Introduction to Memory RAS (Reliability, Availability, and Serviceability) Features on ThinkSystem Servers

Oct. 2023

Jason Liu, Kim Jin, Paul Klustaitis

Abstract

This technical white paper provides an overview of fundamental memory technologies enabled on Lenovo ThinkSystem Servers. It then proceeds to introduce both POST (Power On Self Test) RAS (Reliability, Availability, and Serviceability) features and Runtime RAS features across all generations of Xeon® based systems. This comprehensive whitepaper not only acquaints readers with the concepts behind each unique RAS feature but also supplements the explanations with numerous detailed diagrams, flowcharts, and screenshots of ThinkSystem UEFI setup pages to make the content more accessible to the end user.

Simultaneously, the document enumerates all relevant System Event Log (SEL) events associated with each feature. This listing empowers users to gain a full comprehension of various memory-related events and provides them with in-depth technical insights into memory errors or statuses when such events occur.

Memory Basic Technology Introduction

To facilitate comprehension of the memory RAS features, we begin by providing some foundational context on memory technology. Intel Xeon® CPU incorporates multiple integrated Memory Controllers (iMCs) within the processor die. Each iMC has the capacity to accommodate multiple DDR channels, and each of these DDR channels can support either one or two DIMM slots.

DDR4 DIMM has ECC (Error-Correcting Code) memory type and non-ECC memory type. The ECC memory is a type of computer memory (Random Access Memory - RAM) that includes additional error-checking and correction capabilities beyond what is found in standard, non-ECC memory. The primary purpose of ECC memory is to detect and correct errors that may occur during data storage or processing, particularly in critical computing environments where data accuracy and reliability are essential.

Looking closer at DDR4 JEDEC Serial Presence Detect (SPD) data structure we can use Byte 13 “Module Memory Bus Width” to decode ECC functionality. Figure 1, Bits 4~3 encode the bus extensions such as parity or ECC, for example, 64 bit primary bus, no parity or ECC (64 bits total width): xxx 000 011, and 64 bit primary bus, with 8 bit ECC (72 bits total width): xxx 001 011.

Bits 7~5	Bits 4~3	Bits 2~0
Reserved	**Bus width extension, in bits**	**Primary bus width, in bits**
Reserved; must be coded as 000	Bit [4, 3] : 000 = 0 bits (no extension) 001 = 8 bits All others reserved	Bit [2, 1, 0] : 000 = 8 bits 001 = 16 bits 010 = 32 bits 011 = 64 bits All others reserved

Figure 1: DIMM SPD Byte 13 Definition

ThinkSystem servers only support Enterprise grade ECC DIMMs which take advantage of many memory RAS features. DIMMs with the "x4" designation refers to the number of data lanes used on the Dynamic Random Access Memory (DRAM) module to transfer data between all memory module DRAMs of DIMMs and the memory controller iMC which is located inside the CPU. For example if a x4 ECC DIMM has 18 x4 DRAMs; and each DRAM has its own Bank and Bank Group. A rank is a concept that pertains to the entire DIMM, encompassing all DRAMs within it. For further clarification, please refer to "Figure 2: DIMM Components and DIMM Slots in DDR Channel".

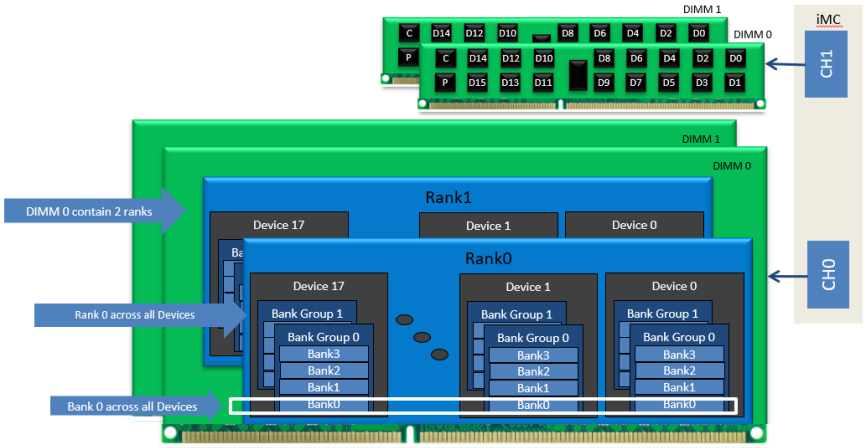


Figure 2: DIMM Components and DIMM Slots in DDR Channel

Intel Xeon® processor's integrated Memory Controller (iMC) will fetch the data from DIMM at cache line level. When the DIMM is x4 type, the iMC will use 8 bursts to fetch 18 x4 DRAMs as one cache line with ECC bits (18 * 4 * 8 = 576 bits = 512 bits <64 bytes> cache line + 64 bits ECC data); Consequently, Xeon® processors can utilize ECC DIMMs to perform cache line-level memory error correction based on the ECC data, see "Figure 3: Fetching Cache Line with ECC from DIMM" for more details.

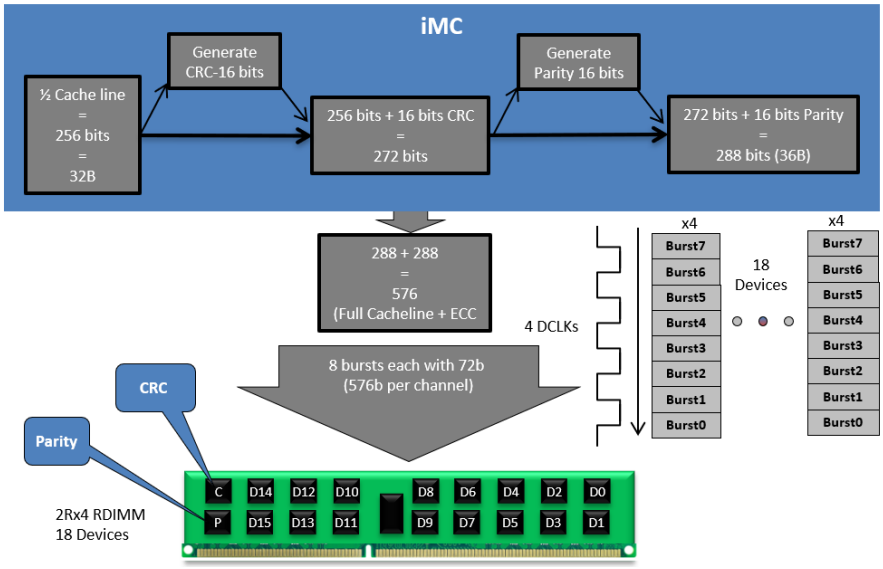


Figure 3: Fetching Cache Line with ECC from DIMM

ThinkSystem Memory RAS Overview

All Lenovo ThinkSystem servers running Intel's 1st Generation Xeon® Scalable processors (formerly codename "Skylake"), and 2nd Generation Xeon® Scalable processors (formerly codename "Cascade Lake"), ThinkSystem 2nd generation servers running Intel's 3rd Generation Xeon® Processors (codename "Ice Lake-SP" and "Cooper Lake"), and ThinkSystem 3rd generation servers running Intel's 4th generation Xeon® Processors (formerly codename "Sapphire Rapids") support DDR4 ECC DIMMs. Lenovo ThinkSystem servers come equipped with comprehensive support for all DDR4 memory RAS (Reliability Availability Serviceability) features offered by the chipset manufacturer. Considering the ThinkSystem boot phases, we can categorize the complete set of RAS (Reliability, Availability, and Serviceability) features into two distinct segments:

- POST (Power On Self Test) RAS Features: These encompass all features managed from the moment the system powers on until the Operating System boots.
- Runtime RAS Features: These encompass all features managed while the Operating System is running.

POST Memory RAS Features

During ThinkSystem power on phase, Lenovo UEFI (Unified Extensible Firmware Interface) will detect and report non-Lenovo branded DIMMs. UEFI will rely on Intel Memory Reference Code (MRC) to carry out memory inintialization. If a DIMM encounters an error during memory training, it will trigger the prompt reporting of a DIMM error. As a precaution, the affected DIMM or even the associated DDR channel may be disabled to prevent the use of a malfunctioning DIMM within the Operating System.

To ensure comprehensive detection of failing DIMMs, the ThinkSystem's UEFI also offers an Advanced Memory Test (AMT) feature within the Lenovo xClarity Provisioning Manager (LXPM). This functionality facilitates thorough memory testing and repair of all DIMMs installed in the server, leading to a reduction in memory errors during OS operation.

