# ThinkSystem Memory Mirroring Configuration Guidelines

**December 2023**

Jason(Zhijun) Liu and Fuzhou Liu

# Abstract

Memory Mirroring is a Reliability Availability Serviceability (RAS) feature of Intel® Xeon® processor family-based platforms. It is a method of duplicating memory between two channels and mirroring the first channel to the second one, creating a redundant copy of memory. In the case of any persistent uncorrectable error in the memory of one channel, the memory controller shifts to the paired channel without causing system downtime. This whitepaper outlines Lenovo's value-added configurable mirroring fail-over feature and optimized Full Mirroring and Partial Mirroring (Address Range Mirroring) features. It explains how they are configured in the Unified Extensible Firmware Interface (UEFI) setup page and operating system of Lenovo ThinkSystem and how operating system level configuration is prioritized over the other.

# Contents

Abstract....2
Introduction to Intel Memory Mirroring....4
Mirroring Correction and Mirroring Fail-over....4
Mirroring configuration in ThinkSystem UEFI setup page....5
Mirroring configuration in operating system for ThinkSystem....10
Mirroring configuration priority between UEFI setup page and operating system....12
References....15

# Introduction to Intel Memory Mirroring

The Intel® Xeon® processor family-based platform offers a Reliability Availability Serviceability (RAS) feature called Memory Mirroring.

Memory Mirroring is a method of keeping a duplicate (secondary or mirrored) copy of the contents of memory as a redundant backup for use if the primary memory fails. Thus, memory mirroring configures the memory in a highly reliable mode so that the system uptime can be maintained even when a memory component is affected by an uncorrectable fault. On the other hand, total available memory is reduced in the mirroring mode as all the memory at the slave is now used as duplicate memory.

Memory mirroring is implemented in the hardware and firmware with the option for the end user to enable the system operation in mirroring mode. The operating system can use the mirrored region for critical code/data. The Intel 1st Generation Xeon Scalable Processor (codenamed Skylake) family-based platform introduces a new layer of support for partial memory mirroring called Address Range Mirroring or Partial Mirroring. Intel 1st or later Generation Xeon Bronze and Silver Processors only support Full Mirroring, but Platinum and Gold Level Processors support both Full and Partial Mirroring.

When partial mirroring is used, the platform allows an operating system to specify a subset of total available memory for mirroring (and to optionally specify whether to mirror the range 0–4 GB). This capability allows a user to make an appropriate trade-off between non-mirrored memory range and mirrored memory range, thus optimizing total available memory while keeping highly reliable memory range available for mission-critical workloads and/or kernel space.

# Mirroring Correction and Mirroring Fail-over

When memory is configured in mirrored mode, all memory-write operations will be performed in both primary and secondary target locations. This guarantees that the mirrored target is always updated with the latest data from the main copy. Dual In-line Memory Modules (DIMMs) populated in a mirror-paired channel need to have identical size of each rank and at least two Double Data Rate (DDR) channels must be populated in each memory controller.

In 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 returns a correctable error in the primary channel, that error will be corrected directly through Error Correction Code (ECC). Then the memory controller writes the good data back to the location of the original read.
- If memory-read returns an uncorrectable error in the primary channel, the memory controller issues a request to shift to the secondary memory channel. If shifting to the secondary memory channel succeeds, the memory controller writes the data back to the primary memory channel. After the corrected data is written, the memory controller re-

reads it to confirm the newly written data does not have any more error. If the error persists, the memory controller triggers a mirroring fail-over event.

Starting from ThinkSystem 3rd generation (with Intel 4th Generation Xeon® Processors – codenamed Sapphire Rapids), Lenovo provides a setup feature called Mirroring Fail-over. When mirroring fail-over event is triggered, the Unified Extensible Firmware Interface's System Management Interrupt (SMI) handler is used to address the whole process of mirroring fail-over. Once it is addressed, mirroring ceases to exist. The primary channel gets disabled, and the secondary channel continues to work as a normal non-mirror case.

The default setting of "Mirror Fail-over" is "Enabled" which implies, a persistent memory uncorrectable error triggers a mirroring fail-over. If "Mirror Fail-over" is set to "Disabled", then Lenovo UEFI skips the mirroring fail-over function even when a persistent uncorrectable error occurs. However, it reports the impacted memory page to the operating system and lets the operating system take the page offline to avoid any further incurring of the persistent uncorrectable error. Thus, the operating system can be allowed to manage the memory address which has the persistent uncorrectable error while the system continues to retain the mirroring-feature of the mirrored memory range.

