John R. Encizo
Timothy M. Wiwel
Loc X. Nguyen

Lenovo System x3850 X6 and x3950 X6 Quick Start Guide

Abstract

The Lenovo® System x3850 X6 and x3950 X6 servers are new four-socket and eight-socket servers that deliver fast application performance thanks to an innovative scalable design and new storage technology that is designed to optimize overall solution performance. These servers are based on the highly scalable Intel Xeon processor E7-4800 v2 and E7-8800 v2 product families and are the first servers designed and optimized for eXFlash™ memory-channel storage, the latest in ultra-low latency flash storage technology.

This quick start guide addresses the preferred practices for setting up the x3850 X6 and x3950 X6 servers. It provides all of the information necessary to unpack a system and be up and running in the minimal amount of time. The paper provides insightful information from experts in the field and in development and covers the system architecture and major subsystems such as memory, I/O, power and cooling. The paper also offers a pre-installation checklist and tips on troubleshooting. This paper is intended for pre-sales and post-sales technical support professionals.

Contents

Introducing the x3850 X5 and x3950 X6

The Lenovo X6 product portfolio represents the sixth generation of servers that are built upon Enterprise X-Architecture. Enterprise X-Architecture is the culmination of generations of Lenovo technology and innovation that is derived from the experience in high-end enterprise servers. Now, with the X6 servers, scalable systems can be expanded on demand and configured by using a building block approach that optimizes system design for your workload requirements. These servers scale to more processor cores, memory, and I/O than previous systems, enabling them to handle greater workloads than the systems that they supersede. Power efficiency and server density are optimized, making them affordable to own and operate.

The Lenovo System x3850 X6 and x3950 X6 servers deliver fast application performance thanks to an innovative scalable design and new storage technology that is designed to optimize overall solution performance. The X6 servers are the first servers designed and optimized for eXFlash memory-channel storage. With eXFlash memory-channel storage, they can deliver up to 12.8 TB of ultra-low latency flash storage. With the new Intel Xeon processor E7-4800 v2 and E7-8800 v2 product families, the x3850 X6 and x3950 X6 servers can deliver up to 6.0 TB or 12 TB of memory and 60 or 120 cores of processing power, respectively.

Figure 1 shows the x3850 X6 server.

Figure 1 x3850 X6 server

The x3850 X6 server has the following key characteristics:

- Up to four Intel Xeon processor E7-4800 v2 or E7-8800 v2 product family processors
- Up to 96 DIMM slots (24 DIMM slots per processor) for up to 6 TB of memory (using 64 GB DIMMs)
- Up to 1600 MHz DDR3 memory speeds and up to 2667 MHz SMI2 link speeds
- Up to 12.8 TB of eXFlash memory-channel storage
- Up to eight 2.5-inch hot-swap drives or up to 16 1.8-inch hot-swap solid-state drives (SSDs)
- Support for 12 Gbps serial-attached SCSI (SAS) connectivity for the internal storage

- Mezzanine LOM (ML) slot for the integrated network interface controller (NIC) functionality (choice of dual-port 10 GbE or quad-port 1 GbE adapters)
- Up to 11 PCIe 3.0 I/O slots
- Internal USB port for the embedded hypervisor
- Three USB ports on the front of the server, 2x USB 3.0 and 1x USB 2.0

Figure 2 shows the x3950 X6 server.

Figure 2 x3950 X6 server

The x3950 X6 server has the following key characteristics:

- Up to eight Intel Xeon processor E7-8800 v2 product family processors
- Up to 192 DIMM slots (24 DIMM slots per processor) for up to 12 TB of memory (using 64 GB DIMMs)
- Up to 1600 MHz DDR3 memory speeds and up to 2667 MHz SMI2 link speeds
- Up to 12.8 TB of eXFlash memory-channel storage
- Up to 16 2.5-inch hot-swap drives or up to 30 two 1.8-inch hot-swap SSDs
- Support for 12 Gbps SAS connectivity for the internal storage
- Two ML slots for the integrated NIC functionality (choice of dual-port 10 GbE or quad-port 1 GbE adapters)
- Up to 22 PCIe 3.0 I/O slots
- Two internal USB ports for the embedded hypervisors
- Six USB ports on the front of the server, 4x USB 3.0 and 2x USB 2.0

The supported operating systems for both x3850 X6 server and x3950 X6 server include:

- Microsoft Windows Server 2008 R2
- Microsoft Windows Server 2012
- Microsoft Windows Server 2012 R2
- Red Hat Enterprise Linux 6.5+
- SUSE Linux Enterprise Server 11 SP1+
- VMware vSphere (ESXi) 5.1 U1+
- VMware vSphere (ESXi) 5.5

Architecture

To ensure proper function of the X6 server, the system should be configured correctly at the time of purchase. The block diagram of the system in Figure 3 shows that proper memory and I/O placement is critical for optimal performance. With the move to the Intel Ivy Bridge processors in the X6 platform, the I/O is now directly attached to the processors. Additionally, memory population rules have become more complex because each channel now has three DIMMs and different memory modes.

There are a few general preferred practices in regards to configuring the system. Memory should be installed in quantities of eight DIMMs per CPU for optimal performance and to avoid any potential issues with memory installation order interfering with the memory mode chosen. For best I/O performance, install the correct amount of adapters so that as many processors as possible have a direct connection to I/O.

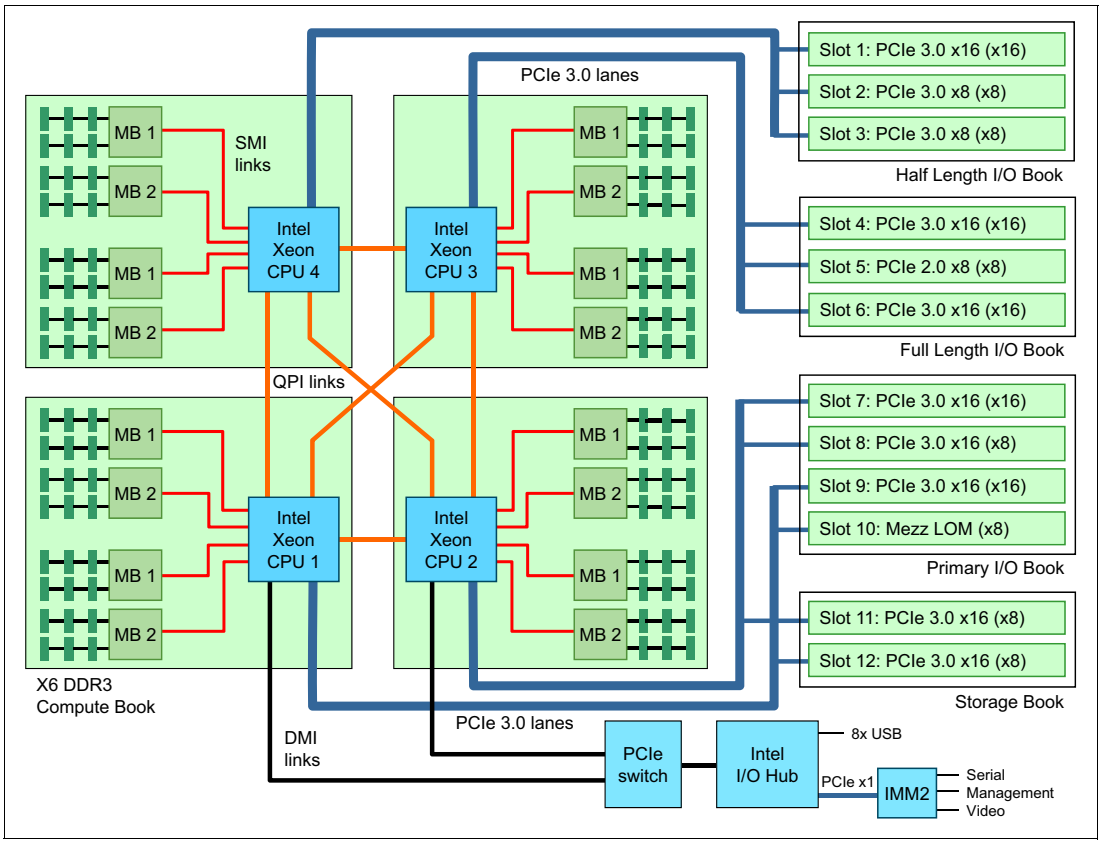


Figure 3 x3850 X6 system architecture

Memory subsystem

The memory subsystem can operate in one of the following modes:

- Independent (Performance) mode
- Lockstep (RAS) mode

In *Performance* mode, each channel on an SMI is addressed independently, essentially providing the ability to run two transactions at once. In *Lockstep* mode, a pair of DIMMs spread across both channels on an SMI is addressed simultaneously. Lockstep mode provides higher memory frequency and memory reliability (Double Device Data Correction +1 bit) but at the expense of higher memory bandwidth. Conversely, Independent mode provides higher memory bandwidth at the cost of advanced memory RAS features and memory frequency (Single Device Data Correction only).

Figure 4 shows the two memory modes.

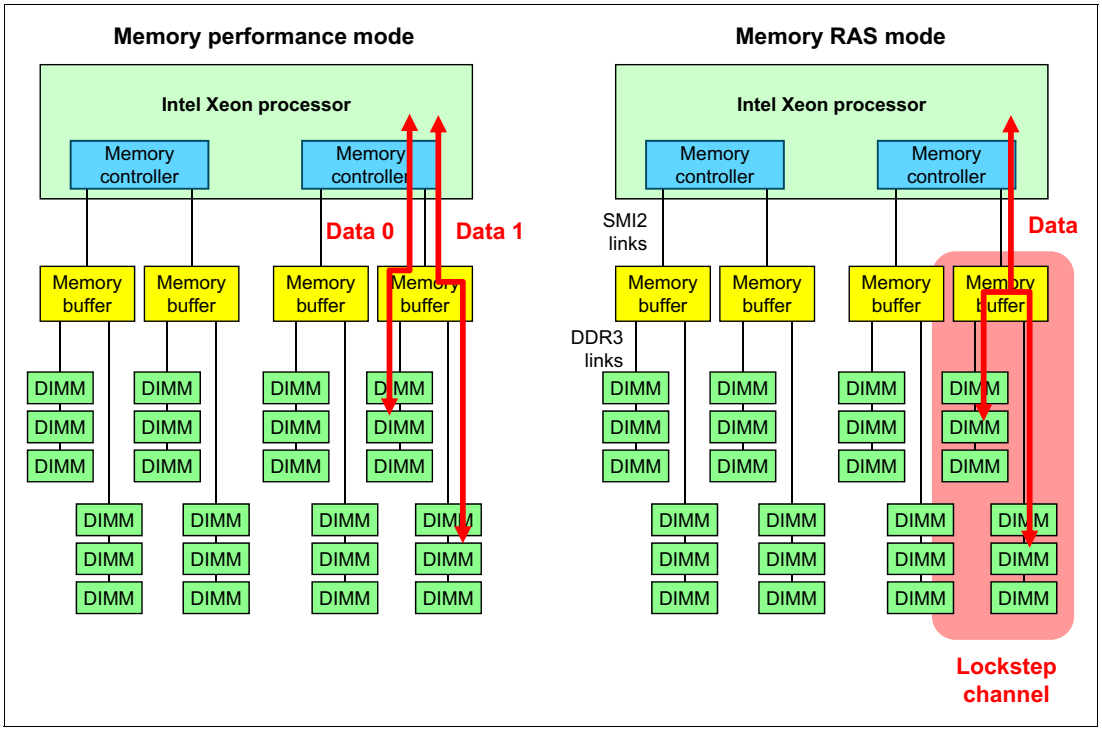


Figure 4 Memory modes: Performance mode (left) and RAS mode (right)

Table 1 shows the characteristics of the supported DIMMs and the memory speeds.