In case some DIMMs exhibit limited hardware issues, such as bad rows, ThinkSystem will use Post Package Repair (PPR) feature to perform hard PPR. This process effectively and permanently repairs a failing DIMM by substituting the access to a bad row with a spare row within the DRAM device, transforming a bad DIMM into a completely healthy DIMM.

Lenovo Branded DIMM Identifier

Lenovo branded DIMMs (a.k.a. TruDDR) receive the highest level of support within Lenovo because they undergo comprehensive validation, including Signal Integrity (SI) assessment for

hardware compatibility, as well as a battery of functional and stress tests before being shipped to the customers.

ThinkSystem UEFI provides a comprehensive solution in identifying whether all DIMMs installed in the ThinkSystem are Lenovo branded or not. When UEFI detects a non-Lenovo branded DIMM, it will generate a SEL (System Event Log) entry labeled “FQXSFMA0006I” for that DIMM but will continue the memory training process as usual as normal. If a non-Lenovo branded DIMM is removed or replaced with Lenovo branded DIMM, the system UEFI will report SEL “FQXSFMA0007I” as reference.

Event ID	Severity	Message String	Notes
FQXSFMA0006I	Info	Unqualified DIMM [arg1] has been detected, the DIMM serial number is [arg2].	[arg1]: DIMM silk label [arg2]: DIMM identifier
FQXSFMA0007I	Info	Unqualified DIMM number [arg1] has been replaced. [arg2]	[arg1]: DIMM silk label [arg2]: DIMM identifier

Failing DIMM Disable and Re-enable

Lenovo ThinkSystem UEFI will report DIMM status in UEFI setup page: “System Settings > Memory > System Memory Details > DIMM Details for Processor x”. Under normal conditions when a DIMM is healthy and configured correctly, the related DIMM will have one of two states:

- ”Present Enabled” state --- DIMM is trained successfully and is used as main system memory
- ”Present, Partially Rsvd. for RAS” state --- DIMM is trained successfully and is used in memory rank sparing or memory partial mirror/full mirror mode

When ThinkSystem UEFI detects any DIMM with memory error, or memory training failure, or runtime error that occured duringlast boot, then the DIMM status in the setup page will be designated either ”Disabled” or ”Present, Not Configured”. Several distinct cases are outlined below to illustrate this scenario:

- DIMM in ”Disabled” state --- DIMM is mapped out by MRC because of one of below reasons:
    - DIMM has failed memory training/test
    - DIMM had encountered runtime uncorrectable error during previous boot.

As illustrated in below Figure 4, the help string will provide an explanation why the DIMM was disabled.

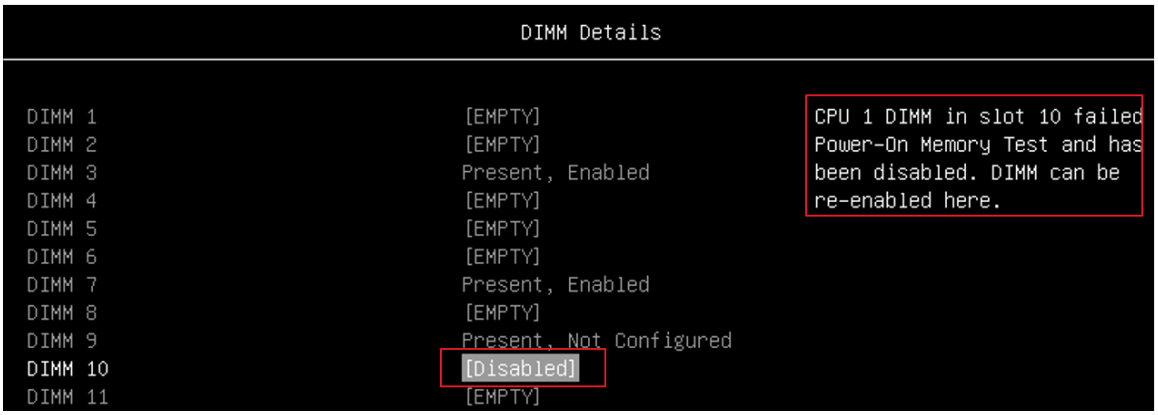


Figure 4: DIMM Status In UEFI Setup

- The paragraph below explains how to re-enable the previously disabled DIMM:
  - User replaces a “bad” DIMM with a new DIMM
  - User has an option to manually re-enable DIMM in the setup by changing the status from "Disable" to "Enable" (see Figure 5 below as one example). This action will trigger memory training during the next boot to determine if the DIMM can be used. The outcome will be either "Present, Enabled" if the DIMM passes the training or "Disabled" if it still fails the memory training.

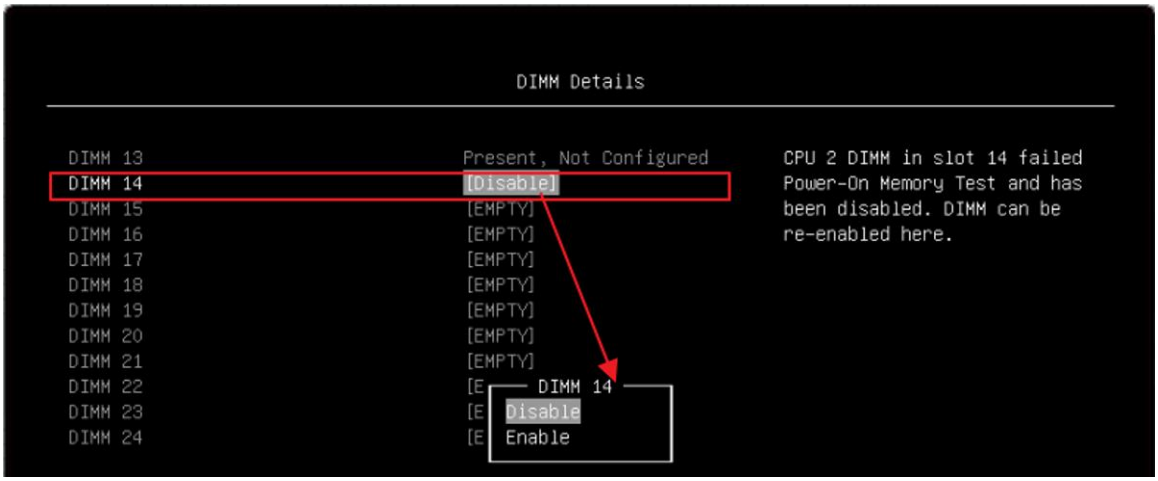


Figure 5: DIMM Re-enable In UEFI Setup

- A DIMM in a "Present, Not Configured" state represents an unusual scenario that requires attention (see Figure 6 below as one example). The DIMM is excluded from memory map by the Memory Reference Code (MRC) due to one of the following conditions: if failed DIMM is disabled and another DIMM in same DDR channel is mapped out by MRC and shown as ”Present, Not Configured”;
- If MRC cannot isolate the error to a specific DIMM in DDR channel then MRC will map out the whole channel and show 'Present, Not Configured' for all DIMM(s) present in that channel.

This situation can be resolved by either replacing the failed DIMM(s) in the DDR channel or by forcing a retraining of the failed DIMM during the next boot.

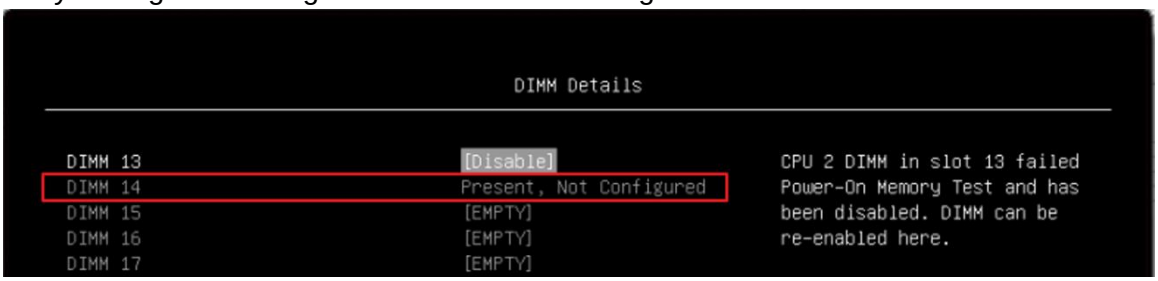


Figure 6: DIMM “Present, Not Configured” In UEFI Setup

There are several distinct scenarios that can lead to the disabling of a DIMM, and ThinkSystem UEFI will generate corresponding SEL entries to assist customers in addressing these cases conveniently:

- **Case 1 – DIMM training failure or test failure has been detected during boot.**

  When a DIMM fails memory training or memory test, UEFI will report SEL ”FQXSFMA0008M” (DIMM x has failed the POST memory test) and ”FQXSFMA0001M” (DIMM x has been disabled due to an error detected during POST).
  Additionally, if the failed DIMM resides in slot 0 (far from CPU) of that DDR channel, the DIMM in slot 1 (near to CPU) of the same channel will also be disabled, and UEFI will report SEL entry ”FQXSFMA0052I” (stating that DIMM x has been disabled due to the error on DIMM y).