"Mirror Fail-over" feature can be accessed and configured at ThinkSystem UEFI setup page -> System Configuration and Boot Management -> System Settings -> Memory->Mirror Configuration, Refer to the below screenshot for a quick reference:

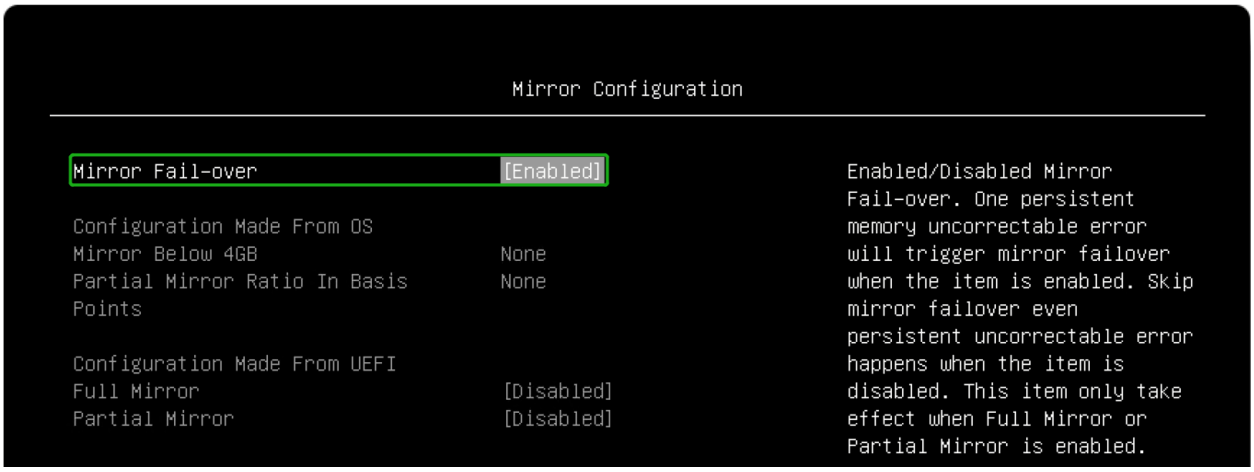



**Figure 1: Mirroring Fail-over setup**

# Mirroring configuration in ThinkSystem UEFI setup page

Intel Xeon processors of 1st (codenamed Skylake) and later generations offer various mirroring feature scopes for different Central Processing Unit categories. Bronze and Silver Processors only support Full Mirroring while Platinum and Gold Processors support both Full and Partial Mirroring.

Memory Mirroring feature has some basic requirements as follows:

1. Full Mirroring configuration is supported by all Xeon processors, while Partial Mirroring is supported only by Platinum and Gold Processors.
2. Mirroring feature is not available for 9x4 Dual In-line Memory Module.
3. If another Reliability Availability Serviceability feature of memory called “Adaptive Double Device Data Correction (ADDDC) Spare” is enabled, the Full Mirroring or Partial Mirroring features get grayed out.

Lenovo ThinkSystem has different UEFI setup options for Full and Partial Mirroring. Both are disabled by default and only one of them can be enabled at a time. The Figure 2 below is a screenshot of the default memory “Mirror Configuration” setup page when Adaptive Double Device Data Correction (ADDDC) Sparing feature is Disabled. As can be seen, either “Full Mirror” or “Partial Mirror” can now be enabled.