Table 1 Characteristics of the supported DIMMs and the memory speeds

DIMM Specification	**RDIMM**				**LR-DIMM**			
Ranks	Single-rank DIMM		Dual-rank DIMM		4R LR-DIMM		8R LR-DIMM	
Part number	00D5024 (4 GB) 00D5036 (8 GB)		46W0672 (16 GB)		46W0676 (32 GB)		46W0741 (64 GB)	
Rated speed	1600 MHz		1600 MHz		1600 MHz		1333 MHz	
Rated voltage	1.35 V		1.35 V		1.35 V		1.35 V	
Operating voltage	1.35 V	1.5 V	1.35 V	1.5 V	1.35 V	1.5 V	1.35 V	1.5 V
Maximum quantity supported	96		96		96		96	
Maximum DIMM capacity	8 GB		16 GB		32 GB		64 GB	
Maximum memory capacity	768 GB		1.5 TB		3 TB		6 TB	
Maximum operating speed: Independent (Performance) mode								
1 DIMM/channel	1333 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz
2 DIMMs/channel	1333 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz
3 DIMMs/channel	1066 MHz	1333 MHz	1066 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz
Maximum operating speed: Lockstep (RAS) mode								
1 DIMM/channel	1333 MHz	1600 MHz	1333 MHz	1600 MHz	1333 MHz	1600 MHz	1333 MHz	1333 MHz
2 DIMMs/channel	1333 MHz	1600 MHz	1333 MHz	1600 MHz	1333 MHz	1600 MHz	1333 MHz	1333 MHz
3 DIMMs/channel	1066 MHz	1333 MHz	1066 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz	1333 MHz

Figure 5 shows the Intel Xeon E7 memory controller architecture. Note that memory is populated from the furthest DIMM (Slot 0) to the closest DIMM slot (Slot 2).

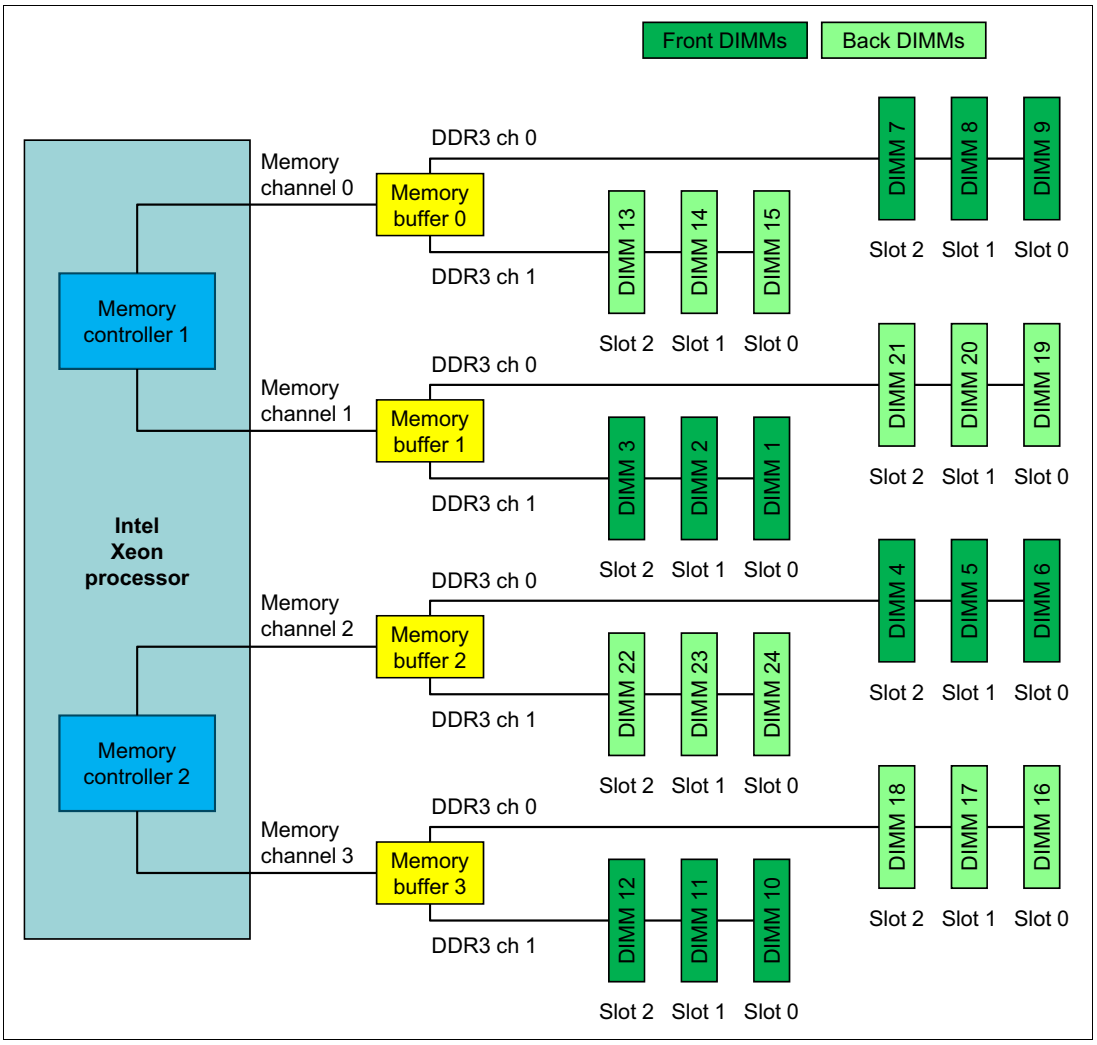


Figure 5 Xeon E7 memory controller architecture

Figure 6 and Figure 7 show the central processing unit (CPU) board layout.

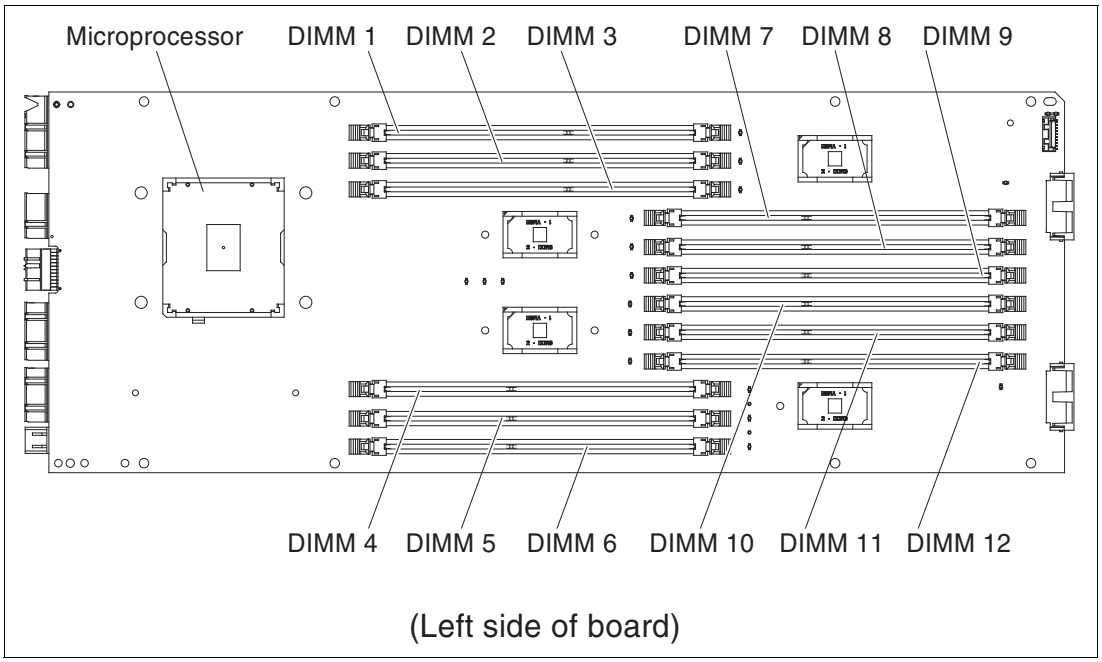


Figure 6 CPU layout (left side of board)

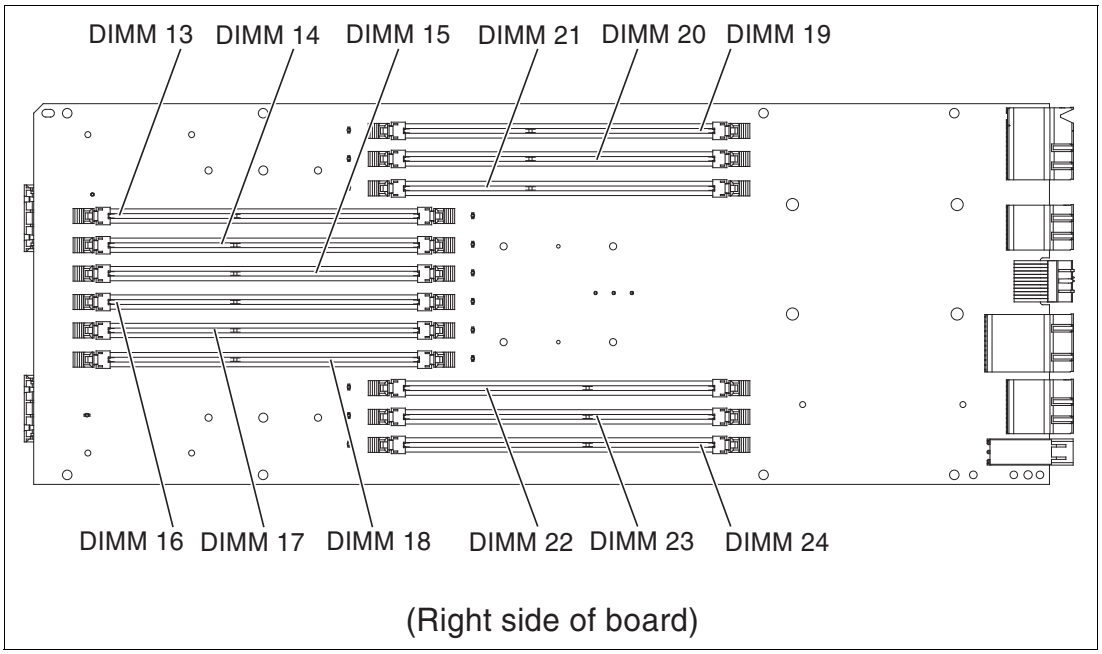


Figure 7 CPU layout (right side of board)

Table 2 shows the memory layout.

Table 2 Memory layout

CPU memory channel	DDR3 channel	DIMM number	DDR3 slot number
Channel 0	DDR3 Channel 0	9, 8, 7	0, 1, 2
	DDR3 Channel 1	15, 14, 13	0, 1, 2
Channel 1	DDR3 Channel 0	19, 20, 21	0, 1, 2
	DDR3 Channel 1	1, 2, 3	0, 1, 2
Channel 2	DDR3 Channel 0	6, 5, 4	0, 1, 2
	DDR3 Channel 1	24, 23, 22	0, 1, 2
Channel 3	DDR3 Channel 0	16, 17, 18	0, 1, 2
	DDR3 Channel 1	10, 11, 12	0, 1, 2

Table 3 shows the Independent mode population sequence.

Table 3 Independent mode population sequence

Installation	Population sequence	DIMM number	DDR3 slot number
1st	1	SMI0, Channel0, Slot0	Front
2nd	2	SMI2, Channel0, Slot0	Front
3rd	3	SMI1, Channel1, Slot0	Front
4th	4	SMI3, Channel1, Slot0	Front
5th	5	SMI0, Channel1, Slot0	Back
6th	6	SMI2, Channel1, Slot0	Back
7th	7	SMI1, Channel0, Slot0	Back
8th	8	SMI3, Channel0, Slot0	Back
...	9 - 16	Repeat order, move to Slot1	Four front, four back
...	17 - 24	Repeat order, move to Slot 2	Four front, four back

Table 4 shows the Lockstep (RAS) mode population sequence.

