform Default for PPR type is Hard PPR. Disabling PPR is not recommended. There is no performance or other impact when PPR is enabled.

Configuring error reporting

Corrected errors can be logged by platform firmware or by the host operating system. Cisco recommends allowing firmware to log errors and disabling OS-based logging to avoid conflicting logs due to differences in log-collection timing.

Firmware-based error logging is enabled by default. It cannot be disabled.

OS-based tools can interfere with firmware monitoring, depending on error register management. Further, some OS-based monitoring subsystems are known to encounter issues under error storms, which can also impact system performance. For these reasons Cisco recommends disabling OS-based error monitoring.

When an error occurs, a Corrected Machine Check Interrupt (CMCI) may be invoked to alert the operating system of the condition. CMCI interrupts can impact performance during error storms. Disabling CMCI can be done in BIOS settings. Errors can still be detected through polling.

Disabling OS-based error monitoring

Linux

Linux supports two features related to error monitoring and logging: EDAC[4] (Error Detection and Correction) and mcelog.[5] Both are common in most recent Linux distributions. Cisco recommends disabling EDAC-based error collection, to allow all error reporting to be handled in firmware.

From kernel versions 3.14 onward, EDAC can be disabled by adding the option “**edac_report=off**” to the kernel command line.

[4] EDAC Project

[5] Linux mcelog daemon

Mcelog is enabled by default in most recent Linux distributions.

Refer to your Linux distribution documentation for further instructions.

For customers who prefer to collect all diagnostic and fault information from OS resident tools, mcelog is recommended. In this case, Cisco recommends disabling CMCI to prevent performance impact. **Firmware logs may be incomplete when OS logging is enabled**.

Testing memory

Enhanced Memory Test on boot

Enhanced Memory Test (EMT)[6] provides additional memory tests during the server boot process. This feature leverages optimized memory test patterns to identify weak or failing memory locations and repair them with Post Package Repair (PPR). EMT can be enabled in the BIOS policy. It is recommended for use during predeployment burn-in testing and diagnostic workflows. EMT is disabled by default because it can add minutes of additional boot time.

The following test-times have been measured in typical system configurations. The times listed are for the expected worst cases. Actual times can be up to 50% lower and will be consistent across boots for a given server configuration. The test-times scale roughly linearly with capacity, but are affected by core count and clock, memory type, and memory configuration. Errors encountered during tests can extend boot time to perform repairs. It is recommended to execute a boot with EMT enabled on a representative system to determine accurate timing for your configuration.

Table 4. Estimated additional boot time for EMT completion

Total memory	Additional boot time for EMT completion
768 GB	8 minutes
1.5 TB	15 minutes
3 TB	30 minutes

Enhanced Memory Testing was introduced as a BIOS policy option in Cisco UCS Manager and standalone Cisco Integrated Management Controller (IMC) Release 4.1(3). Once enabled, EMT will run on subsequent system boots. If the policy is set to “Auto” the test will run on any boot if there were memory errors during the previous system uptime.

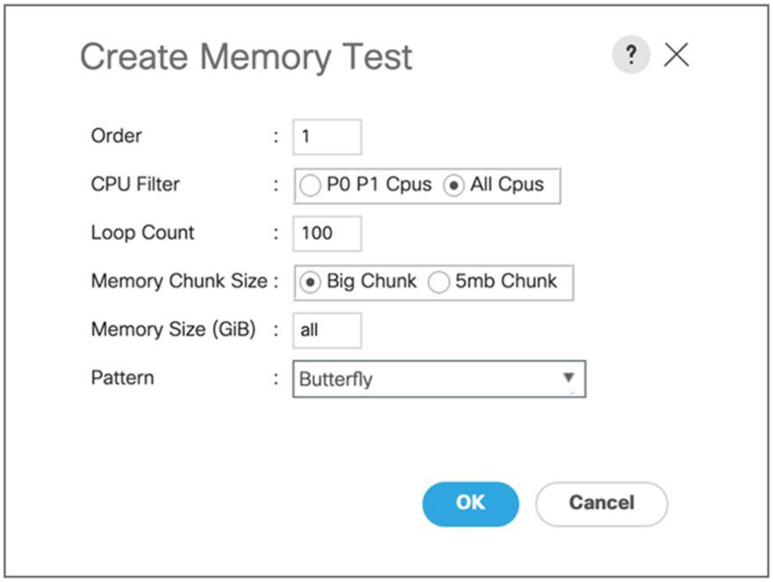
[6] Named Advanced Memory Test in 4.1(3c) and earlier firmware releases.