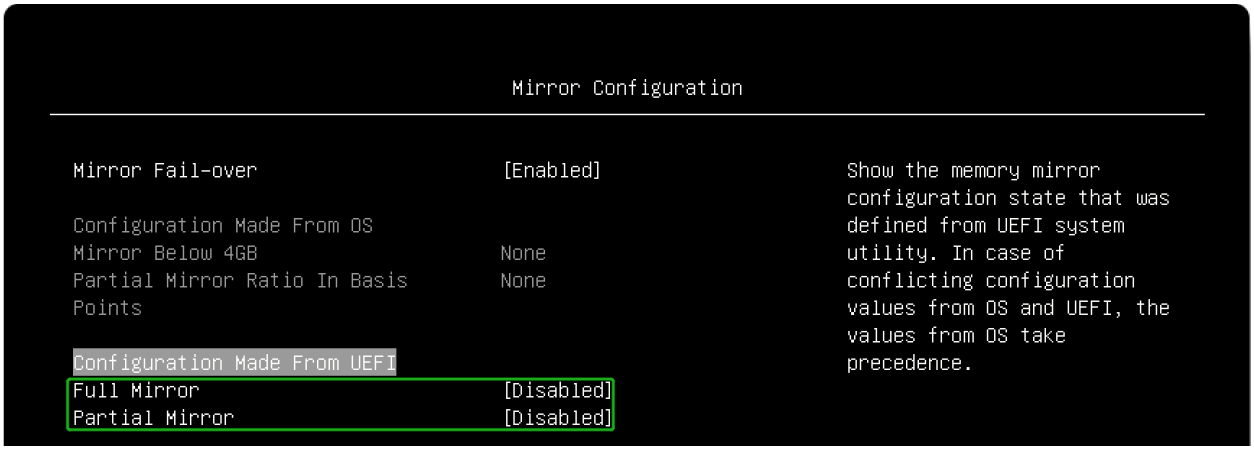



**Figure 2: Default Memory Mirroring Setup Page (ADDDC Sparing Disabled)**

When Adaptive Double Device Data Correction Sparing is enabled, both Full Mirroring and Partial Mirroring setup options get grayed out and mirroring feature cannot be enabled as seen in Figure 3 below:

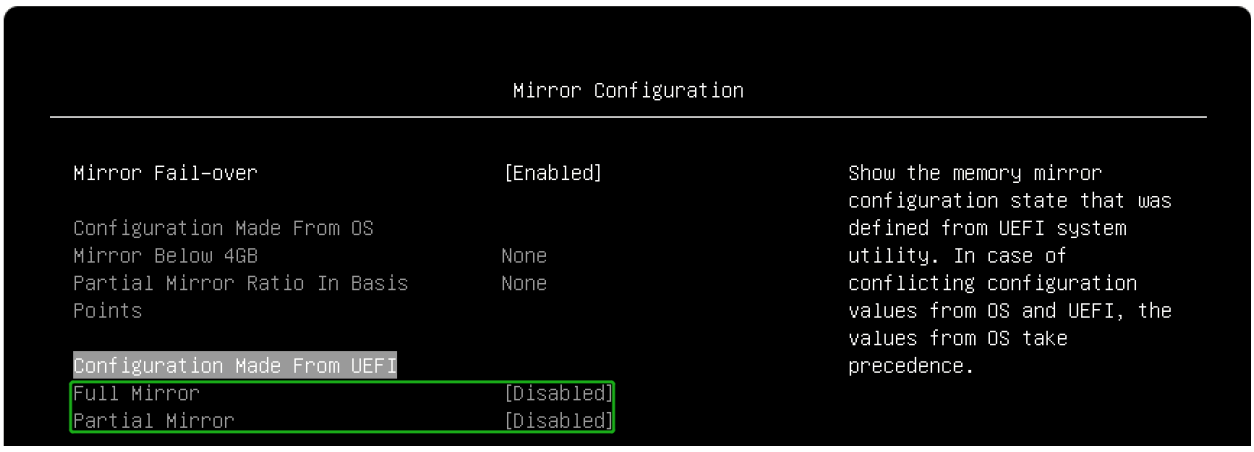



**Figure 3: Default Memory Mirroring Setup Page (ADDDC Sparing Enabled)**

### Full Mirroring Configuration

In full memory mirroring, half of the installed memory is reserved for redundancy. Therefore, only half of the total memory is made available to the operating system.