Table 4 Lockstep (RAS) mode population sequence

Installation	Population sequence	Controller location	Physical location
1st	1 & 2	SMI0, Channel0, Slot0 & SMI0, Channel1, Slot0	Front & back
2nd	3 & 4	SMI2, Channel0, Slot0 & SMI2, Channel1, Slot0	Front & back
3rd	5 & 6	SMI1, Channel1, Slot0 & SMI1, Channel0, Slot0	Front & back
4th	7 & 8	SMI3, Channel1, Slot0 & SMI3, Channel0, Slot0	Front & back
5th	9 & 10	Repeat order, move to Slot1	Front & back
6th	11 & 12	Repeat order, move to Slot1	Front & back
7th	13 & 14	Repeat order, move to Slot1	Front & back

Installation	Population sequence	Controller location	Physical location
8th	15 & 16	Repeat order, move to Slot1	Front & back
...	18 - 24	Repeat order, move to Slot2	Front & back

Table 5 shows the memory population order for both Independent (Performance) mode and Lockstep (RAS) mode.

Table 5 Memory population order for both Independent (Performance) mode and Lockstep (RAS)

DIMM installation order	DIMM slots			
	Performance mode	Performance + mirroring	RAS mode	RAS + mirroring
1	DIMM 9	DIMM 9, 19	DIMM 9, 15	DIMM 1, 9, 15, 19
2	DIMM 6	DIMM 6, 16	DIMM 6, 24	DIMM 6, 10, 16, 24
3	DIMM 1	DIMM 1, 15	DIMM 1, 19	DIMM 2, 8, 14, 20
4	DIMM 10	DIMM 10, 24	DIMM 10, 16	DIMM 5, 11, 17, 23
5	DIMM 15	DIMM 8, 20	DIMM 8, 14	DIMM 3, 7, 13, 21
6	DIMM 24	DIMM 5, 17	DIMM 5, 23	DIMM 4, 12, 18, 22
7	DIMM 19	DIMM 2, 14	DIMM 2, 20	Not applicable
8	DIMM 16	DIMM 11, 23	DIMM 11, 17	....
9	DIMM 8	DIMM 7, 21	DIMM 7, 13	
10	DIMM 5	DIMM 4, 18	DIMM 4, 22	
11	DIMM 2	DIMM 3, 13	DIMM 3, 21	
12	DIMM 11	DIMM 12, 22	DIMM 12, 18	
13	DIMM 14	Not applicable	Not applicable	
14	DIMM 23	....	....	
15	DIMM 20			
16	DIMM 17			
17	DIMM 7			
18	DIMM 4			
19	DIMM 3			
20	DIMM 12			
21	DIMM 13			
22	DIMM 22			
23	DIMM 21			
24	DIMM 18			

Important notes

Keep in mind the following important aspects of the memory subsystem when setting up and using your server:

- The optimal memory configuration is to install DIMMs in quantities of eight per CPU. This configuration ensures that every channel on every memory buffer is populated. Additionally, it ensures that no matter what memory mode is chosen, the system functions correctly.
- Proper memory population rules for DIMMs include:
  - Higher capacity (ranked) DIMMs must be installed first. Follow the population sequence for the appropriate mode.
  - The server supports a maximum of eight ranks (octal-rank) per DDR3 channel. Note that LR-DIMMs might exceed eight ranks per channel through rank multiplication.
  - The server supports 1.35-volt (low-voltage) registered DIMMs, depending on the memory configuration settings in the Setup utility. The server can also operate at 1.5-volt.
  - RDIMMs and LR-DIMMs cannot be mixed in the same system. Make sure to populate the largest DIMMs in the channel first to enable optimal performance.
    - RDIMMs are available in 4 GB, 8 GB, and 16 GB.
    - LR-DIMMs are available in 32 GB and 64 GB.
  - A minimum of one DIMM must be installed for each Compute Book. Depending on the memory mode (for example, Lockstep), this configuration might not be sufficient for operation.
- The default memory mode for the X6 system is Lockstep, which is important for those scenarios where less than the optimal number of DIMMs per processor are installed. Always be sure to properly configure the memory mode in Unified Extensible Firmware Interface (UEFI) before changing the DIMM installation order, but prior to a reboot, for this configuration to take effect.

  When not fully populating system memory according to preferred practices, changing the memory mode can cause the system to fail to boot. The memory population label on the CPU book is for Performance mode. Lockstep mode DIMMs should be installed in pairs. You can use the Advanced Settings Utility™ (ASU) tool to remotely flip the memory mode. For example:

  ```
  asu64 Memory.MemoryMode Independent OR ./asu64 Memory.MemoryMode Lockstep
  ```
- The maximum operating speed of the server's memory is determined by the slowest DIMM in the server.
- If swapping out a failed DIMM, make sure you move the DIMM to a different SMI and re-enable the failed DIMM slot.
- If planning to install eXFlash DIMMs, note that eXFlash DIMMs are RDIMMs and, therefore, cannot be used with either the 32 GB or 64 GB LR-DIMMs.
- No more than one eXFlash DIMM per channel is supported, and at least one RDIMM must already be present in the same channel as the eXFlash DIMM that it will be installed in.

- General eXFlash DIMM population rules are as follows:
  - A maximum of four processor books can be configured with eXFlash DIMMs.
  - For the x3950 X6, the primary node (lower processor books) must host the eXFlash DIMM modules.
- Population rules for up to eight eXFlash DIMMs and up to eight memory modules in a Processor Book (two DIMMs per channel is 1333 MHz) are as follows:
  - RDIMMs install to DIMM slots 9, 6, 1, 10, 15, 24, 19, 16
  - eXFlash DIMMs install to DIMM slots 8, 5, 2, 11, 14, 23, 20, 17
- Population rules for up to eight eXFlash DIMM and up to 16 memory modules in a Processor Book (two DIMMs per channel is 1333 MHz) are as follows:
  - RDIMMs install to DIMM slots 9, 6, 1, 10, 15, 24, 19, 16, 8, 5, 2, 11, 14, 23, 20, 17
  - eXFlash DIMMs install to DIMM slots 7, 4, 3, 12, 13, 22, 21, 18

I/O subsystem

Figure 8 shows an overview of the I/O layout of the x3850 X6 server.

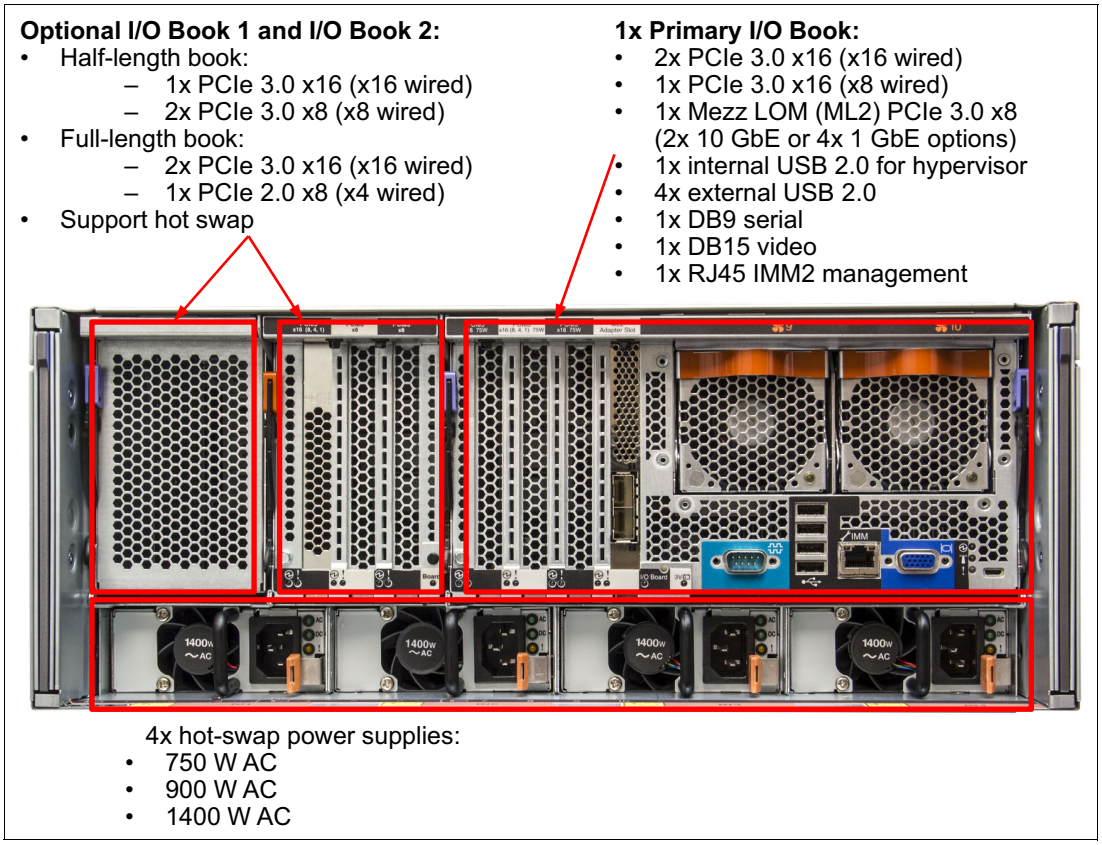


Figure 8 I/O layout of the x3850 X6 server

The I/O layout of the x3950 X6 server is essentially the same as the x3850 X6, just with twice the components.

The I/O subsystem is directly attached to a processor.

- Socket 1 (the CPU book closest to the Storage Book) is responsible for the ML2 slot, Slot 9 and Slot 12.
- Socket 2 is responsible for Slots 7- 8 and Slot 11.
- Both sockets 1 and 2 drive their respective I/O through the Primary I/O Book.
- Socket 3 is attached to Slots 4-6, and Socket 4 is connected to slots 1-3.
- Slots 1-6 are configurable in how they present I/O. In the case of the FH/HL I/O Book, this is presented as 1x16 slot and two x8 slots. For the FH/FL I/O Book, this configuration is presented as 2x16 slots and 1x8 slot.
- You can opt for any combination of the two possible I/O Books.
- In an eight-socket configuration, Socket 1 is in the bottom node.

Table 6 shows the processor and PCIe slot association.

Table 6 Processor and PCIe slot association

Processor	PCIe slot association
CPU 0 (Compute Book 1)	9, 10 (Primary I/O Book) and Storage Book Slot 12
CPU 1 (Compute Book 2)	7, 8 (Primary I/O Book) and Storage Book Slot 11
CPU 2 (Compute Book 3)	4, 5, and 6 (I/O Book 2)
CPU 3 (Compute Book 4)	1, 2, and 3 (I/O Book 1)

Slot 10 configuration: Slot 10 is intended for Mezzanine LOM 2 (ML2) network adapters only and cannot be used for regular PCIe adapters.

Figure 9 shows the numbering of the PCIe slots on the x3850 X6.

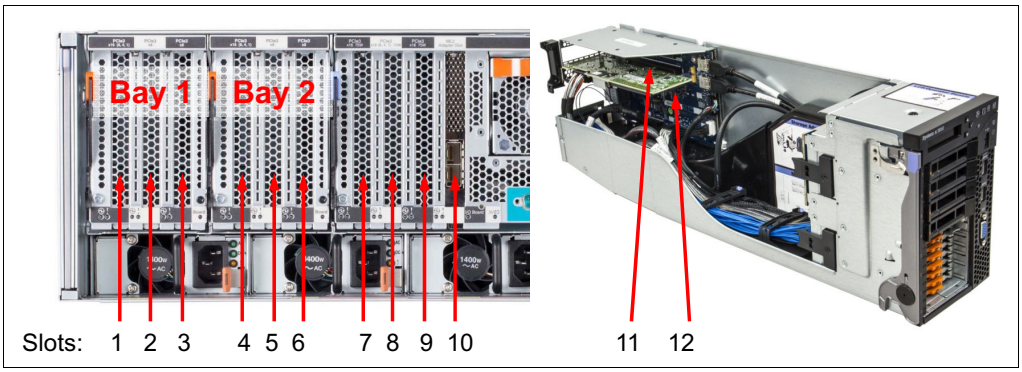


Figure 9 x3850 X6 PCIe slot numbering

Figure 10 shows the numbering of the PCIe slots of the x3950 X6.