- **Case 2 – DIMM has experienced runtime uncorrectable error during previous boot.**
  ThinkSystem UEFI retains information about a failing DIMM when it reports an uncorrectable error during runtime. In the subsequent boot, UEFI will disable the failing DIMM and generate a corresponding SEL FQXSFMA0001M” (DIMM x has been disabled due to an error detected during POST).

- **Case 3 - DIMM configuration doesn’t match any of the supported DIMM population rules.**
  ThinkSystem UEFI checks the arrangement of DIMMs against predifined rules supported by Intel processors and validated by Lenovo. If the current DIMMs population does not align with the supported rules, UEFI will attempt to disable the minimum number of DIMMs necessary to conform to an appropriate supported population rule. Any DIMM disabled due to a violation of this rule will be indicated as ”Present, Not Configured” in ”DIMM Details” in UEFI F1 Setup, accompanied by the relevant SEL entry "FQXSFMA0053M".

Event ID	Severity	Message String	Notes
FQXSFMA0001M	Error	DIMM [arg1] has been disabled due to an error detected during POST. [arg2]	[arg1]: DIMM silk label [arg2]: DIMM identifier
FQXSFMA0008M	Error	DIMM [arg1] has failed the POST memory test. [arg2]	[arg1]: DIMM silk label [arg2]: DIMM identifier
FQXSFMA0052I	Error	DIMM [arg1] has been disabled due to the error on DIMM [arg2].[arg3]	[arg1]: Good DIMM silk label [arg2]: Failing DIMM silk label [arg3]: Good DIMM identifier
FQXSFMA0053M	Warning	DIMM [arg1] not defective but disabled due to unsupported memory module combination on CPU [arg2].	[arg1]: DIMM silk label. [arg2]: CPU socket number

Advanced Memory Test

The standard memory training primarily conducts a basic memory test. However, to ensure comprehensive detection of failing DIMMs, ThinkSystem UEFI offers an Advanced Memory Test (AMT). This advanced test is more extensive compared to POST MRC memory test and can take up to 30 minutes or longer, depending on the system's DIMM capacity and the number of DIMMs installed.

Additionally, during AMT, if any weak cells or rows are detected, it utilizes Post Package Repair to perform row repairs. This valuable feature is accessible through Lenovo's graphical UEFI setup engine, known as Lenovo XClarity Provisioning Manager (LXPM).

For further information and in-depth details, please refer to the Lenovo public white paper titled "ThinkSystem Advanced Memory Test Guidance." The precise link to this resource can be found in the References section of this document.

Post Package Repair

Based on JEDEC spec, DDR4 offers optional fail row address repair support for 4Gb DRAMs and mandates it for 8Gb DRAMs and beyond. This feature allows for the repair of one row per bank group. PPR (Post Package Repair) presents a straightforward and efficient system-level repair method, with two distinct PPR types:

- Hard Post Package Repair (hPPR): repair the fail row element in a Bank Group on a DDR4 DRAM device by the electrical programming of electrical-fuse scheme; it is permanent repair of a row element.
- Soft Post Package Repair (sPPR): a way to quickly, but temporarily, repair a row element in a Bank Group on a DDR4 DRAM device; If system power is removed or the system is rebooted (reset), the DIMM reverts to its original state.

The presence of support for PPR can be determined by examining DIMM SPD Byte 9, which corresponds to "Other SDRAM Optional Features." The table below in Figure 7, sourced from the DDR4 SDRAM data sheet, provides details on Byte 9:

Bits 7~6	Bit 5	Bits 4~0
Post Package Repair (PPR)	**Soft PPR**	**Reserved**
00: PPR not supported 01: Post package repair supported, one row per bank group 10: Reserved 11: Reserved	0 = Soft PPR **not** supported 1 = Soft PPR supported	Reserved; must be coded as 00000
Note: If PPR is supported, Hard PPR is always supported by the device. Optional additional support for Soft PPR is indicated in bit 5.		

Figure 7: DIMM SPD Byte 9 Definition

As a default setting, Lenovo ThinkSystem employs a hard Post Package Repair (PPR) approach to ensure the permanent repair of failing or faulty rows within DIMMs, and restoring (when possible) them to a healthy state. When UEFI activates PPR for a specific DIMM, ThinkSystem will generate the following PPR SEL entries to inform customers and guide them on any necessary subsequent actions:

Event ID	Severity	Message String	Notes
FQXSFMA0026I	Info	DIMM [arg1] Self-healing, attempt post package repair (PPR) succeeded. [arg2]	[arg1]: DIMM silk label. [arg2]: DIMM identifier
FQXSFMA0029I	Info	The PFA of DIMM [arg1] has been deasserted after applying PPR for this DIMM. [arg2]	[arg1]: DIMM silk label. [arg2]: DIMM identifier
FQXSFMA0027M	Warning	DIMM [arg1] Self-healing, attempt post package repair (PPR) failed at Rank [arg2] Sub Rank [arg3] Bank [arg4] Row [arg5] on Device [arg6]. [arg7]	[arg1]: DIMM silk label. [arg2]: Rank number [arg3]: Sub Rank number [arg4]: Bank number [arg5]: Row number [arg6]: Device number [arg7]: DIMM identifier
FQXSFMA0028M	Warning	DIMM [arg1] Self-healing, attempt post package repair (PPR) exceeded DIMM level threshold [arg2] at Rank [arg3] Sub Rank [arg4] Bank [arg5] Row [arg6] on Device [arg7]. [arg8]	[arg1]: DIMM silk label. [arg2]: PPR Attempt Threshold [arg3]: Rank number [arg4]: Sub Rank number [arg5]: Bank number [arg6]: Row number [arg7]: Device number [arg8]: DIMM identifier

Runtime Memory Error Handling

DIMM Rank Sparing

Lenovo ThinkSystem servers running Intel's 1st Generation Xeon® Scalable processors (formerly codename "Skylake"), and 2nd Generation Xeon® Scalable processors (formerly codename "Cascade Lake") support Rank Sparing. This RAS feature was dropped from ThinkSystem 2nd generation servers running Intel's 3rd Generation Xeon® Processors (codename "Ice Lake-SP" or "Cooper Lake").

Rank Sparing feature allows dynamic fail-over of failing Ranks to spare Ranks behind the same memory controller (within same DDR4 channel). An unused spare rank on the channel can be used by copying the contents of a failing rank on that channel to the spare rank.

Rank sparing enables a failing rank to be replaced by an unused spare rank which is located in the same DDR channel. ThinkSystem UEFI will detect that a rank's operation becomes marginal in such a way that a single device is experiencing excessive corrected errors, it can preemptively copy the rank contents over before the affected rank may experience multi-device errors which would typically be detected as uncorrected error and results in fatal machine check exception error (MCERR). For an illustrative example of rank sparing configuration within a DDR channel and the copying process when a failing rank is activated, please refer to Figure 8.