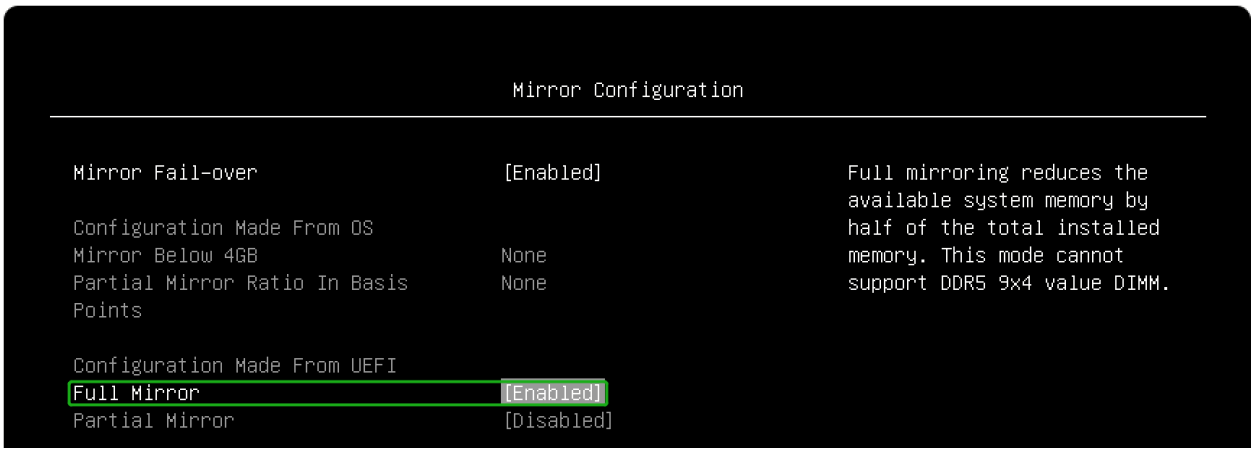



**Figure 4: Full Mirroring enabled in the setup page**

After Full Mirroring is configured, in the subsequent reboot, the ThinkSystem splash screen displays a total of 128GB memory but with a usable capacity of only 64GB for the system as shown below:

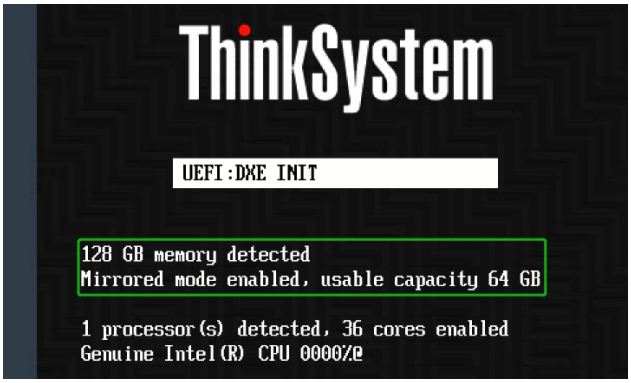



**Figure 5: ThinkSystem splash screen with full mirroring**

### Partial Mirroring Configuration

In this mode, a subset of memory is mirrored leaving the rest of the memory in non-mirrored mode. Partial memory mirroring mode allows mirroring of critical memory address range (also called more reliable memory) without incurring the cost of full memory mirroring.

Partial Mirroring offers the following two options:

- ”Mirror below 4G”: This mirroring option is specifically for the memory range that is below 4GB for the first processor as shown in Figure 6 below.

- ”Partial Mirror Ratio In Basis Points”:  This option is used for percentage setting of memory for each processor.

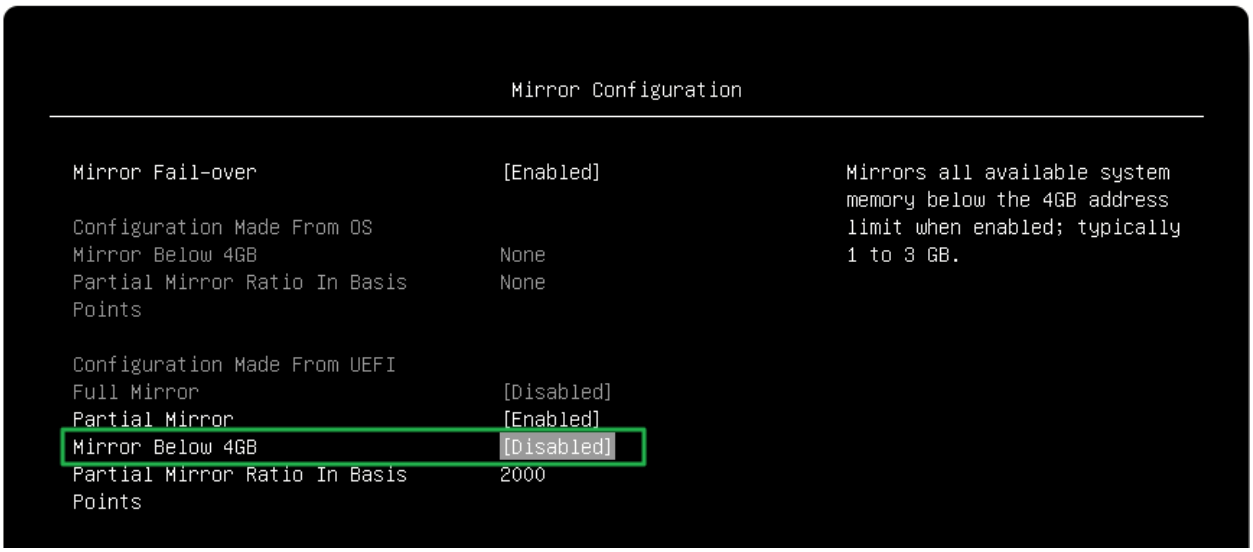



**Figure 6: Mirror Below 4GB in Mirror Configuration Setup Page**

The system address for a memory range that is below 4GB can mostly be split into two parts. One of them is allocated for several different sub-regions such as below 4GB Memory Mapped Input Output (IO) range, PCIe Memory Mapped(MM) Configuration space, some Platform Controller Hub (PCH) special decode regions and so on. The second part of the system address is allocated for memory region.