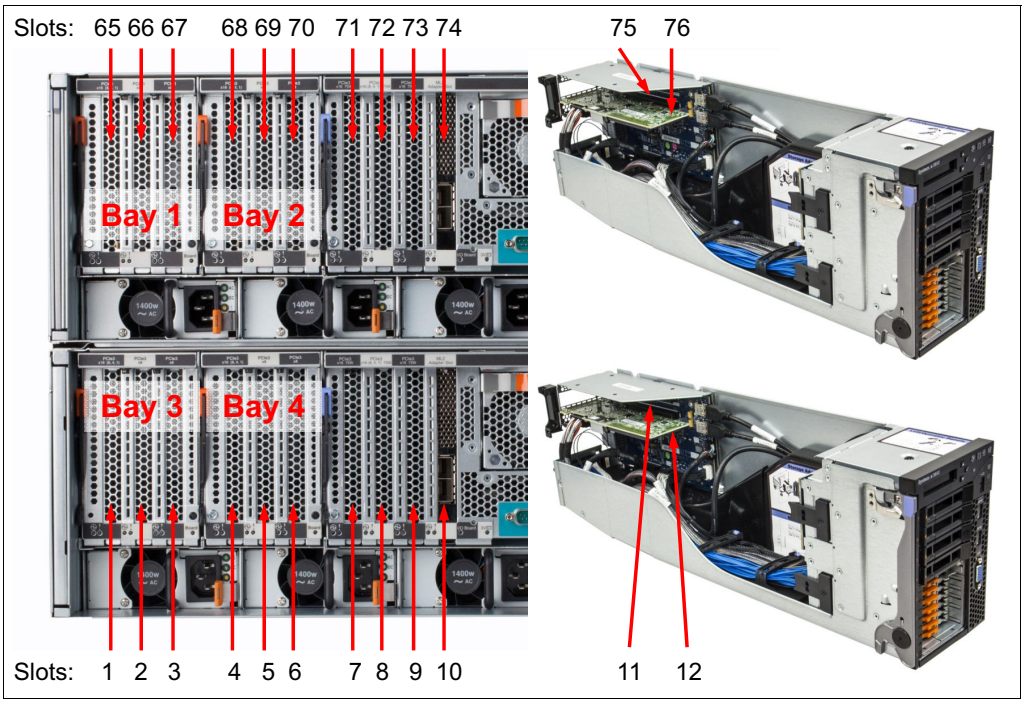


Figure 10 x3950 X6 PCIe slot numbering

Important notes

Keep in mind the following important aspects of the I/O subsystem when setting up and using your server:

- Proper I/O placement is dependent on the types of adapter cards. Have at least two of any type of adapter in the system to maintain enough I/O to keep the processors busy. For low latency or high bandwidth applications, consider using one adapter per processor to give each CPU direct access to I/O. In general, at least one host bus adapter (HBA), if HBAs are used, and one NIC should be present in the Primary I/O Book (Slots 7-9 in a four-socket configuration or Slots 39-41 in an eight-socket configuration) and the additional adapters should be populated so that you have one adapter type per book. If using an eight-socket configuration, have at least one HBA in slots 39-41 and Slots 7-9 first. After the initial one adapter type per CPU population is complete, go back and populate from higher numbered slots to lower numbered slots using the same population rule as before (one adapter type per I/O Book).
- Take care when inserting I/O Books to avoid damaging either the connector on the I/O Book itself or the midplane. Inserting the I/O Book at an angle or with I/O cables attached can lead to the I/O Book damaging the midplane.

  To install the Full Length I/O Book, do the following, as shown in Figure 11:

  a. Install the support plate.

  b. Insert the long I/O Book.

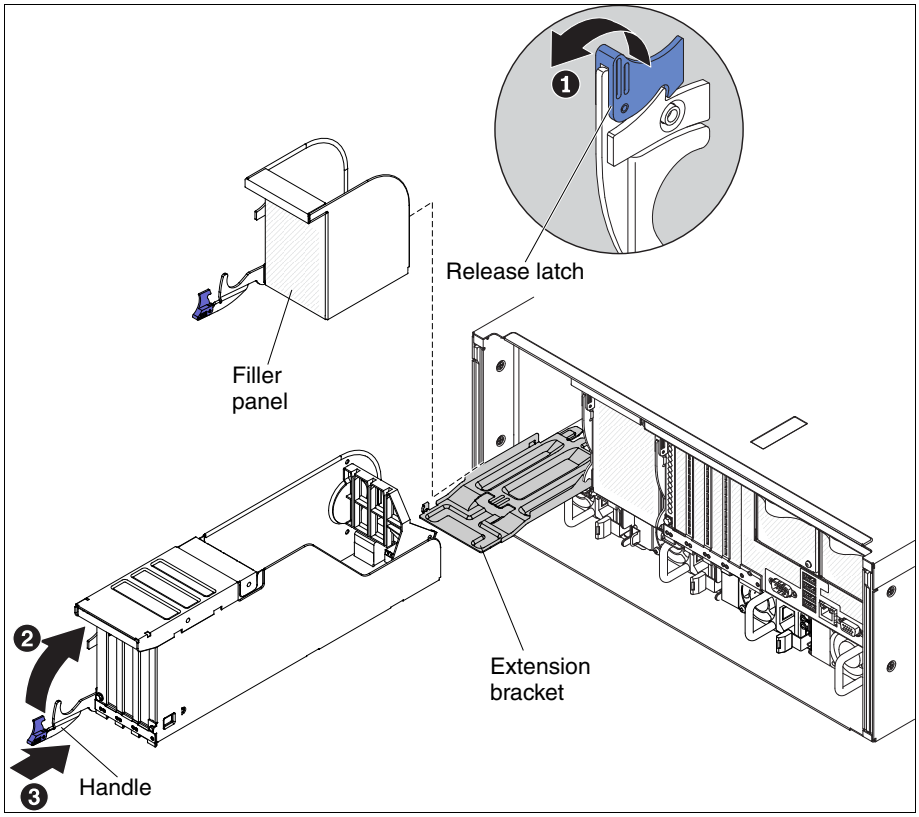


Figure 11 Install the F/L I/O Book

To install the Primary I/O Book, do the following: as shown in Figure 12 on page 16:

a. Validate that the system has AC power removed.

b. Place the I/O Book into the system. There are two (non-hot swap) handles on the Primary I/O Book. These handles should be engaged at the same time to lock the board into place.

c. Reapply the AC power.

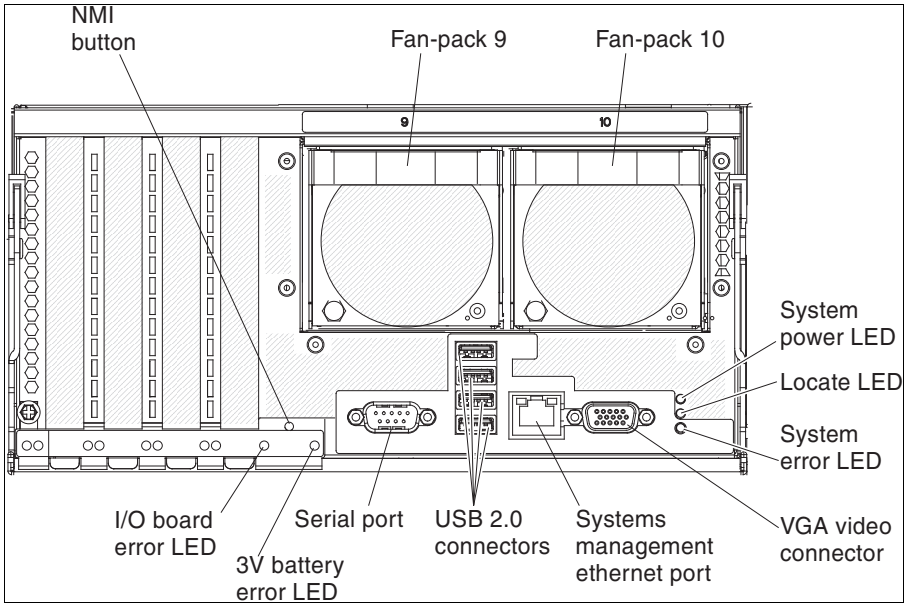


Figure 12 Install the Primary I/O Book

- If an I/O Book is removed without first pressing the attention button (Figure 13) on the I/O Book to shut off power to all adapters in that book, I/O Books can cause the system to experience instability.

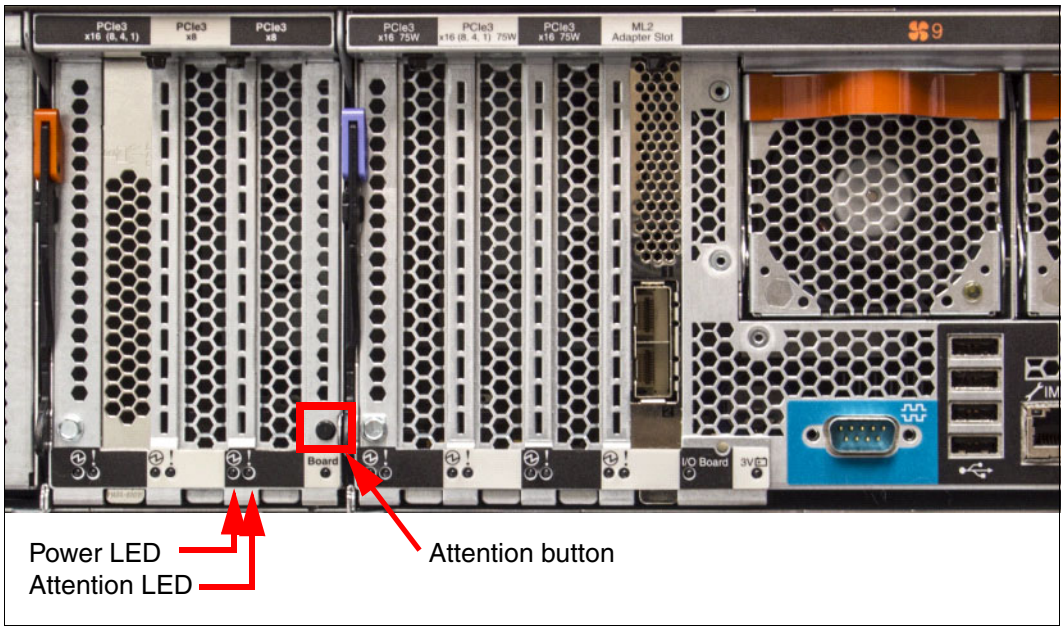


Figure 13 Attention button

- Use care when installing PCIe cards so that the metal PCIe card retention latch is fully seated. Failure to properly align and seat the retention PCIe card retention latch can lead to either PCIe cards becoming loose (causing system stability issues) or the PCIe retention latch to break when the I/O Book is installed.
- With back-level firmware, the middle slot on FH/FL I/O Books can become disabled. This is connected in UEFI 1.0 (BuildID A8E112A).

- The Integrated Management Module (IMM) defaults to acquire a DHCP IP address. In those cases where the IMM has been previously assigned an IP address or you are unsure what IP address it has acquired, the front LCD panel (Figure 14) can be used to list the IP address of the IMM. The default IMM IP address is 192.168.70.100.