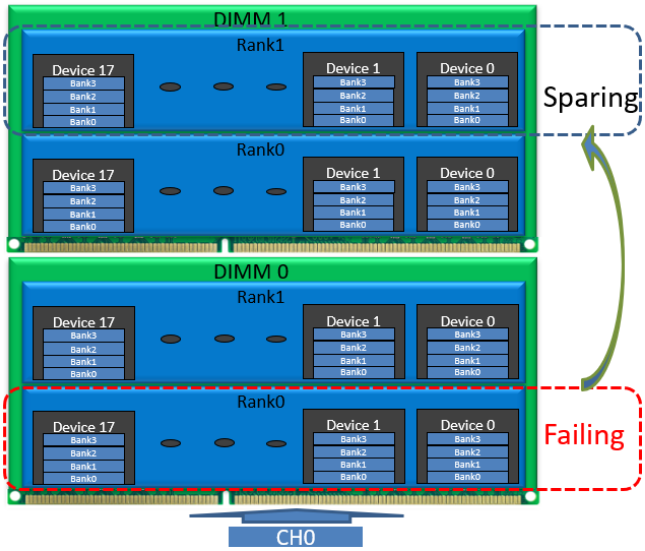


Figure 8: Rank Sparing

The initial generation of Intel Xeon® Processors, known as Skylake, as well as the second generation, referred to as Cascade Lake, offer Rank Sparing capabilities. These processors allow for the allocation of up to two spare ranks out of a maximum of eight ranks within a channel. If one rank is assigned as spare rank, then this rank is reserved and cannot be utilized within the primary system's memory space. Consequently, it remains hidden from the Operating System, resulting in a reduction in the total available memory capacity due to the Rank Sparing

configuration. Therefore, users have the option to enable Rank Sparing on ThinkSystem, which reserves one spare rank for each channel. Refer to the configuration settings in the setup page, as shown in Figure 9 below:

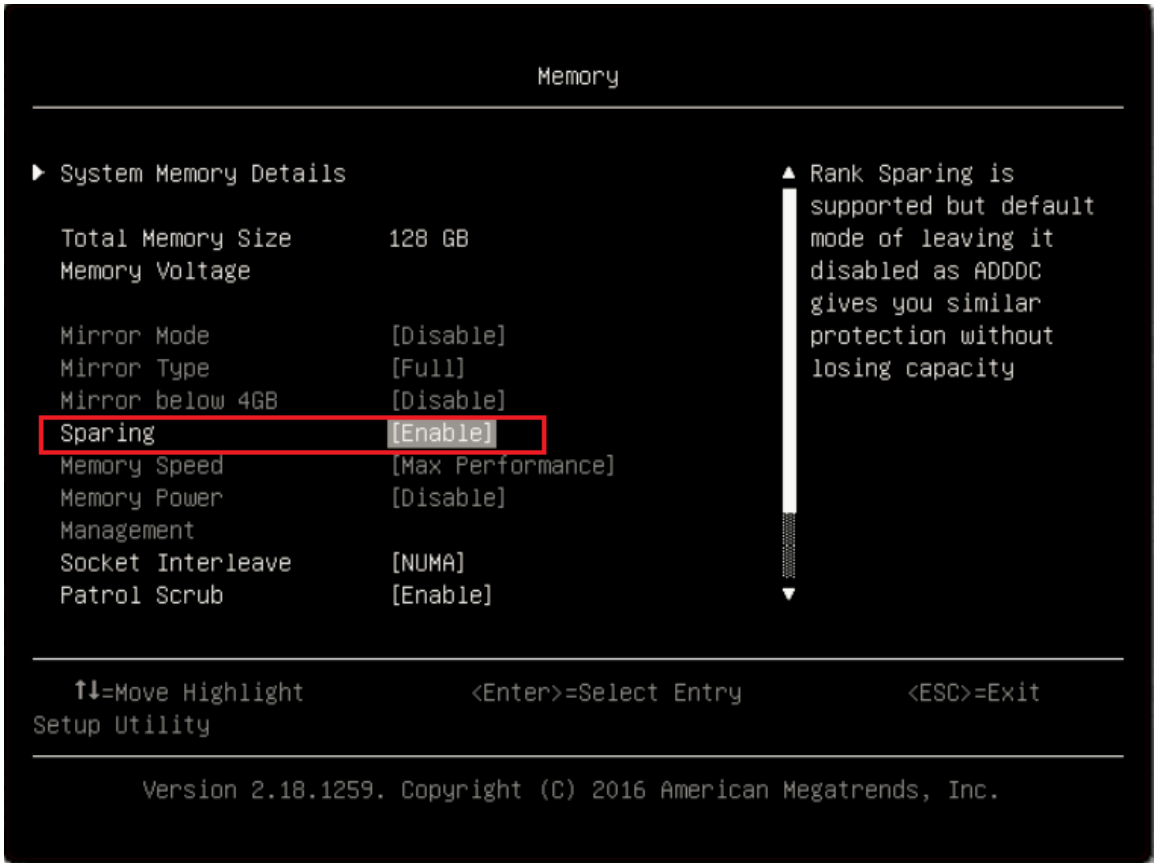


Figure 9: Rank Sparing Configuration in UEFI Setup

When the Rank Sparing feature is activated, ThinkSystem UEFI employs a predictive approach to assess the device's health status based on correctable errors and hardware leaky bucket algorithm. When the relevant threshold is reached, UEFI initiates the Rank Sparing process to transfer data from the failing rank to the spare rank. It then generates SEL entry “FQXSFMA0014I” to inform the customer that the Rank Sparing migration has commenced, and then OS workload can be resumed. Once the Rank Sparing migration is completed, another System Management Interrupt (SMI) is triggered, and UEFI generates SEL entry "FQXSFMA0015I" to notify the customer that the Rank Sparing migration has concluded successfully, with all data successfully transferred from the failing rank to the spare rank.

The SEL events associated with DIMM rank sparing are detailed as follows:

Event ID	Severity	Message String	Notes
FQXSFMA0014I	Info	Memory spare copy initiated. [arg1]	[arg1]: DIMM identifier

FQXSFMA0015I	Info	Memory spare copy has completed successfully. [arg1]	[arg1]: DIMM identifier

ADC-SR and ADDDC-MR

ThinkSystem servers configured with Intel Xeon® based processors support both single and double DRAM device correction capabilities, achieved through adaptive virtual lockstep. The operation occurs within a single DDR channel and is observable during system runtime.

Intel's Xeon® Bronze and Silver processors provide support for the Adaptive Data Correction - Single Region (ADC-SR) feature. ADC-SR feature only supports one region of adaptive virtual lockstep in one channel, with this region being at the bank level.

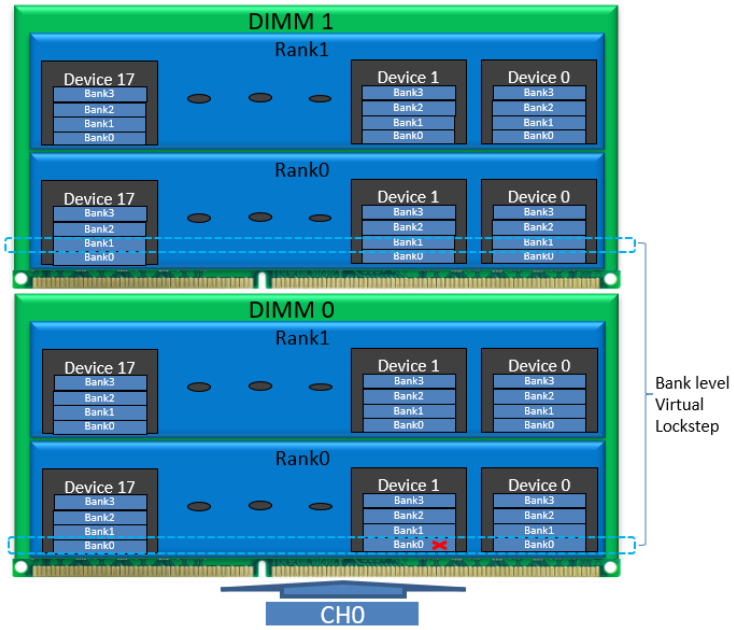


Figure 10: ADC-SR

Figure 10 provides an illustration of a scenario featuring two identical Dual Rank DIMMs within one same DDR channel. As the system continues to operate, Device1 of DIMM0 Rank0 begins to exhibit issues in Bank0, resulting in some errors. The ThinkSystem UEFI SMI handler actively monitors the health of DRAM devices. Once UEFI detects a failing DRAM device, it proceeds to configure the Intel Xeon® CPU to execute device mapping using an algorithm referred to as adaptive virtual lockstep.

In the context of Figure 10, Bank0 of DIMM0 Rank0 is paired virtually with Bank1 of DIMM1 Rank0 through virtual lockstep. This virtual lockstep mechanism involves the hardware reorganizing cache lines stored within the DRAM device, ideally distributing them across two logical banks. This data reconfiguration enables the hardware to transfer data from the failing device to a spare device as a benefit of the virtual lockstep. Intel's Xeon® Gold and Platinum processors support Adaptive Double Device Data Correction - Multiple Regions (ADDDC-MR)

feature. ADDDC-MR facilitates the implementation of two regions of virtual lockstep, with each region capable of operating at either the bank level or the rank level.

Figure 11 serves as an example featuring two identical Dual Rank DIMMs within one same DDR channel. Similar to the ADC-SR feature, the ThinkSystem UEFI SMI handler orchestrates the configuration of the Intel Xeon® CPU to execute bank-level virtual lockstep. In this specific case, it addresses the failing Bank1 of DIMM1 Rank1 Device0 by establishing virtual lockstep between Bank0 of DIMM0 Rank1 and Bank1 of DIMM1 Rank1.