Server memory diagnostics

Execution of memory diagnostics is dependent on the UCS management mode being used. For UCSM-managed servers, Cisco UCS Manager includes built-in memory-diagnostic capabilities. For standalone Cisco UCS C-Series Rack Servers, UCS Diagnostics bootable media are available from Cisco.com software downloads.

Field experience has shown positive benefits from running 100 loops of the memory butterfly pattern with a “Big Chunk” size, as shown in the example below. This test will take approximately eight hours for a server with 768GB of memory across all 24 DIMMs. Runtime will scale roughly linearly with loop count and memory capacity.

Performing an Enhanced Memory Test is recommended in combination with a running of memory diagnostics.



Create Memory Test

Order : 1

CPU Filter : ○ P0 P1 Cpus ● All Cpus

Loop Count : 100

Memory Chunk Size : ● Big Chunk ○ 5mb Chunk

Memory Size (GiB) : all

Pattern : Butterfly

Figure 1.
Example of diagnostic policy test for UCSM-managed servers

For complete details on running Cisco UCS memory diagnostics, refer to the following documents.

- For UCSM-managed servers, consult the chapter titled “Diagnostics configuration” of the **UCS Manager Server Management Guide** for the installed version of UCSM.
- For standalone Cisco UCS C-Series Rack Servers, consult the **Cisco UCS Server Diagnostics Utility User Guide**.

Handling RAS events

BANK-level and RANK-level RAS events are noncritical faults indicating that a potentially bad region of memory has been dynamically swapped out during runtime. During this state, the system is operational but incurs a marginal performance penalty to memory bandwidth and latency. Systems can run reliably for extended periods of time with one or more DIMMs in ADDDC state.

When BANK-level or RANK-level RAS events are observed (and PPR is enabled):

1. Verify that no other DIMM faults are present (for example, an uncorrectable error).
2. Schedule a maintenance window (MW).
3. During MW, put the host in maintenance mode and reboot the server to attempt a permanent repair of the DIMM using Post Package Repair (PPR).
    a. If no errors occur after reboot, PPR was successful, and the server can be put back into use.
    b. If new ADDDC events occur, repeat the reboot process to perform additional permanent repairs with PPR.
4. If an uncorrectable error occurs after reboot, replace the DIMM.

Table 5. RAS events by firmware release

Firmware Version	CPU SKU	Bank Event	Rank Event
4.0(4)	Standard	N/A	N/A
	Advanced	N/A	F1706
4.1(1)	All	F1706	F1706
4.1(2), 4.1(3)	Standard	F1706	N/A
	Advanced	F1705	F1706
4.2(1)	Standard	F1705	N/A
	Advanced	F1705	F1706

Related documentation

- Cisco UCS C220/C240/B200 M5 Memory Guide
- Cisco UCS B480 M5 Memory Guide
- Cisco UCS C480 M5 Memory Guide

Conclusion

Cisco UCS M5 servers incorporate microcode updates and BIOS enhancements that improve management of memory faults by enabling additional RAS features.

Cisco recommends a firmware upgrade to enable ADDDC Sparing and to expand the memory fault coverage. (Refer to Table 1.)

Cisco recommends running memory diagnostics prior to placing servers into production to mitigate early runtime errors. Running Enhanced Memory Test on boot is also recommended.

For mission-critical applications that are unable to tolerate a memory failure, memory mirroring should be considered.

Document history

Date	Updates
July 7, 2020	Initial release
July 13, 2021	Memory test enhancements and updated firmware recommendations
January 29, 2022	Updated firmware recommendations. Technical reference for failure modes. Details for firmware events related to ADDDC.

Americas Headquarters
Cisco Systems, Inc.
San Jose, CA

Asia Pacific Headquarters
Cisco Systems (USA) Pte. Ltd.
Singapore

Europe Headquarters
Cisco Systems International BV Amsterdam,
The Netherlands

Cisco has more than 200 offices worldwide. Addresses, phone numbers, and fax numbers are listed on the Cisco Website at **https://www.cisco.com/go/offices**.

Printed in USA C17-743902-04 01/22