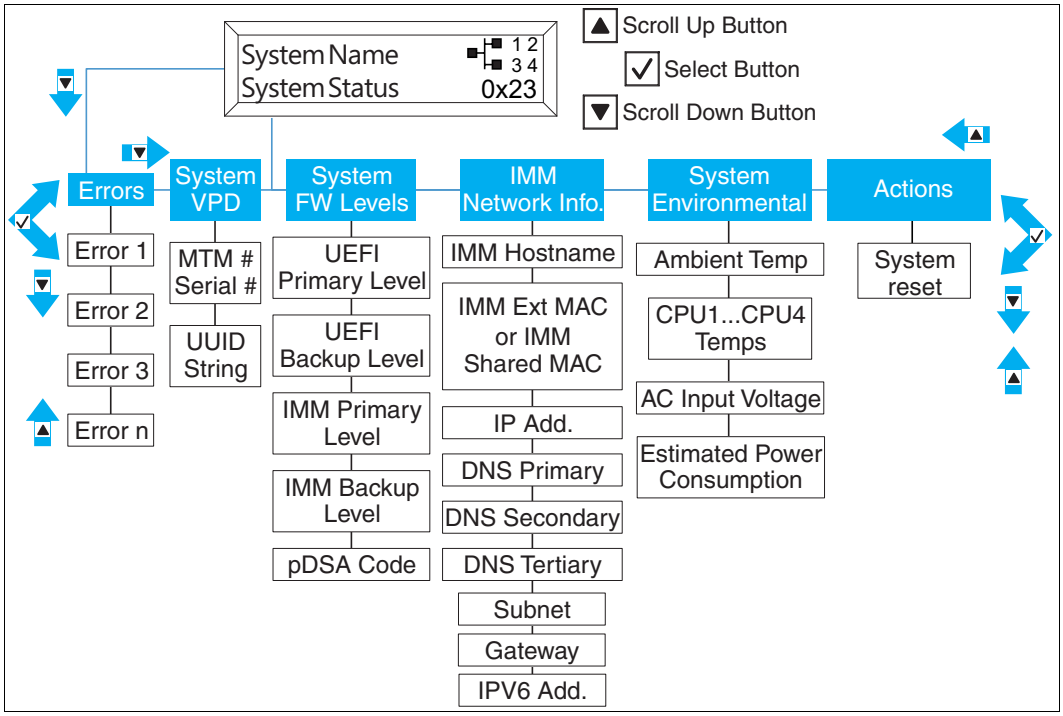


Figure 14 System properties on the LCD system information panel

- The Primary I/O board is required in any system. The Primary I/O board hosts the IMM and UEFI code, as well as the system battery, USB hypervisor port, and the Mezzanine LAN-on-motherboard generation 2 (ML2) adapter slot. The ML2 slot (Slot 10, as highlighted in Figure 15) is designed exclusively for ML2 adapters. ML2 network adapter cards only work in the ML2 slot.

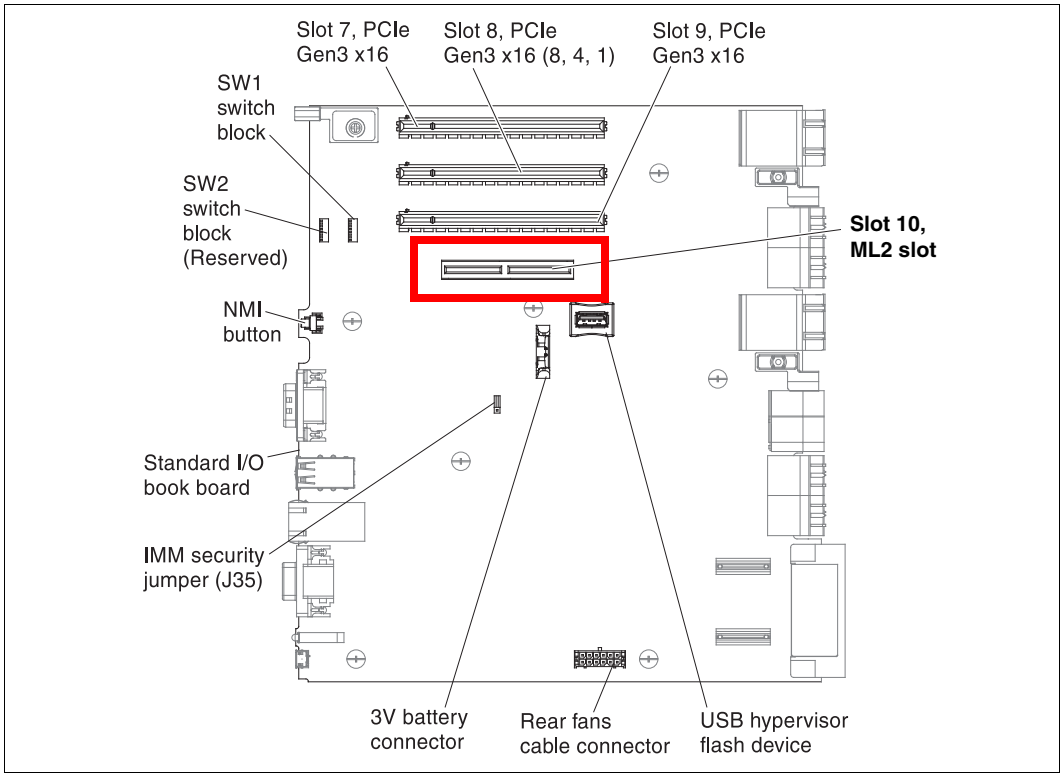


Figure 15 ML2 slot

Power subsystem

The x3850 X6 and x3950 X6 systems are designed with modular power subsystem. This design allows for one to choose the optimally-sized power supplies for your system configuration. The power subsystem is designed around two zones to ensure that in the event of a loss of a power feed the system continues to function.

Power supplies are numbered (from left to right) 4, 3, 2, and 1, as shown in Figure 16. The even power supplies are in one power feed zone, and the odd power supplies are in the other power feed zone.

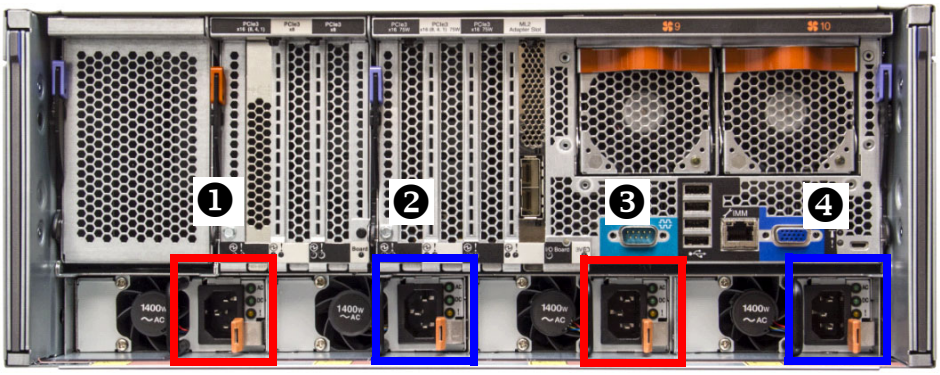


Figure 16 Power supplies

For the power subsystem, keep in mind the following notes:

- Power supply unites (PSUs) are installed in the following order for the x3850 X6:
  a. Power supply bay 3
  b. Power supply bay 2
  c. Power supply bays 1 and 4
- If systems have mixed 900W and 1400W PSUs, make sure the PSUs are populated correctly: 1400W PSUs in the odd numbered bays and 900W PSUs in the even numbered bays.
- A reboot is required to recalibrate system power any time settings are changed that impact overall power draw on the system.
- Systems plugged into 110/120V power have their 1400W PSUs operate at 900W. Keep this in mind when calculating the system power.
- When changing power redundancy settings, be aware that this impacts overall power requirements for a system. Setting a system to a power redundancy level that its input load cannot support can cause the system to go into a reboot loop. If the system fails to power and reboot after the “Recalibrating System Power …” message displays on the screen, complete the following steps:
  a. Log in to the IMM.
  b. Click the **Server Power Management** tab, as shown in Figure 17 on page 20.
  c. Select **Power Supply Redundancy**. The Power Policies window opens, as shown in Figure 18 on page 20.
  d. Make sure there is enough power to power the system.
  e. Make sure there is enough power to run the system with the level of redundancy required.

- In an eight-socket system, the top four supplies power the top four CPUs, and the bottom four supplies power the bottom four CPUs.

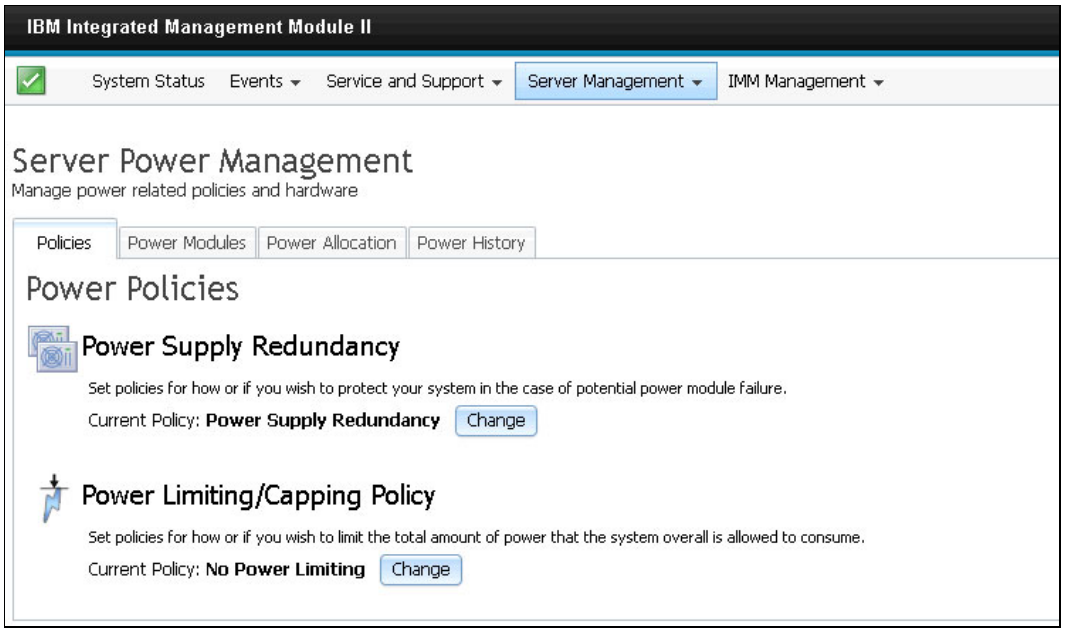


Figure 17 Server Power Management tab

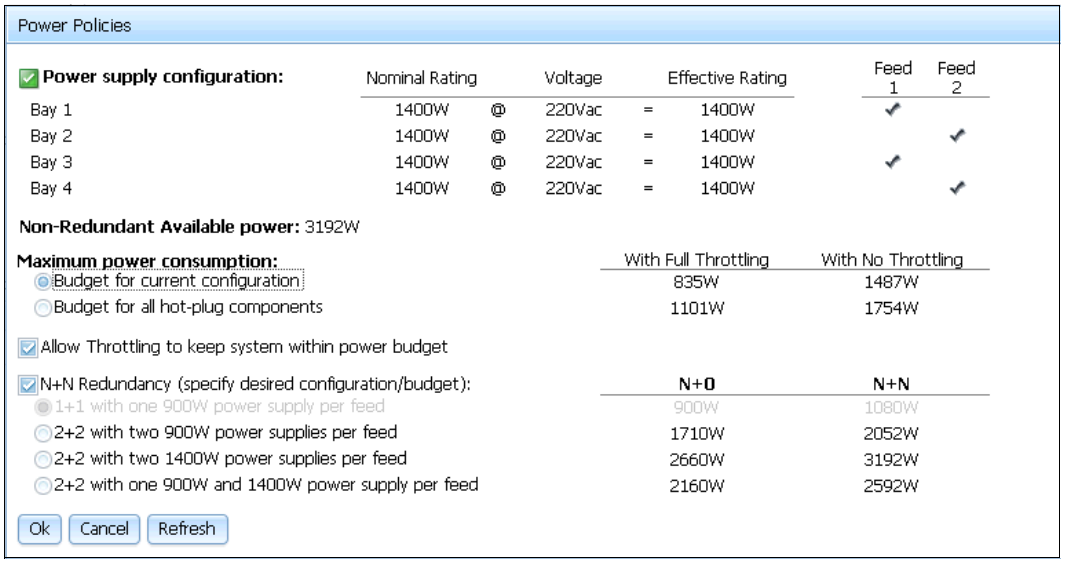


Figure 18 Power Policies window

Cooling subsystem

There are five cooling zones (Figure 19) fed by two different types of fan packs. The Storage Book is fed by a fan pack on the Primary I/O Book, and each CPU book is fed by a pair of fan packs (Figure 19).