In the ThinkSystem UEFI setup page, at System Configuration and Boot Management->System Settings->Devices and I/O Ports, there is a setup feature called MM Config Base. If MM Config Base is 2GB, then the system address range of 0~2GB is below 4GB memory range and 2GB~4GB is allocated for different sub-regions. The default value of MM Config Base is “Auto**”** as shown in Figure 7 below. Its value is automatically assigned based on the resource allocation of PCIe devices.

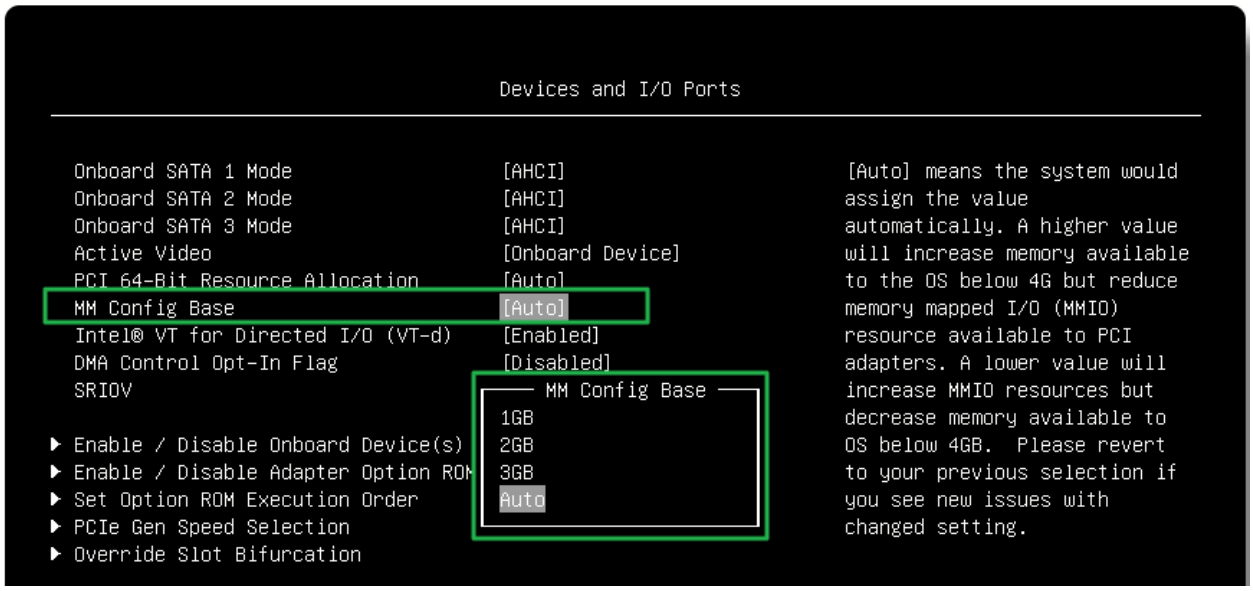



**Figure 7: MM Config Base in Setup Page**

Figure 8 below is a screenshot of the UEFI setup page that shows how to enable partial mirroring at 20% with the system having one processor and the detected memory of 128GB:

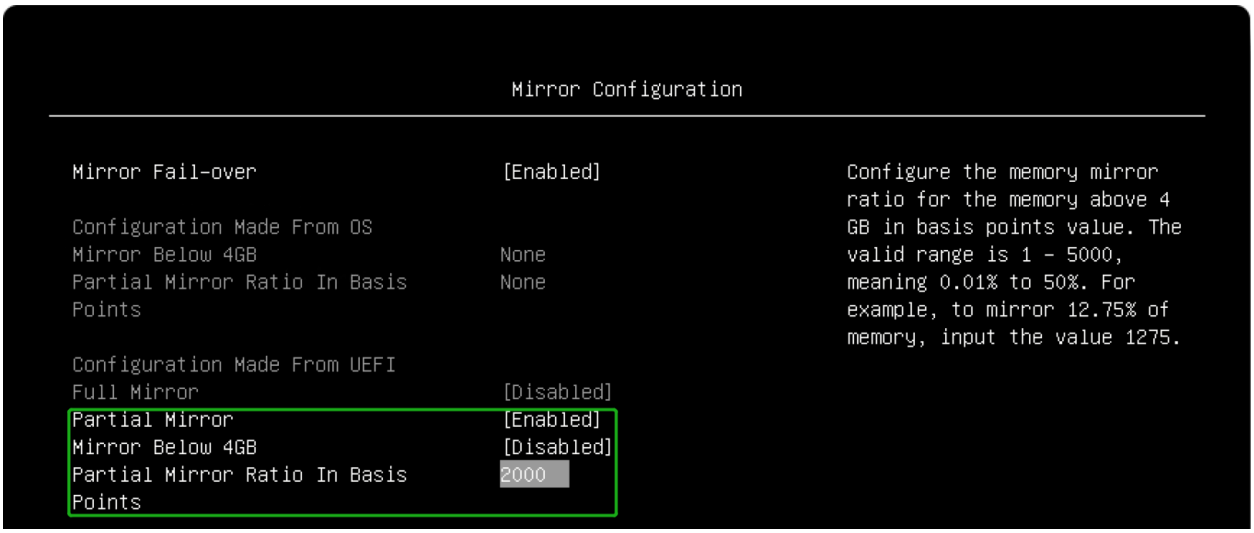



**Figure 8: Enabling Partial Mirroring setup page**

The actual value of partial mirroring ratio is in basis points and one hundredth of its value is the percentage by which partial mirroring is enabled. For example, if the value of the setting is 1275, the partial mirroring percentage is 12.75%. After partial mirroring is configured, in the subsequent reboot, the following ThinkSystem splash screen shows up with 20% partial mirroring configuration. For the system having 128GB of detected memory, 20% of it amounts to 25GB (with 1GB alignment) memory in mirroring. Therefore, the usable capacity for the system is 128GB-25GB = 103GB.

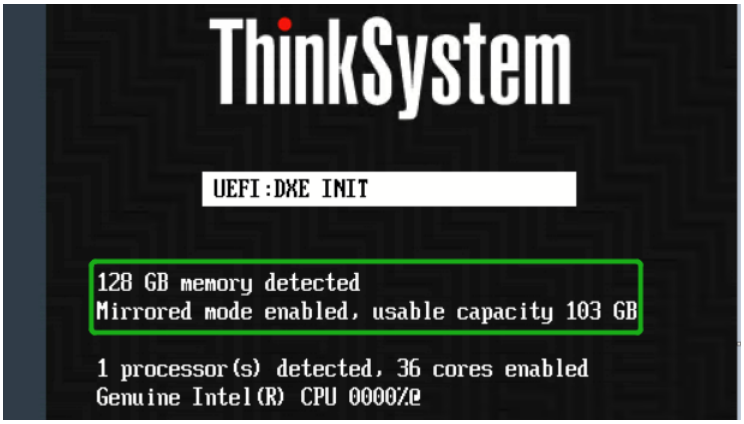



**Figure 9: ThinkSystem splash screen with partial mirroring**

# Mirroring configuration in operating system for ThinkSystem

Legacy full memory mirroring is transparent to the operating system. However, partial memory mirroring requires a Firmware-operating system interface for the user to specify the desired subset of memory for mirroring. UEFI based operating system calls its service GetMemoryMap() to get the UEFI memory-map information. Partial memory ranges that are set to be mirrored are indicated by the Attribute field as: EFI_MEMORY_MORE_RELIABLE in the EFI_MEMORY_DESCRIPTOR.

Following is the Memory Attribute Definition in UEFI specification:

```
//*******************************************************
// Memory Attribute Definitions
//*******************************************************
#define EFI_MEMORY_UC                   0x0000000000000001
#define EFI_MEMORY_WC                   0x0000000000000002
#define EFI_MEMORY_WT                   0x0000000000000004
#define EFI_MEMORY_WB                   0x0000000000000008
#define EFI_MEMORY_UCE                  0x0000000000000010
#define EFI_MEMORY_WP                   0x0000000000001000
#define EFI_MEMORY_RP                   0x0000000000002000
#define EFI_MEMORY_XP                   0x0000000000004000
#define EFI_MEMORY_NV                   0x0000000000008000
#define EFI_MEMORY_MORE_RELIABLE        0x0000000000010000
#define EFI_MEMORY_RO                   0x0000000000020000
#define EFI_MEMORY_RUNTIME              0x8000000000000000
```

In addition to  the memory attribution, there are the UEFI runtime variables that enable the operating system or system software to identify the memory blocks which the platform is mirroring. They are also used to permit the operating system requests for mirroring on a subsequent boot.