At another instance, the UEFI SMI handler intervenes again, this time configuring the CPU to implement rank-level virtual lockstep. This action resolves multiple failing banks (Bank0 and Bank2 of Device1) within DIMM0 Rank0 by coordinating virtual lockstep between DIMM0 Rank0 and DIMM1 Rank0.

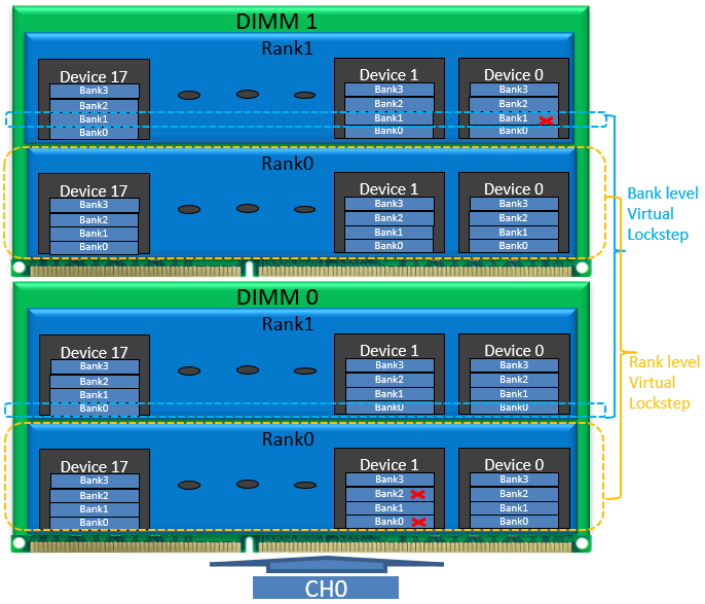


Figure 11: ADDDC-MR

Both ADC-SR and ADDDC-MR come with specific prerequisites for DIMM configuration within ThinkSystem. These requirements include the use of x4 ECC DIMMs for ThinkSystem, as well as the following conditions:

1. If only one DIMM is present in a channel, it should be a Dual or Quad Rank DIMM.
2. If two DIMMs are installed in a channel, both of them must have the same rank size.

Also, it is important to note ADC-SR or ADDDC-MR cannot be activated concurrently with the Rank Sparing or Mirroring features.

ThinkSystem UEFI incorporates a single setup option titled "System Settings -> Memory -> ADDDC Sparing," which encompasses both ADC-SR and ADDDC-MR functionalities. The

activation of ADC-SR or ADDDC-MR is determined by the system's CPU capabilities. By default, the ADDDC Sparing feature is disabled, allowing users to enable or disable the ADDDC feature through the setup page, as depicted in Figure 12 below:

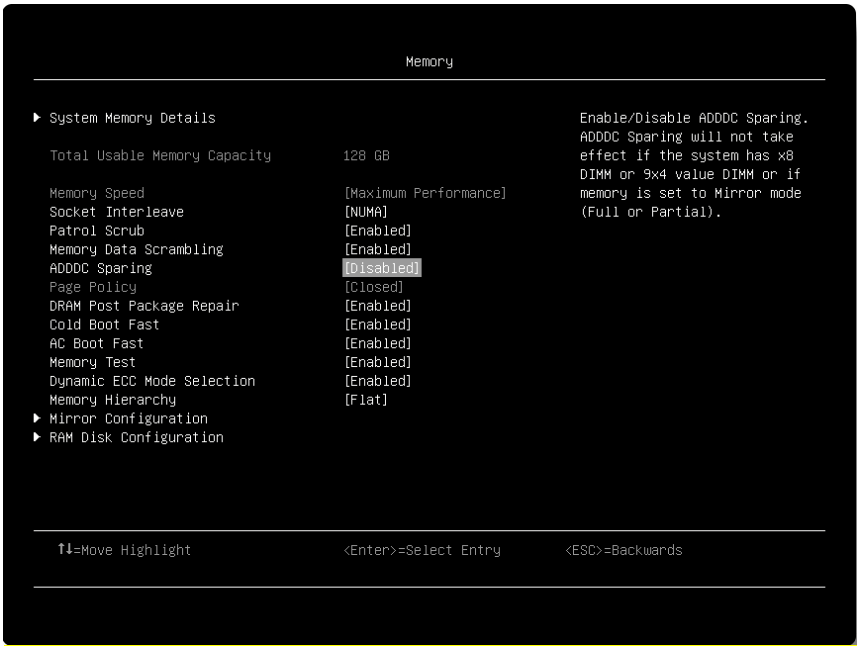


Figure 12: ADDDC Sparing Item in UEFI Setup

Upon enabling the ADDDC Sparing feature, ThinkSystem UEFI initiates a process that involves predicting the health status of devices based on correctable errors and hardware leaky bucket data. Subsequently, it configures Intel CPU to execute the virtual lockstep region specified by ADC-SR or the first and second virtual lockstep regions defined by ADDDC-MR.

In certain scenarios, ThinkSystem UEFI may execute more intricate procedures, including the reversal of two bank-level virtual lockstep regions, followed by a conversion to a single rank-level virtual lockstep region, as necessitated by ADDDC-MR.

Memory Full Mirroring and Partial Mirroring

Memory mirroring feature has more complex address decoding, this chapter will introduce some basic knowledge of how Intel Xeon® processor decodes system addresses, and then explain more details of memory mirroring address decoding. Intel Xeon® processors employ four distinct types of system address decoders, each of these four decoder types serves a unique role in decoding system addresses for different regions and directing them to their respective components:

- Legacy Decoder: This decoder is responsible for decoding specific legacy system address regions, such as the flash region, VGA region, and IO APIC.
- MMCFG Decoder: The MMCFG decoder is utilized to decode the system address region which is used for Memory Mapped PCIe Configuration space.

- MMIO-L / MMIO-H Decoders: These decoders are responsible for decoding the system address regions which are used for Memory Mapped IO Low region (below 4GB) and the Memory Mapped IO High region (above the highest memory region).
- DRAM Decoders: DRAM Decoders are tasked with decoding all memory regions.

Each Intel Xeon® processor incorporates multiple DRAM decoders, capable of managing various levels of interleaves. These interleaves can range from the socket level to the integrated Memory Controller (iMC) level, DDR channel level, DIMM level, and rank level.

For instance, consider Figure 13 as an example. One Intel Xeon® processor contains two integrated Memory Controllers (iMC), with each iMC featuring two DDR channels. This results in four channels: channel A, B, C, and D, each equipped with the same DDR DIMMs. The entire socket level memory space spanning channel A, B, C, and D can be interleaved at the channel level. This interleaving arranges four different color-coded blocks sequentially to form the complete socket level memory space.

However, due to the presence of legacy regions (such as the flash region, VGA, IO APIC, etc.), the MMFG region, and the MMIO low region within the system space below 4GB, the memory space needs to be divided into two regions. SAD0 (System Address Decoder entry 0) manages the decoding of the memory region below 4GB (excluding the memory hole), while SAD1 oversees the decoding of the memory region above 4GB.

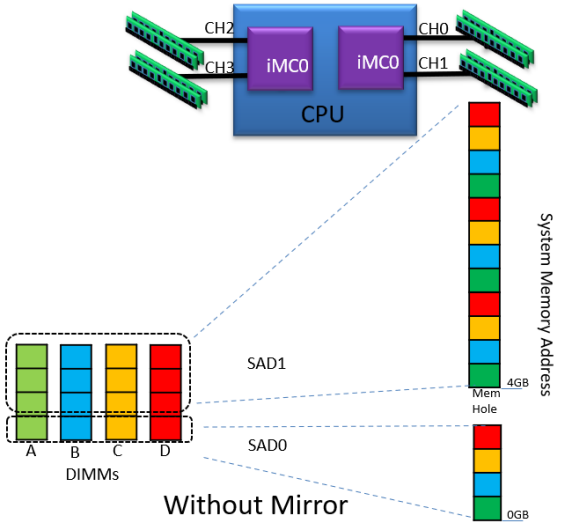


Figure 13: System Memory Address without Mirror

Intel Xeon® processors offer support for the data Memory Mirroring feature, which involves splitting memory into two channels and duplicating the contents of the first channel onto the second channel. This duplication creates a redundant copy of memory. In the event of a memory error occurring in one channel, the memory controller seamlessly switches to the paired channel, ensuring continuous system operation without any downtime.