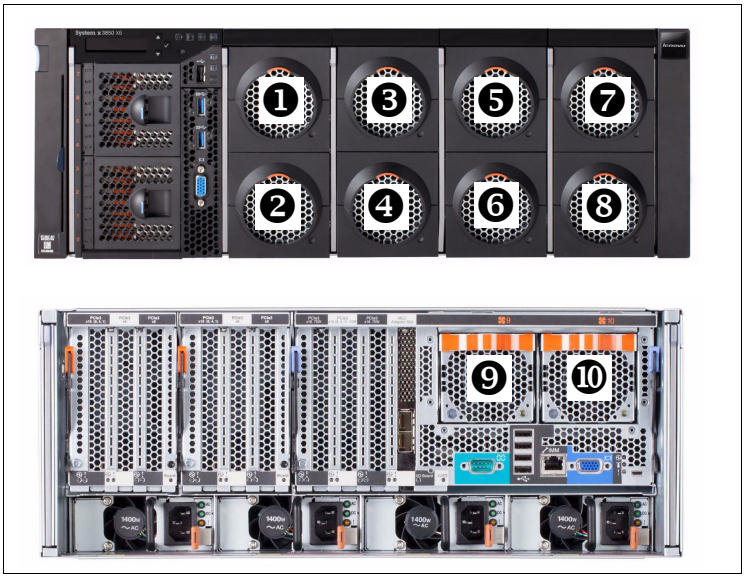
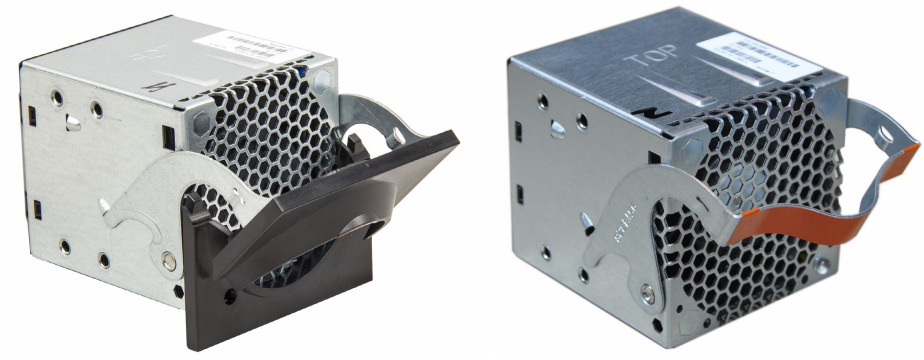


Figure 19 Cooling zones

Keep in mind the following notes regarding the cooling subsystem:

- If a fan on a CPU book is removed for more than 10 seconds, the system begins to generate thermal events, and the system might shut itself down in approximately 5 minutes if the fan is not replaced.
- A failed fan pack on a system causes it to fail to power on. If you have a fan installed and the system fails to power on, remove the fan pack (Figure 20) and validate that the power connector on the fan has not come loose.



Compute Book fan Primary I/O Book fan

Figure 20 Compute and Primary I/O Book fans

- With only four drives installed, it is mandatory to leave the top filler (Figure 21) in place for cooling and EMC reasons.

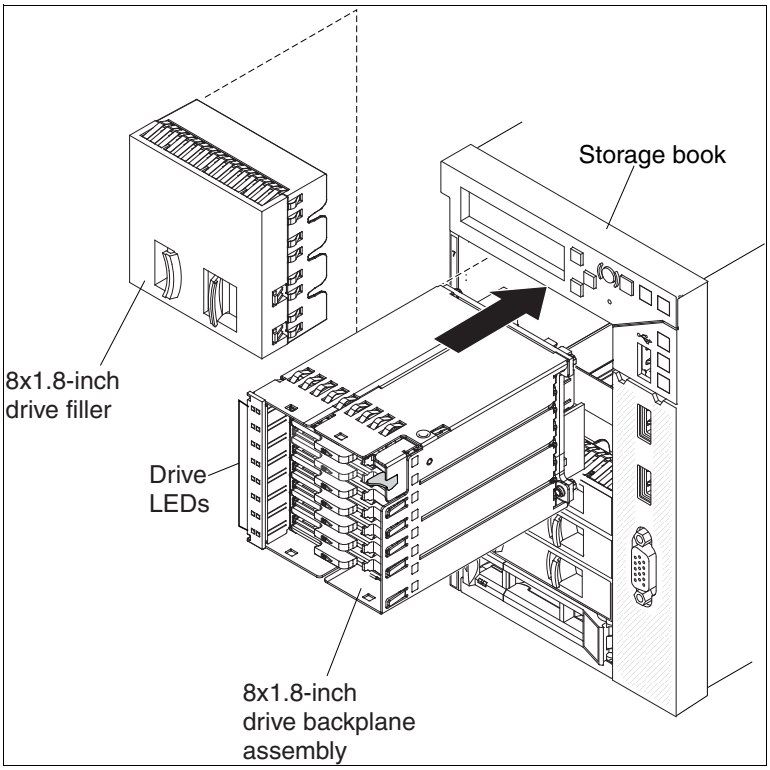


Figure 21 Drive filler

Pre-installation checklist

Before you install the server and get it up and running, complete the following steps:

1. Validate the type of power feeds for the system. Make sure that both the number and types of power supplies are present to support the level of PSU redundancy required.
2. Prior to installing any additional hardware, apply AC power and log into the IMM. Update the base system firmware (IMM/UEFI/Dynamic System Analysis (DSA)) via the IMM. Update the firmware in the following order:
   a. Update both IMM banks and restart the IMM, which can take up to 15 minutes.
   b. Update both UEFI banks and start the system. (This process can take up to 30 minutes for the firmware update to complete.) The F1 screen displays when UEFI has completed the update.
   c. Press F1, go into UEFI, and load the defaults. This process can also be done via the ASU tool by entering:

      ```
      /asu64 loaddefaults UEFI --host xxx.xxx.xxx.xxx --user USERID --password
      PASSWORD
      ```

   d. After the IMM and UEFI have been updated via the IMM interface, we recommend using Bootable Media Creator™ (BoMC) to update the system to the current versions of firmware.

If using a M5210 RAID controller, prior to drives being addressable as installation targets, you must first create a RAID array. RAID arrays can be created by using the **F1** → **System Settings** → **Storage** selection:

1. Select the RAID controller that you want to configure arrays for, as shown in Figure 22. This opens the main menu for that controller, as shown in Figure 23. If there are no virtual drives created when you select that component, the main menu is empty.

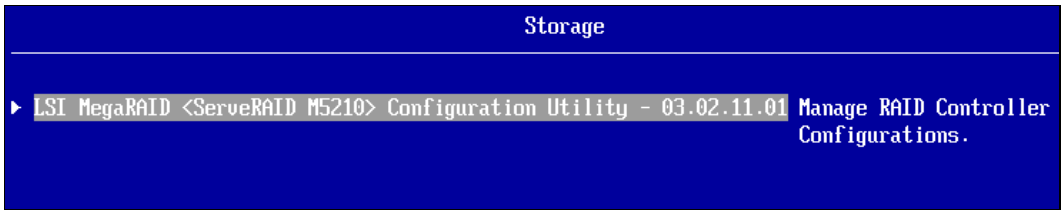


Figure 22 Select the RAID controller

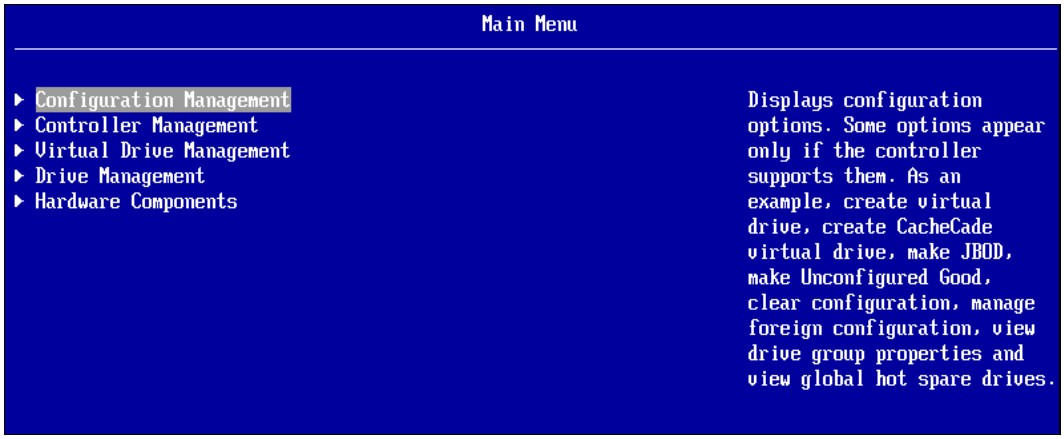


Figure 23 Controller main menu

2. To create a RAID array, select Configuration Management, as shown in Figure 24 on page 24. You can pick from the following options:
   - The Create Virtual Drive option does basic RAID configuration for Generic RAID 0, 1, 5, and 6 (if applicable). This basic configuration gives you little control in how arrays are laid out.
   - The Make JBOD option is supported in those scenarios where you have a M5210 adapter, but do not have the cache module installed. This option sets the controller in pass-through mode, which allows the drives to become directly addressable to the OS. JBOD mode and virtual drive (RAID) mode are mutually exclusive.
   - The Create Virtual Drive - Advanced option (Figure 24 on page 24) allows you to make multiple arrays and configure the array parameters. Any parameters that are grayed out are not available either because of the number of drives present in the system or, as shown in Figure 24 on page 24, because of the lack of a cache module. Be sure to always save the array creation at the end to commit the work. Note that when selecting drives (Figure 25 on page 24), only those drives not already in an array or of the drive type selected are displayed as targets.