There are two UEFI runtime variables with the following names:

1. MirrorCurrent - This variable is written by the BIOS and read by the operating system to communicate the status of address range mirroring on the platform.
2. MirrorRequest - This variable is written by the operating system to request a new mirroring configuration on the next boot.

The application can be used to set runtime variable MirrorRequest in the operating system. In the subsequent boot, the UEFI reads this variable and configures mirroring and updates the value to MirrorCurrent. The operating system then reads the value of MirrorCurrent to get the current mirroring configuration that has been applied by UEFI during the boot.

Both the runtime variables use the following GUID:

#define ADDRESS_RANGE_MIRROR_VARIABLE_GUID { 0x7b9be2e0, 0xe28a, 0x4197, 0xad, 0x3e, 0x32, 0xf0, 0x62, 0xf9, 0x46, 0x2c }

The variable data consists of the following structure:

```
typedef struct{
UINT8       MirrorVersion;
BOOLEAN     MirrorMemoryBelow4GB;
UINT16      MirroredAmountAbove4GB;
UINT8       MirrorStatus;


} ADDRESS_RANGE_MIRROR_VARIABLE_DATA;
```

- MirrorVersion is set to 1 in the initial implementation.
- MirrorMemoryBelow4GB is set to **True** to mirror memory below 4GB.
- MirroredAmountAbove4GB is the amount of available memory above 4GB that needs to be mirrored. It is measured in basis points (one hundredth of the value in percentage, that is,1500=15.00%).
- MirrorStatus indicates whether the desired parameters have been successfully applied while setting up the mirroring.

The Linux based tool “efibootmgr” is used for managing UEFI boot entries. UEFI boot entries can be created, reshuffled, and removed using this tool. “efibootmgr” also supports partial mirroring configuration.

To configure partial mirroring, two arguments are passed to efibootmgr:

1. -m to set ‘t(true)|f(false)’ to mirror memory below 4GB.
2. -M to set X Percentage memory to mirror of the system.

Following is an example which sets 15% partial mirroring to the system:

**[root@Gold Desktop]# ./efibootmgr -m t -M 15**

BootCurrent: 0004
Timeout: 2 seconds
BootOrder: 0004,0003,0002,0000,0001
Boot0000* Enter Setup
Boot0001* UEFI Internal Shell
Boot0002* Boot Device List
Boot0003* UEFI INTEL SSDSC2BF240A5 CVTS4261009B240JGN
Boot0004* Red Hat Enterprise Linux
**MirroredPercentageAbove4G: 0.00**
**MirrorMemoryBelow4GB: false**
**RequestMirroredPercentageAbove4G: 15.00**
**RequestMirrorMemoryBelow4GB: true**

# Mirroring configuration priority between UEFI setup page and operating system

The operating system administrator should have knowledge about the memory mirroring size requirement for operating system kernel and critical application usage. Therefore, if a system has both UEFI setup page memory mirroring configuration and operating system memory mirroring configuration, Lenovo ThinkSystem 3rd generation systems offer a value-added feature – retaining both the configurations for the system but assigning operating system mirroring configuration a higher priority in the subsequent system-reboot than the UEFI setup configuration.

Here is an example of how it works. As shown in Figure 6 above, suppose a user configures the system (with one processor and 128GB detected memory) with 20% partial mirroring in the UEFI setup page at first. This leads to the system having 103GB usable memory capacity as shown in Figure 7 above.

Assuming that the user then configures the memory mirroring again with enabling “Mirror Below 4GB” and 15% partial mirroring in the operating system, the system will have 128GB*15% = 19GB (with 1GB alignment) memory in mirroring. The mirroring memory will be in the following two regions:

- One region is of 2GB memory range which is below 4GB because MM Config Base on this system is set to 2GB.
- The other memory region is 19GB – 2GB = 17GB which is above 4GB.

Thus, the system will have a total of 128GB – 19GB = 109GB of available memory.