Intel Xeon® processors provide support for both RAS features Full Mirroring and Address Range/Partial Mirroring (note: in some old documents Intel used partial mirroring to indicate only CPU of the system enables mirroring but other CPUs of the system do not enable mirroring; or only some channels of one CPU enable mirrors; but this kind of partial mirror is out date, normally Intel Xeon® processors only support Address Range Mirroring now, and also just treats Partial Mirroring same as Address Range Mirroring.) In Full Mirroring, a duplicate (secondary or mirrored) copy of memory contents is maintained for all memory regions. Address Range/Partial Mirroring takes this a step further by allowing the firmware or Operating System to specify a specific range of memory addresses for mirroring while leaving the remaining memory in the socket in non-mirrored mode. This offers increased granularity for mirroring within a socket at partial address range level.

Using the memory configuration illustrated in Figure 13 as a reference, when the system is set up with full memory mirroring, Figure 14 below demonstrates the full mirroring scheme on the left side. In this scenario, the green blocks in channel A represent the primary block, while the corresponding green blocks in channel B serve as the secondary (backup) for the primary green blocks. Similarly, the blue blocks in channel B are designated as the primary blocks, with the corresponding blue blocks in channel A serving as its secondary (backup). Under this configuration, SAD0 is responsible for decoding the memory mirror region below 4GB, while SAD1 handles the decoding of the memory mirror region above 4GB.

On the right side of Figure 14, the partial mirroring configuration is displayed. Here, SAD0 and SAD1 are responsible for decoding the non-mirrored regions, while SAD2 is tasked with decoding the specific region designated for partial mirroring.

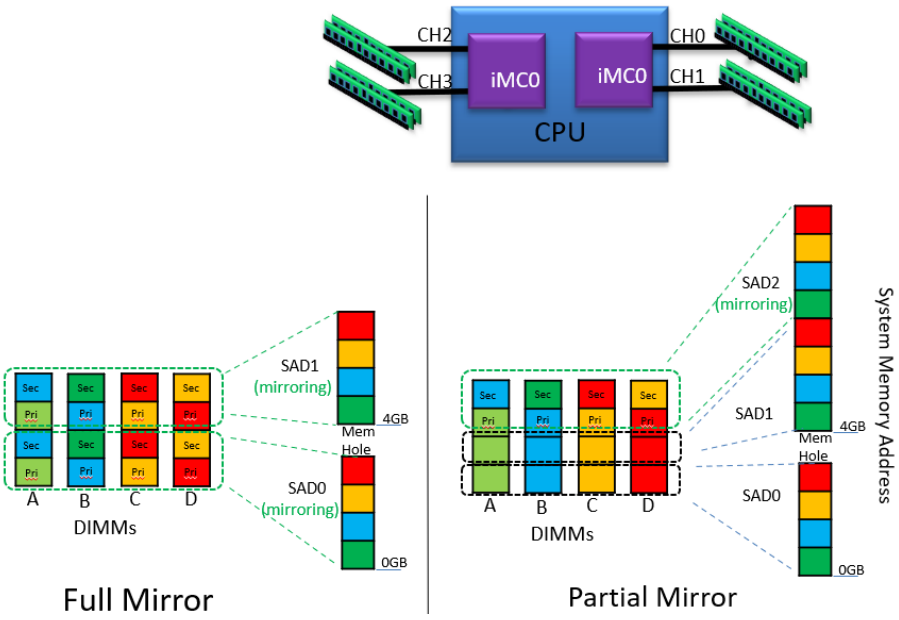


Figure 14: System Memory Address without Mirror

ThinkSystem UEFI has a dedicated setup page for configuring memory mirror related RAS features and can be found under “System Settings -> Memory ->Mirror Configuration.” This page encompasses both Full Mirroring (as shown in Figure 15) and Partial Mirroring (as shown in Figure 16) configurations. For a more comprehensive understanding of memory mirroring configurations available on ThinkSystem servers, please refer to the Lenovo white paper linked in the References section of this document.

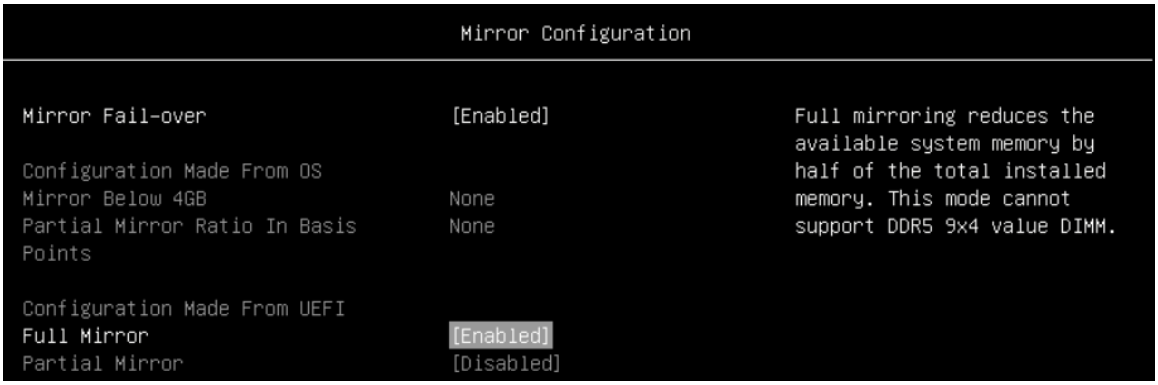


Figure 15: Full Mirror Configuration in Setup

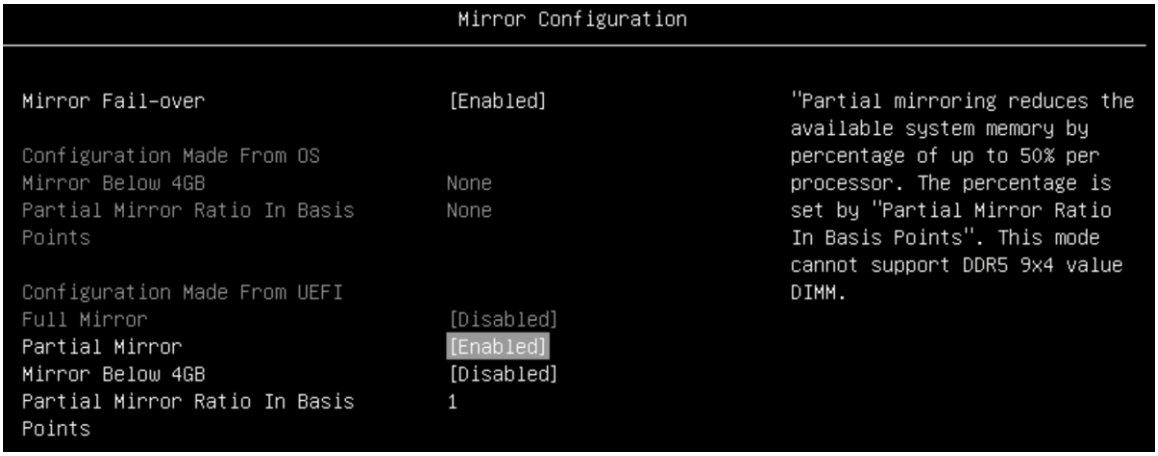


Figure 16: Partial Mirror Configuration in Setup

In Full Mirroring or Partial Mirroring mode, when a memory-read returns an error, it can be corrected through various flows based on whether the error is correctable or uncorrectable as follows:

- If memory-read operation yields in a correctable error in the primary channel, that error will be corrected through ECC. Then the memory controller writes the corrected data back to the original read location directly, without impact to secondary channel.
- If memory-read operation yields in an uncorrectable error within the primary channel, the memory controller issues a request to switch to the secondary channel. If the secondary channel successfully retrieves the data, the memory controller then writes the data back to the primary channel. Following the data correction process, the memory controller performs a re-read to ensure that the newly written data is free from

errors. If the error persists despite these efforts, then the memory controller triggers a mirror fail-over event.

Upon the activation of a mirror fail-over event, the ThinkSystem UEFI SMI handler takes charge of overseeing the entire fail-over process. Once this process is successfully completed, SEL entry "FQXSFMA0013I" is generated as a notification. It is important to note that following the resolution of the mirror fail-over, the mirroring functionality ceases to exist. The primary channel is deactivated, while the secondary channel continues to operate in a standard non-mirrored configuration.