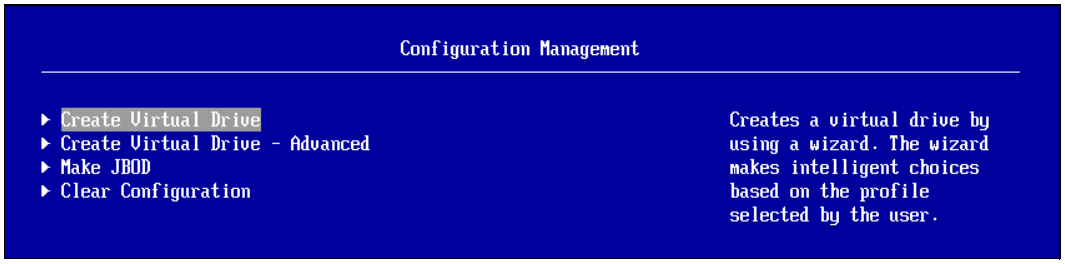


Figure 24 Create Virtual Drive - Advanced

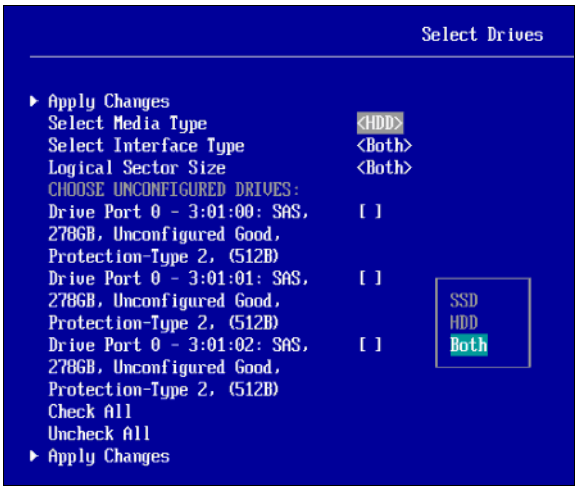


Figure 25 Select Drives

3. Select the drives.
4. Configure the type of RAID.
5. Configure the size of the array to create on the drives. (You can opt to not use the entire size for the array if you want to create more than one logical partition on the drive.)
6. Set the default initialization type (use “Fast” for most purposes).
7. Save the configuration.
8. Select **Yes** to confirm the settings and create the array, as shown in Figure 26 on page 25.

The M5210 RAID card requires a driver to both install and function correctly in older Windows (2008 and 2012) and Linux (RHEL 6.4, 6.5, SLES 1 and 1 SP1) operating systems. The driver `.inf` files for Windows and RPM and DUD files are included on the x3850 X6 driver download site.

If installing Windows 2012+, it is required to install the OS in UEFI mode. Failure to perform a UEFI mode install results in the OS installer aborting prior to getting to the graphical portion of the install. A UEFI mode install is required since the X6 includes the hardware for implementing Secure Boot.

Requirements: A UEFI install requires the OS to be installed on a GPT disk and places an entry in the Boot Order pointing back to the UEFI full path boot path of the .EFI loader.

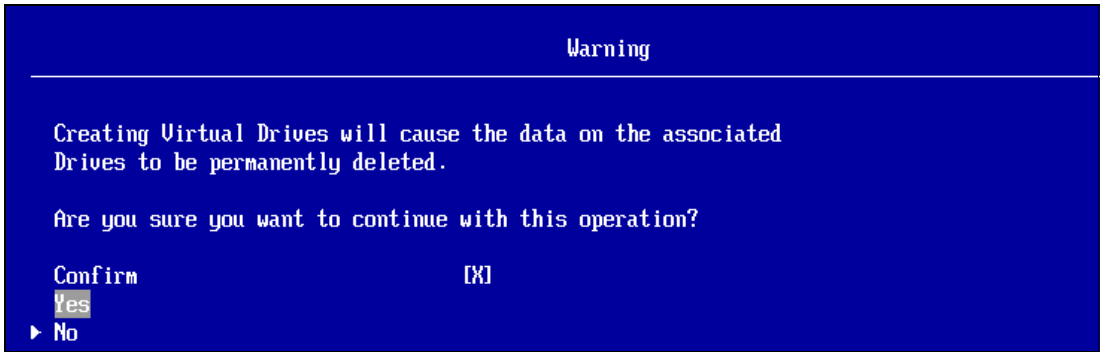


Figure 26 Confirm the settings

Troubleshooting tips

If you are having problems with the server, review this list of troubleshooting tips to aid with problem determination.

- If the fan speed on the CPU book ramps up, but there is no error on the LCD display:
  - Check the other fan on the same CPU book. The power connector might be broken or pushed in the housing, as shown in Figure 27.

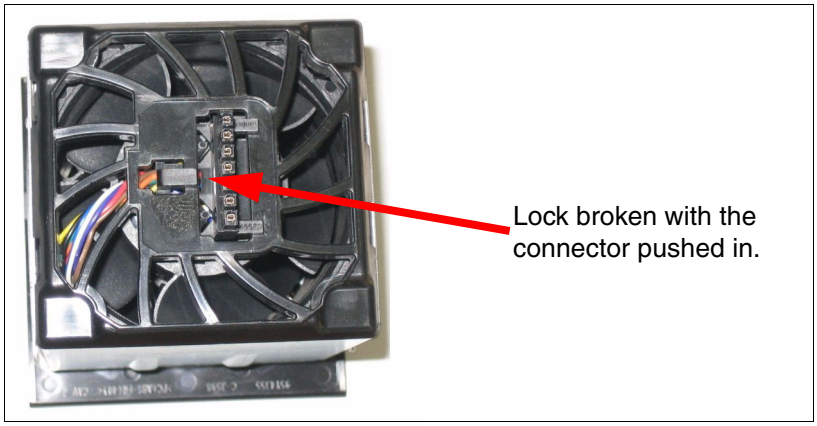


Figure 27 Fan unit with broken lock mechanism

– Check the IMM log, as shown in Figure 28.

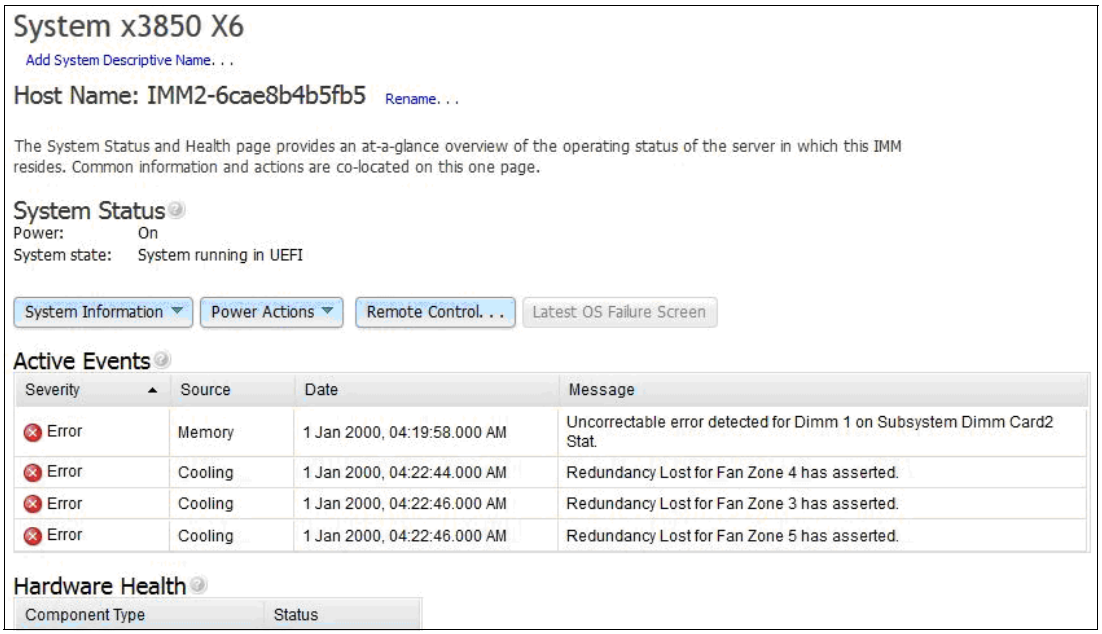


Figure 28 IMM log

- If the system fails to boot to Setup (F1), complete the following steps:
  a. Remove the AC power.
  b. Remove the main I/O Book.
  c. Pull the CMOS battery and let the system discharge (two minutes). This process is equivalent to setting the system to default.
  d. Replace the CMOS battery and reinstall the I/O Book. The system should now boot correctly.
- If the system fails to power and reboot after the “Recalibrating System Power …” message displays:
  a. Log in to the IMM.
  b. Click the **Server Management** tab.
  c. Select **Power Redundancy** and check the following:
     i. Make sure there is enough power to power the system.
     ii. Make sure there is enough power to run the system with the level of redundancy required.
- Downlevel RAID firmware can cause the system to panic when MegaCLI is used to query the adapter. Update to the newest firmware on the x3850 X6 support site (ibm_fw_sraidmr_5210-24.2.1-0027).
- If the system is in a state where the system will not power on and you see firmware errors:
  a. Remove the AC power.
  b. Remove all but CPU0 and the primary IO module (also remove all adapters from the primary IO module and use the CMOS battery tip to discharge the primary IO Modules capacitor).
  c. Reapply power and log into the IMM to power on the system.

d. If the system has been in a firmware “Pending” state, allow the system to go through multiple reboot cycles to update the CRTM. This can take upwards of 30 minutes. The F1 prompt displays when the system reboot is complete.

e. Load the defaults, save, and exit. The system goes down to reboot.

f. Remove the AC power, remove the primary IO book, and re-install the adapters.

g. Insert the Storage Book.

h. Reapply power and log into the IMM to power on the system. The system comes up to the F1 prompt.

i. Load the defaults, save, and exit. The system goes down to reboot. When the system goes down, remove the AC power.

j. Insert the remainder of the CPU books and IO Modules.

k. Reapply power and log into the IMM to power on the system. The system comes up to the F1 prompt.

l. Load the defaults, save, and exit. The system goes down to reboot. Allow the system to successfully complete a power recalibration cycle. When the system successfully comes up to the F1 prompt, use the IMM or front power button to turn off the system power.

m. Remove AC power and reinstall the remainder of the option cards.

n. Allow the system to boot to F1, load the defaults, save, and reboot.

- Before adding CPU books, power off the server and remove the AC power cords. If you did not remove the AC power when adding a CPU book and if the server fails to restart properly (either fails to power on or goes into an endless boot loop), the IMM data might have corrupted. The solution is to reflash both IMM banks again.

  Quick method: Re-flashing the IMM can take up to 15 minutes, but you can still ping the IMM and log in via SSH. You can use a quick “hack” method to update it. Download and extract the UXSPI tool, grab the “iflash” directory and your `.uxz` file, and run the following command to iflash it remotely out of band:

  ```
  iflash64.exe --package IMMfirmwarefile.uxz  --host XXX.XXX.XXX.XXX --user
  USERID --password PASSWORD --unattended [--reboot ]
  ```

- Make sure fillers are installed in the empty drive cages and backplanes or the system will generate thermal alerts.
- Windows 2012 R2 Installations have issues in Legacy mode. This is due to the inclusion of the Secure Boot functionality in the system. To install Windows 2012, you must use UEFI mode. Legacy mode installation fails with a “`winload.exe`” error prior to the graphical part of the install commencing.
- If the remote install SLES 11 does not boot the ISO file using the virtual, the workaround is to:

  a. Log in to the IMM and start the virtual media.

  b. Mount the SLES 11 ISO file.

  c. Apply power to the x3850 X6 and boot to setup using F1.

  d. Select **Boot Manager** → **Boot from File**. The USB CDROM device, which is the virtual media device, is displayed.

  e. Select the device and press Enter.

  f. Select **EFI** and press Enter.

g. Select **Boot** and press Enter.

h. Select **bootx64.efi** and press Enter. The install process should start.

- The blank front panel LCD might have a loose connection on the Storage Book. Remove the Storage Book from the chassis and verify that the cable is seated properly.
- If attempting to log in to the IMM and the web screen is not loading fully:
  a. Log in to the IMM using SSH.
  b. Reset the IMM to the default setting.
- If the PCIe hot plug is not working in the current mode (UEFI/IMM):
  a. The system does not see adapters installed in full length I/O Book slot 1 and 4.
  b. The system does not see adapters installed in half length I/O Book slot 2 and 5.
  c. These issues are fixed in the new code stack (57S (58S) and 107M (108M).
- If attempting to manually configure RAID with an M5210 card in UEFI, then it is required to upgrade the UEFI, IMM, and RAID firmware to the latest firmware levels prior to attempting to configure arrays.
- For an eight-socket system, if for any reason one node turns on and the other one does not, DC off the node that is on, then AC cycle both nodes.
- If the fault LED is lit, look at both the IMM web log and the front panel to determine the failure cause. In some failure modes, an error message is not printed to the LCD, but is logged to the IMM. All error messages from the IMM might not populate to the front LCD panel.
- When attempting to AC cycle the system, be aware that PSUs can retain power for up to 30 seconds after AC line cords are removed. For a successful AC cycle, ensure that all LEDs are off on the primary I/O module and installed PSUs before re-installing the line cords.
- When the DC fan is turned on, if a fan is not present, of if one rotor is bad, the system powers off within 3 minutes.
- The top operator panel might not show the same status as the bottom operator panel. The primary nodes operator panel should be the one that is used to get the system information.
- All HDD fillers must be installed if a drive back plane is not installed. If the fillers are not in place, then the HDDs could overheat in extreme conditions.
- Install VMware in legacy mode.