In this case, Lenovo UEFI initializes the system partial mirroring with operating system configuration. Therefore, in the subsequent reboot, the ThinkSystem splash screen shows up with 109GB usable capacity as shown below:

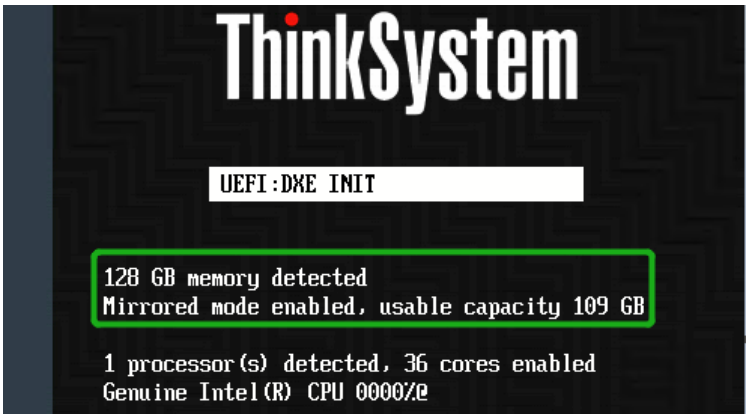



**Figure 10: ThinkSystem Splash Screen with Partial Mirroring via Operating System Configuration**

After booting up, the “Configuration Made From OS” gets displayed in the UEFI setup page as shown below:

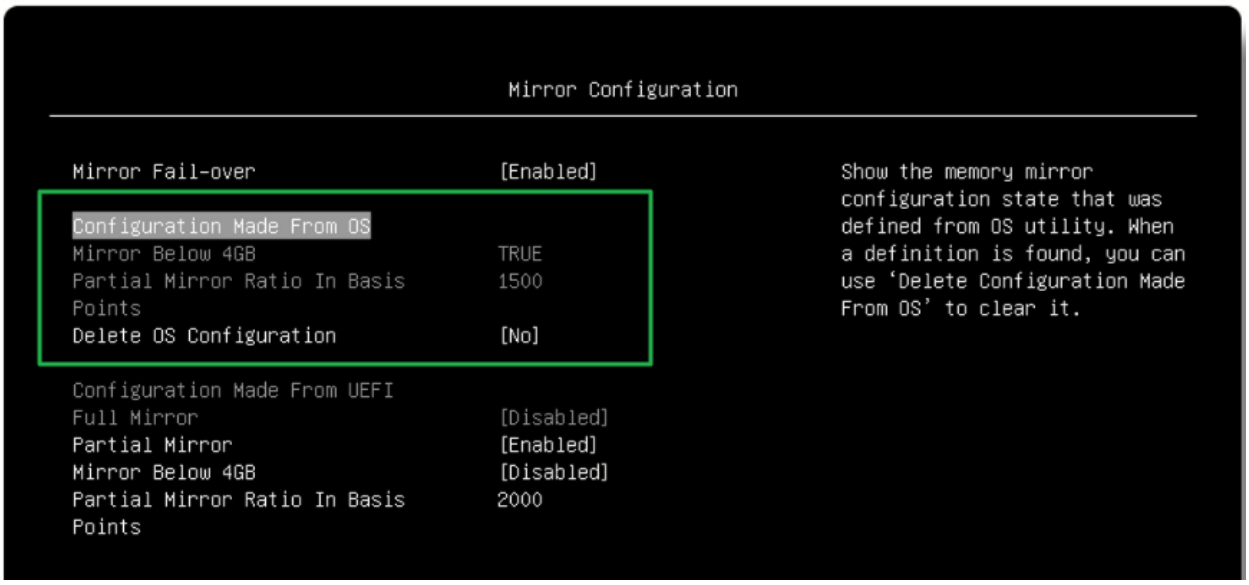



**Figure 11: Memory Mirror Setup Page with Operating System Configuration**

In case the operating system mirroring configuration needs be deleted, it can be done by either of the following methods:

1. Using efibootmgr tool: ./efibootmgr -m f -M 0
   In this method, the operating system mirroring configuration variable is not deleted but is just set to “False” for “Mirror Below 4GB” and 0 percentage for partial mirroring. By doing so, even though memory mirroring settings continue to exist in UEFI setup page, UEFI assigns priority to operating system mirroring settings and configures the system in non-mirror mode in the subsequent reboot.

2. Using “Delete OS Configuration” option (Refer to Figure 11) in UEFI setup page.

In this method, the operating system mirroring configuration variable is deleted. By doing so, if memory mirroring settings exist in UEFI setup page, UEFI applies that configuration in the subsequent reboot.

# References

For further information, refer to the following documents:

UEFI specification: https://uefi.org/specifications

Address Range Partial Memory Mirroring:

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

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 ThinkSystem, are trademarks of Lenovo in the United States, other countries, or both.

Intel, the Intel logo, and Intel Xeon are trademarks of Intel in the United States, other countries, or both.