Beginning with the ThinkSystem 3rd generation servers powered by Intel's 4th generation Xeon® Processors (formerly known as "Sapphire Rapids"), Lenovo offers a setup feature known as Mirror Fail-over, as depicted in Figure 17 below. By default, this feature is set to "Enabled," which means that in the event of a persistent memory uncorrectable error, a mirror fail-over is triggered. In such cases, UEFI reports SEL entry "FQXSFMA0054G" as a notification to end user; if mirror fail-over is not finished successfully by processor, then UEFI reports SEL entry "FQXSFMA0055G" instead.

However, users have the option to set Mirror Fail-over to "Disabled" if customer's Operating System (OS) supports memory page retire request from UEFI. When this setting is chosen, Lenovo UEFI will bypass the mirror fail-over function even when a persistent uncorrectable error occurs. Instead, it reports the affected memory page to the OS, allowing the OS to take the problematic page offline to prevent any further access to the persistent uncorrectable error. SEL entry "FQXSFMA0055G" (Mirror failover operation was unsuccessful. DIMM [arg1] cannot fail over again.[arg2]) is generated in this scenario (note: actually in this scenario, mirror fail-over is not applied at all, so will change to use another new SEL event ID and string for this scenario in future UEFI release).

This flexibility enables users to let the OS handle memory addresses containing persistent uncorrectable errors while still retaining the mirror-feature for the mirrored memory range.

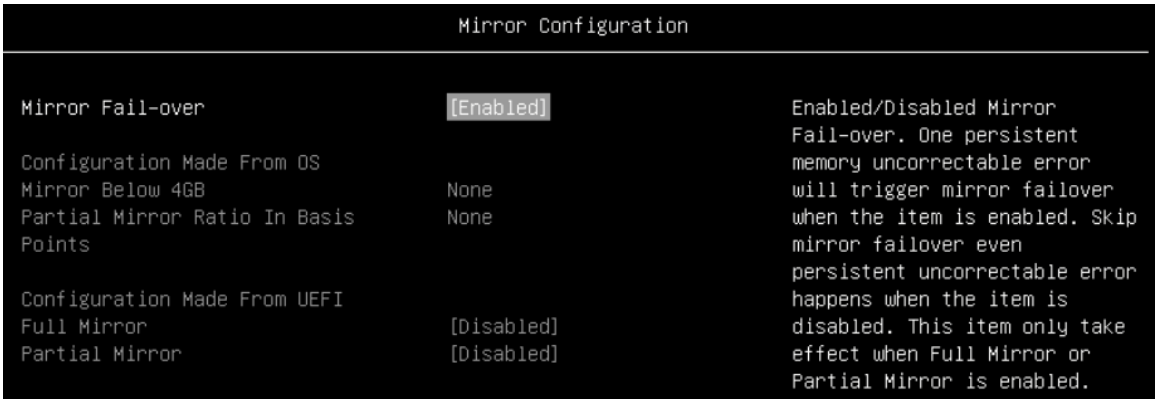


Figure 17: Mirror Fail-over Configuration in Setup

Memory mirror related SEL event is listed as below:

Event ID	Severity	Message String	Notes
FQXSFMA0013I	Info	Mirror Fail-over complete. DIMM [arg1] has failed over to the mirrored copy. [arg2]	[arg1]: DIMM silk label [arg2]: DIMM identifier
FQXSFMA0053G	Warning	An uncorrected memory error has been recovered by mirror on DIMM [arg1] at address [arg2].[arg3]	[arg1]: DIMM silk label [arg2]: Error address [arg3]: DIMM identifier
FQXSFMA0054G	Warning	Mirror failover operation was successful. DIMM [arg1] has failed over to the mirrored DIMM [arg2].[arg3]	[arg1]: DIMM silk label [arg2]: DIMM silk label [arg3]: DIMM identifier
FQXSFMA0055G	Warning	Mirror failover operation was unsuccessful. DIMM [arg1] cannot fail over again.[arg2]	[arg1]: DIMM silk label [arg2]: DIMM identifier

Partial Cache-line Sparing (PCLS)

ThinkSystem 2nd generation servers running Intel's 3rd Generation Xeon® Processors (codename "Ice Lake-SP" and "Cooper Lake") are equipped with Intel's Partial Cache-line Sparing (PCLS) feature, designed to enhance system uptime in the event of a persistent single-bit memory errors. This functionality enables the identification and mapping out of failed single bits by utilizing the available nibble-level spare capacity within each DDR channel of the Xeon® processor's integrated Memory Controllers (iMC)."

When ThinkSystem UEFI identifies the occurrence of single-bit errors within a specific cache line of a DIMM, and these errors surpass the defined threshold during runtime, UEFI will proceed to record the precise address and nibble location of the single-bit error in the CFG_DATA_INFO of a designated PCLS entry. Subsequently, the integrated Memory Controller (iMC) will initiate nibble sparing for cache line read/write operations that match the address specified in the PCLS entry. For a more detailed understanding of this process, please refer to Figure 18 below.

It is important to note that this feature will not be activated in channels where ADC-SR or ADDDC-MR is in use. Additionally, this feature is available on all 3rd Xeon® Scalable Processors and is enabled by default, eliminating the need for making changes to system setup configuration.

High-Level Overview of Runtime Code Logic:

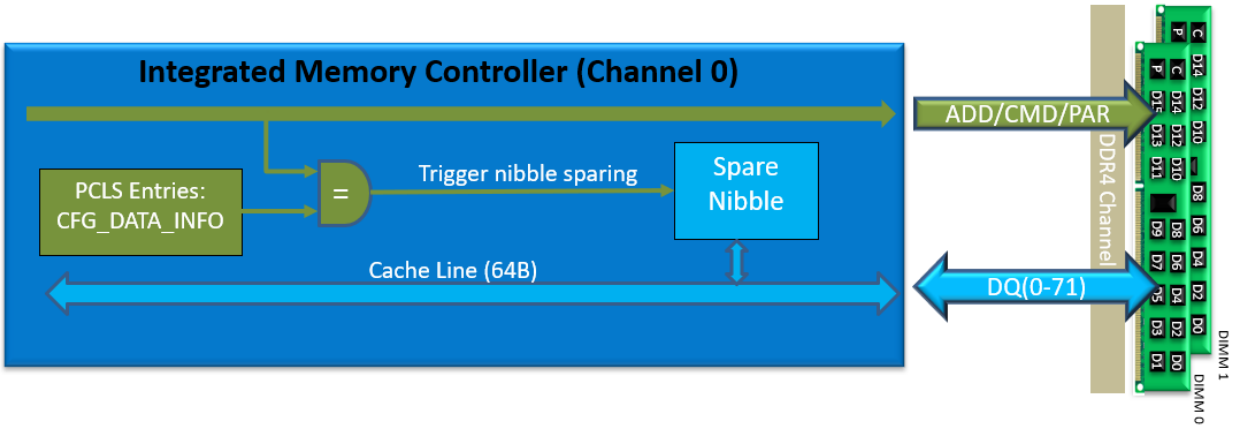


Figure 18: PCLS Process in integrated Memory Controller

Lenovo Value-added Page Retire PFA

The Memory Page (4KB) Retire Operation utilizes the ACPI feature integrated into UEFI to establish communication with the Operating System (OS). This communication serves to signal the OS when a page has accrued a significant number of correctable memory errors, prompting the OS to cease using that particular page. This action effectively retires or takes the page offline within the OS when memory is not locked by the critical process or application. Upon successfull Page Retire opeartion any future errors linked to this page will no longer be reported in system hardware log, though the OS logs will retain those log entries.

Once the OS confirms the successful retirement of a specified memory page range to the platform, it assures that the retired memory range has been flushed and invalidated for all caches and all tables in all caching agents. These agents include I/O device caches, processor caches, TLBs, and any chipset tables or caches such as VT-D or IOMMU. Consequently, the memory page will not be accessed by the OS for the remainder of the boot session.

To support the page-retire mechanism, Lenovo ThinkSystem UEFI establishes a dedicated Generic Hardware Error Source Structure (GHESS) that is designated for page-retire requests, that is activated even before the OS has booted. The GHESS is registered in the Hardware Error Source Table (HEST) as described by the ACPI specification.

When the number of correctable errors on a single memory page surpasses the defined threshold, UEFI initiates a page-retire request to the OS. This process entails UEFI employing the SCI (System Control Interrupt) method to notify the OS Power Management (OSPM). Furthermore, UEFI communicates the specific failing page address by means of the GHESS table. Consequently, if the OS is equipped to handle page-retire requests from UEFI, it proceeds to retire or take the failing memory page offline.

Many Operating System versions, including Red Hat (kernel 3.12 or later) and VMware (ESXi 7.0U3/7.0.3 or later), offer support for page retire requests originating from UEFI. For specific information about OS compatibility and page retire support, it is recommended to reach out to your respective OS vendors.