- The quick path interconnect (QPI) debug consists of a set of 20 TX lanes plus a CLK and a set of 20 RX lanes plus a clock. If any one lane fails, the system picks 10 good lanes and runs in half width. If the clock fails, the system uses one of the data lanes as the CLK and runs in half width. This *faildown* happens independently on TX and RX. That is, the RX can run in half width, and the TX port runs in full length. Each CPU has three QPI interfaces labeled 0, 1, and 2.

  In a four-socket or eight-socket system, the QPI meshes all CPUs together, as shown in Figure 29 and Figure 30.

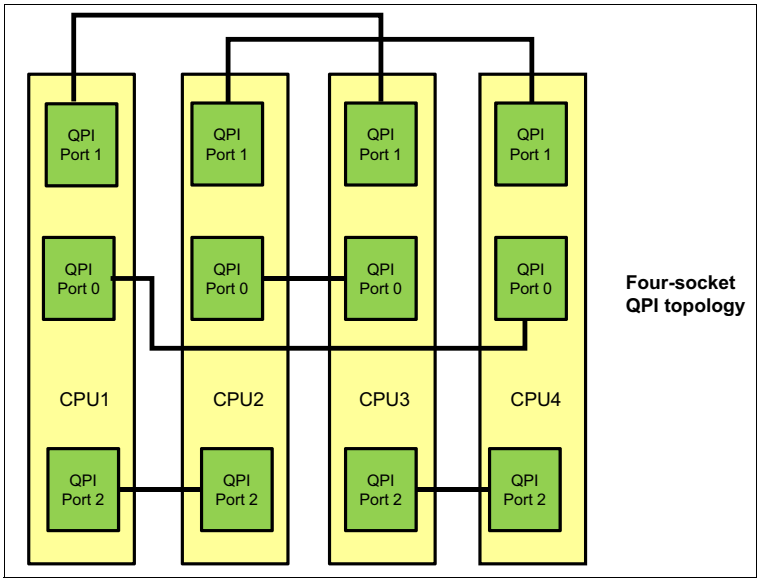


Figure 29 Four-socket system QPI topology

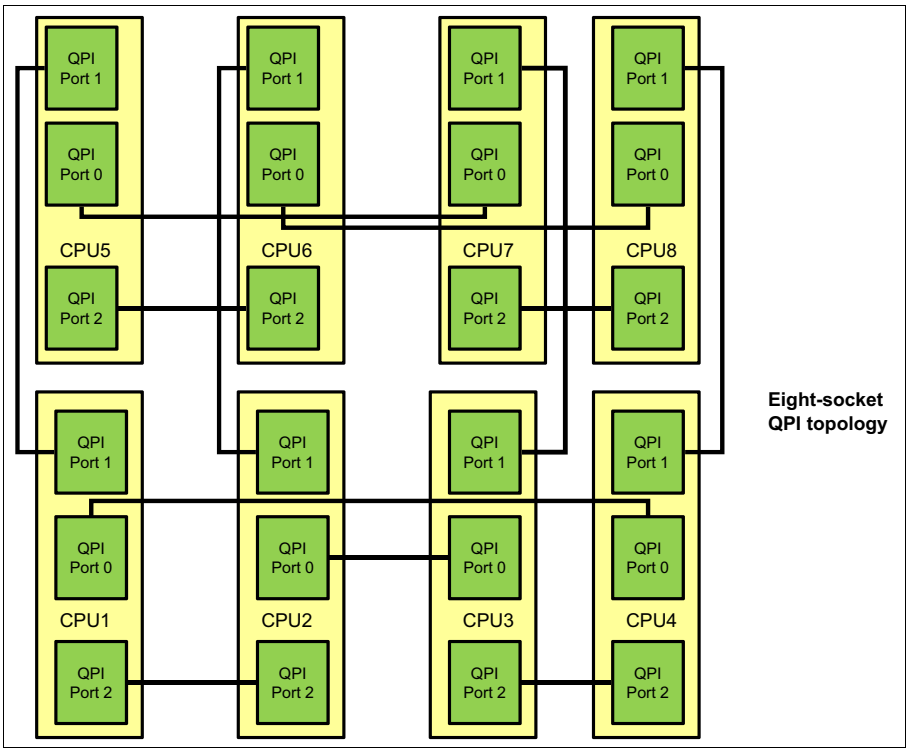


Figure 30 Eight-socket system QPI topology

- When a QPI fails down to half width, the IMM logs two errors messages. For example, “Sensor CPU 1 QPILink has transitioned to critical from a less severe state” and “Sensor CPU 2 QPILink has transitioned to critical from a less severe state.” Figure 34 shows that if the IMM flags CPU1 and CPU2, QPI Port1 has an issue. When the IMM reports this kind of error, the issue can be in one of several places:
  - CPU1 or CPU2
  - CPU socket 1 or 2
  - CPU book 1 or 2
  - Midplane

  IMM/HW cannot detect where the QPI is broken. It can only see that the RX is not in full width. This is why the IMM must report two CPUs when a link goes to half width.

  To isolate which FRU is defective, you need to complete a series of tests:

  a. Check for bent pins on the midplane to avoid a the situation where this one hardware problem can cascade into addition hardware failures. Perform this check before starting the defective FRU isolation process.
  b. Swap one of the failing books with a good book in the same system.
  c. If the IMM still calls out the same two CPU books, swap out the other failing book.
  d. If the failure follows the book, it is the CPU or book.
  e. If the failure follows the CPU book slot, the issue is with the midplane.
  f. After you have isolated the error to a single book, swap CPUs with one of the good books.
  g. If the error follows the CPU, replace the CPU. If it follows the book, replace the book.

To date, most QPI errors have been related to damaged CPU socket pins. To inspect the pins, look at the socket while moving the CPU book to get a view of the sockets from many angles. Look for non-uniform light reflection off the pins. Pin damage can be subtle and hard to see even under a microscope.

Related publications

The following Lenovo Press publications provide additional information about the x3850 X6 and x3950 X6 servers.

- Product documentation (manuals) for System x3850 X6

  http://ibm.com/support/entry/portal/documentation_expanded_list/system_x/system_x3850_x6?productContext=-40720963

- Planning and installation documentation for System x3850 X6

  http://ibm.com/support/entry/portal/plan_install/system_x/system_x3850_x6?productContext=-40720963

- eXFlash DIMM Configuration and Support Requirements

  http://ibm.com/support/entry/portal/docdisplay?lndocid=SERV-FLASHDM

- Lenovo System x3850 X6 and x3950 X6 Planning and Implementation Guide

  http://www.redbooks.ibm.com/abstracts/sg248208.html?Open

- Lenovo ServerProven®

  http://ibm.com/systems/info/x86servers/serverproven/compat/us/xseries/3837.html

- BoMC PXE image fails - System x3850 X6, RETAIN tip: H212116

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5094856

- System hangs after uncorrectable memory error - System x3850 X6, RETAIN tip: H212099

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5094814

- DIMMs may appear as present, not configured under UEFI F1 setup menus - System x3850 X6, RETAIN tip: H212105

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5094838

- Memory status on IMM web does not match web log entries - System x3850 X6, RETAIN tip: H212103

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5094836

- Secure boot enable or disable Recovery - System x3850 X6, RETAIN tip: H212054

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5094723

- The first start after SAN install of VMware in UEFI mode fails on some expansion ports - System x3850 X6, RETAIN tip: H212120

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5094859

- System hang during legacy boot when UEFI install present - System x3850 X6, RETAIN tip: H212104

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5094837

- Abnormal PCI error handling behavior when LER Is disabled - System x3850 X6, RETAIN tip: H212079

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5094752

- Emulex EFI utility 'Back to display Adapter' menu hang - System x3850 X6, RETAIN tip: H212078

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5094751

Authors

This paper was produced by a team of X6 development personnel and field specialists.

John R. Encizo is an Enterprise System Sales Engineer for Lenovo.

Timothy M. Wiwel is a hardware specialist in the Lenovo System x® Server Development team.

Loc X. Nguyen is a hardware engineer in the Lenovo System x Server Development team.

Portions of this paper are from the Lenovo Press publication *IBM System x3850 X6 and x3950 X6 Planning and Implementation Guide*, SG24-8208, written by:

- David Watts
- Rani Doughty
- Ilya Solovyev

Thanks to the following people for their contributions to this project:

- Tisha R Jefferson, Lenovo System x Marketing
- David Watts, Lenovo Press

Now you can become a published author, too!

Here's an opportunity to spotlight your skills, grow your career, and become a published author—all at the same time! Join an ITSO residency project and help write a book in your area of expertise, while honing your experience using leading-edge technologies. Your efforts will help to increase product acceptance and customer satisfaction, as you expand your network of technical contacts and relationships. Residencies run from two to six weeks in length, and you can participate either in person or as a remote resident working from your home base.

Find out more about the residency program, browse the residency index, and apply online at:

ibm.com/redbooks/residencies.html

Notices

Lenovo may not offer the products, services, or features discussed in this document in all countries. Consulty our local Lenovo representative for information on the products and services currently available in yourarea. Any reference to a Lenovo product, program, or service is not intended to state or imply that only that Lenovo product, program, or service may be used. Any functionally equivalent product, program, or service that does not infringe any Lenovo intellectual property right may be used instead. However, it is the user's responsibility to evaluate and verify the operation of any other product, program, or service.

Lenovo may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:

Lenovo (United States), Inc.
1009 Think Place - Building One
Morrisville, NC 27560
U.S.A.
Attention: Lenovo Director of Licensing

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. Some jurisdictions do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically 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.

The products described in this document are not intended for use in implantation or other life support applications where malfunction may result in injury or death to persons. The information contained in this document does not affect or change Lenovo product specifications or warranties. Nothing in this document shall operate as an express or implied license or indemnity under the intellectual property rights of Lenovo or third parties. All information contained in this document was obtained in specific environments and is presented as an illustration. The result obtained in other operating environments may vary.

Lenovo may use or distribute any of the information you supply in any way it believes appropriate without incurring any obligation to you.

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.

Any performance data contained herein was determined in a controlled environment. Therefore, the result 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.

Trademarks

Lenovo, the Lenovo logo, and For Those Who Do are trademarks or registered trademarks of Lenovo in the United States, other countries, or both. These and other Lenovo trademarked terms are marked on their first occurrence in this information with the appropriate symbol (® or ™), indicating US registered or common law trademarks owned by Lenovo at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of Lenovo trademarks is available on the Web at http://www.lenovo.com/legal/copytrade.html.

The following terms are trademarks of Lenovo in the United States, other countries, or both:

Advanced Settings Utility™	Lenovo®	System x®
Bootable Media Creator™	Lenovo(logo)®	
eXFlash™	ServerProven®	

Intel, Intel Xeon, Intel logo, Intel Inside logo, and Intel Centrino logo are trademarks or registered trademarks of Intel Corporation or its subsidiaries in the United States and other countries.

Linux is a trademark of Linus Torvalds in the United States, other countries, or both.

Microsoft, Windows, and the Windows logo are trademarks of Microsoft Corporation in the United States, other countries, or both.

Other company, product, or service names may be trademarks or service marks of others.