

Cisco UCS M7 Memory Technical Overview – Memory RAS Features

Contents

Abstract	3
Software requirements and recommendations	3
Overview of memory errors	3
Soft errors	4
Hard errors	4
Correctable errors	4
Uncorrectable errors	4
Minimizing early production runtime errors	5
Cisco UCS M7 Memory RAS features	5
System-level ECC	5
Virtual Lock-Step (VLS) / Adaptive Double Device Data Correction (ADDDC) Sparing	5
On-die ECC	6
Error Check and Scrub (ECS)	6
Boot-Time Post Package Repair (PPR)	6
Command/Address Parity Check and Retry	6
Memory Mirroring	6
Cisco UCS M7 memory error handling through RAS features	6
Enhanced Memory Test	6
Initiation of VLS/ADDDC	7
DIMM Blocklisting	7
Conclusion	7
References	7

Abstract

Modern servers, including Cisco UCS® M7 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, require additional memory RAS features to maintain system operation and performance.

Intel® 4th Gen Xeon® Scalable Processors (formerly “Sapphire Rapids Servers”) and 5th Gen Xeon Scalable Processors (formerly “Emerald Rapids Servers”) have implemented changes to incorporate DDR5-supported Reliability Availability and Serviceability (RAS) features. And these features are now available on all Cisco M7 platforms.

This paper describes the classification and handling of memory errors on Cisco UCS M7 servers with 4th and 5th generation Intel Xeon scalable processors. **Cisco UCS firmware supports features such as 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. Memory mirroring is recommended for mission-critical applications unable to tolerate memory failure.

Software requirements and recommendations

Refer to Table 1 below for the baseline firmware version supporting the RAS features described in this document. Before upgrading firmware, the Memory RAS BIOS policy should be reviewed and set appropriately to avoid additional reboots after the upgrade. See section “Cisco UCS M7 memory error handling,” below, for details and recommendations on selecting a policy setting.

Table 1. Baseline firmware

Software product	Minimum server firmware supporting ADDDC Sparing and PPR	Recommended server firmware
Cisco UCS M7 blade servers and integrated Cisco UCS M7 Rack Servers	4.3(3a)	4.3(5c)

Overview of memory errors

Memory errors are among the most common types of errors on modern servers. Errors are often discovered 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. DDR Memory is increasingly made with smaller process nodes run at higher frequency with reduced voltages, which increases the potential for errors, so mitigating these errors is critical to basic functionality.

Soft errors

Errors caused by brief electrical disturbances within the DRAM or on an external memory 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 regularly 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 (SEL).

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.

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 usually 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) 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 the EMT tool 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.

It is recommended to enable PPR while running EMT. This gives the system the ability to immediately repair weak or faulty DIMM rows instead of experiencing failures during runtime.

Cisco UCS M7 Memory RAS features

Cisco UCS M7 servers have a robust set of RAS features, detailed below, to minimize the impact of memory errors on performance and system uptime.

System-level ECC

All Cisco UCS M7 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. This is now referred to as system-level ECC, as in older-generation servers.

Virtual Lock-Step (VLS) / 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 (“sparing”) 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 the DRAM bank-level using bank VLS or the DRAM device-level using rank VLS.

Intel CPUs support ADDDC in Single-Region (SR) and Multi-Region (MR). ADDDC-MR is another feature where ADDDC can correct Correctable Errors (CEs) in different banks or ranks on two DIMM regions in a memory channel.

This feature is supported only while using DDR5 x4 DIMMs. CPU families supporting only Standard RAS features can support only ADDDC-SR, whereas CPU families with Advanced RAS features support ADDDC-SR and ADDDC-MR.

Platinum and Gold Intel CPUs support both bank and rank VLS. Silver and Bronze Intel 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 virtual lockstep process. ADDDC Sparing does not require allocation or usage of the spare main memory regions and does not reduce the overall memory available to the operating system.

On-die ECC

On-die ECC is a new feature in DDR5. This feature is enabled by default. All single-bit errors (hard and soft) are corrected by DRAM before data is transmitted to the host. However, this corrected data is not written back to DRAM. Error Check and Scrub (ECS) is the feature used to scrub and correct single-bit errors in memory.

Error Check and Scrub (ECS)

ECS checks for errors in the background by scrubbing each DRAM die periodically (every 24 hours), correcting them by writing data back to the array and providing a count of the errors found during the scrub. This feature is enabled by default.

Boot-Time Post Package Repair (PPR)

Post package repair is a feature where spare rows are used to replace a bad cell or row in a DRAM device. Cisco UCS M7 servers with Intel CPUs support “hard” PPR. This is a permanent repair and is carried out during reboot based on the error data collected during the previous runtime or if any row errors are encountered during EMT.

Note:

- Boot-time PPR relies on data stored from the previous boot to repair the DIMMs; therefore, rebooting the system once before doing a BIOS update is highly recommended so that data requesting PPR is not lost.
- PPR status is logged to BMC.

Command/Address Parity Check and Retry

Command/Address (CA) parity protects the command and address interface. If a CA parity error occurs on a write cycle, the DIMM will alert the host and will not write data to memory. The host will have an opportunity to retry the operation. If a CA parity error occurs on a read cycle, the host will drop the data and retry the read. This feature is enabled by default.

Memory Mirroring

Memory mirror is a final backup to handling memory failures that cannot be handled by the features listed above. A duplicate copy of all data is maintained in case the original data is corrupted beyond repair.

Cisco UCS M7 memory error handling through RAS features

Cisco UCS M7 servers have additional features, described below, that are used to decide when a particular RAS feature is utilized to manage DIMM-related errors.

Enhanced Memory Test

Enhanced memory test (EMT) is a suite of test patterns provided by Intel to stress the system memory at boot time.

If correctable or uncorrectable errors occur at runtime, EMT, by default, if set to **Auto** in BIOS, is enabled to run on the next boot to test the entire memory. If PPR is enabled, EMT PPR can attempt to repair damaged regions on the DIMM immediately and return the DIMM to operation. Note that this operation will add a significant delay in the boot time. Depending on the outcome, the DIMM may become disabled if the DIMM fails EMT and cannot be repaired through PPR.

Initiation of VLS/ADDDC

VLS will be triggered depending upon the number of errors detected in the same row. VLS will attempt to temporarily swap the bank or rank with a spare at runtime.

DIMM Blocklisting

Cisco Memory Blocklisting, when enabled, will disable any DIMM that previously had an uncorrectable error until it is either removed from the system or PPR is successful, and memory errors are cleared from UCSM/IMM.

Conclusion

The features described above work together to help prevent correctable and uncorrectable errors from severely degrading system performance. Uncorrectable errors may, but not always, bring a system down. Single-bit correctable errors are expected to occur sporadically and fixed by ECC before any corruption occurs. Single-bit errors on their own are not a cause for concern and do not require user intervention. Cisco UCS M7 servers employing DDR5 DIMMs use multiple memory RAS features for improved system uptime and reduced service actions compared to previous-generation servers.

References

- Intel document Doc. No.: 638563, Rev.: 1.75

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>.

Cisco and the Cisco logo are trademarks or registered trademarks of Cisco and/or its affiliates in the U.S. and other countries. To view a list of Cisco trademarks, go to this URL: <https://www.cisco.com/go/trademarks>. Third-party trademarks mentioned are the property of their respective owners. The use of the word partner does not imply a partnership relationship between Cisco and any other company. (1110R)