Within ThinkSystem UEFI, certain thresholds are monitored for Predicted Failure Analysis (PFA) associated with page retire requests. In the event that multiple page retire requests surpass the established Page Retire PFA thresholds, UEFI will generate a Page Retire PFA SEL (System Event Log) event, specifically identified as FQXSFMA0012L, for the affected DIMM; and then Lenovo LXCI will use VMWare Proactive Hardware Management to trigger VM migration, to avoid the potential VM crushing. See more details flow in below Figure 19:

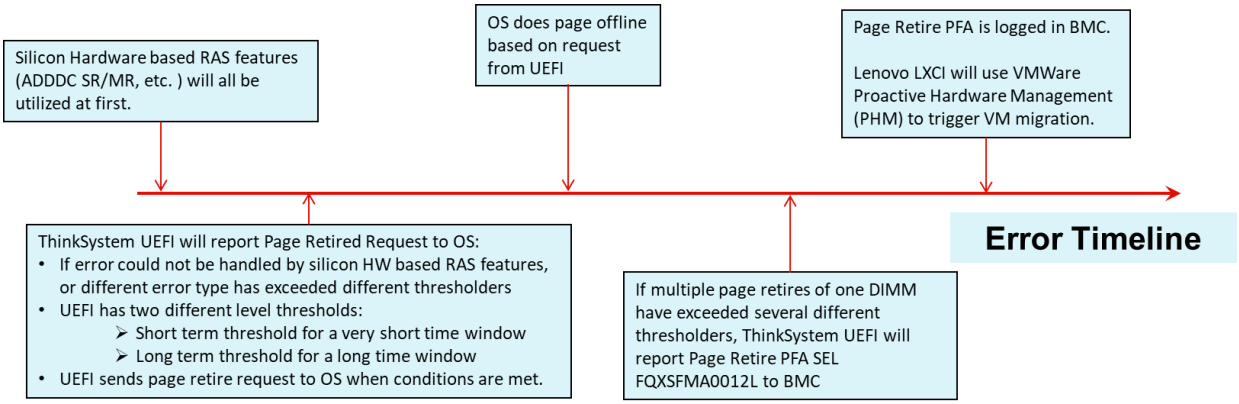


Figure 19: Page Retire PFA Flow

Listed below is the SEL event related to Page Retire PFA:

Event ID	Severity	Message String	Notes
FQXSFMA0012L	Warning	The [arg1] PFA Threshold limit has been exceeded on DIMM [arg2] at address [arg3]. [arg4]	[arg1]: high/low [arg2]: DIMM silk label [arg3]: Address where error occurred [arg4]: DIMM identifier

Memory Uncorrectable Error Handling during System Runtime

When an uncorrectable error occurs during the operation of the Operating System, the Intel Xeon® processor categorizes the error based on the error scenario into two main types: DUE (Detected but Uncorrectable Error) and UCR (Un-Corrected Recoverable).

For DUE errors, which include catastrophic and fatal errors, the processor immediately triggers IERR (Internal Error) and results in a system hang. In such cases, no System Management Interrupt (SMI) handler is executed to address the error. Instead, Lenovo's Fatal Error Handling (FEH) feature, implemented at the Baseboard Management Controller (BMC), takes charge. It collects error data through out-of-band interfaces, isolates the error, and generates System Event Log (SEL) entries accordingly.

Un-Correctable Recoverable Errors (UCR) are further classified into three sub-categories: UCNA (Uncorrected No Action), SRAO (Software Recoverable Action Optional), and SRAR (Software Recoverable Action Required). The specifics of this error classification are illustrated in the Figure 20 below.

Note: Intel Xeon® processors (from 1st generation to 4th generation) silicon and micro-code has reverted SRAO error type to UCNA.

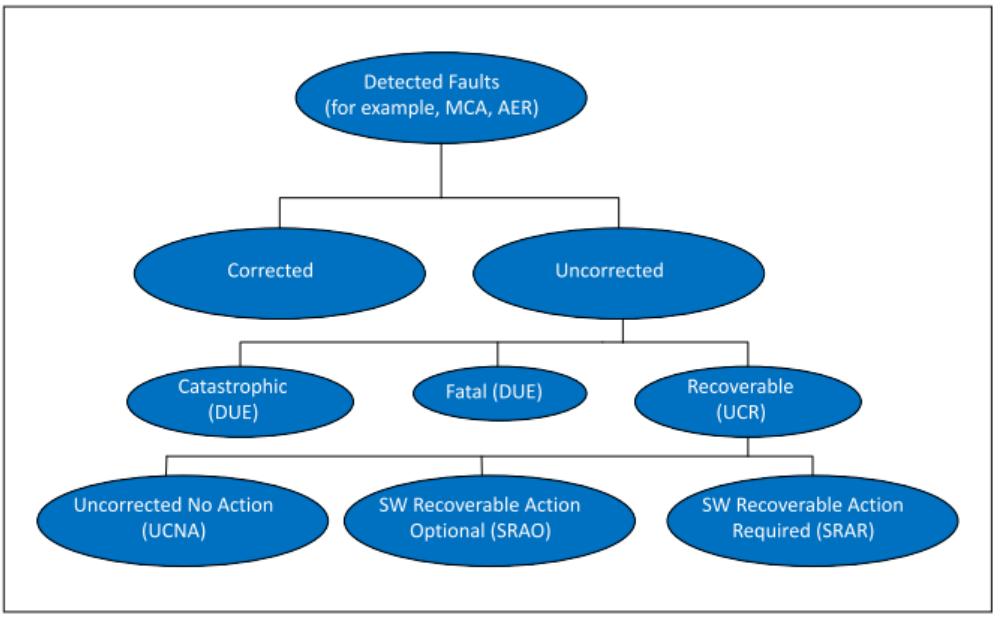


Figure 20: Error Classification in Intel based ThinkSystem

ThinkSystem UEFI has been configured to initiate System Management Interrupts (SMIs) for Uncorrected Recoverable Errors (UCR) falling under the UCNA and SRAR categories. The SMI handler is responsible for processing UCNA errors with a correctable error procedure, while it addresses SRAR errors differently.

In the case of SRAR errors, the UEFI SMI handler conducts a check on the machine check bank registers to determine the recoverability of the error. If the error is deemed recoverable, UEFI reports SEL FQXSFMA0056M. Otherwise, it reports FQXSFMA0002M. Following this, control is returned to the Operating System (OS), which proceeds with the recovery process. This may involve terminating the application responsible for the uncorrectable error or triggering a kernel panic.

SEL entries related to memory uncorrectable errors are listed below:

Event ID	Severity	Message String	Notes
FQXSFMA0002M	Error	An uncorrectable memory error has been detected on DIMM [arg1] at address [arg2]. [arg3]	[arg1]: DIMM silk label [arg2]: Memory error address [arg3]: DIMM identifier

FQXSFMA0056M	Error	An uncorrected recoverable memory error has been detected on DIMM [arg1] at address [arg2].[arg3]	[arg1]: DIMM silk label [arg2]: Memory error address [arg3]: DIMM identifier

References

For more information, see the following documents:

– Address Range Partial Memory Mirroring:

https://www.intel.com/content/www/us/en/developer/articles/technical/address-range-partial-memory-mirroring.html

– ThinkSystem Advanced Memory Test Guidance:

https://download.lenovo.com/servers_pdf/lenovo-amt-on-thinksystem.pdf

– ThinkSystem memory mirroring configuration:

https://download.lenovo.com/servers_pdf/thinksystem_memory_mirroring_config.pdf

LENOVO PROVIDES THIS PUBLICATION “AS IS” WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. This information could include technical inaccuracies or typographical errors. Changes may be made to the information herein; these changes will be incorporated in new editions of the publication. Lenovo may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any performance data contained herein was determined in a controlled environment; therefore, the results obtained in other operating environments may vary significantly. Some measurements may have been made on development-level systems, and there is no guarantee that these measurements will be the same on generally available systems. Furthermore, some measurements may have been estimated through extrapolation. Actual results may vary. Users of this document should verify the applicable data for their specific environment.

Any references in this publication to non-Lenovo Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this Lenovo product, and use of those Web sites is at your own risk.

Lenovo, the Lenovo logo, and ThinkServer, are trademarks of Lenovo in the United States, other countries, or both.

Red Hat is a trademark of Red Hat, Inc. in the United States, other countries, or both.

VMware Server is a trademark of VMware, Inc. in the United States, other countries, or both.