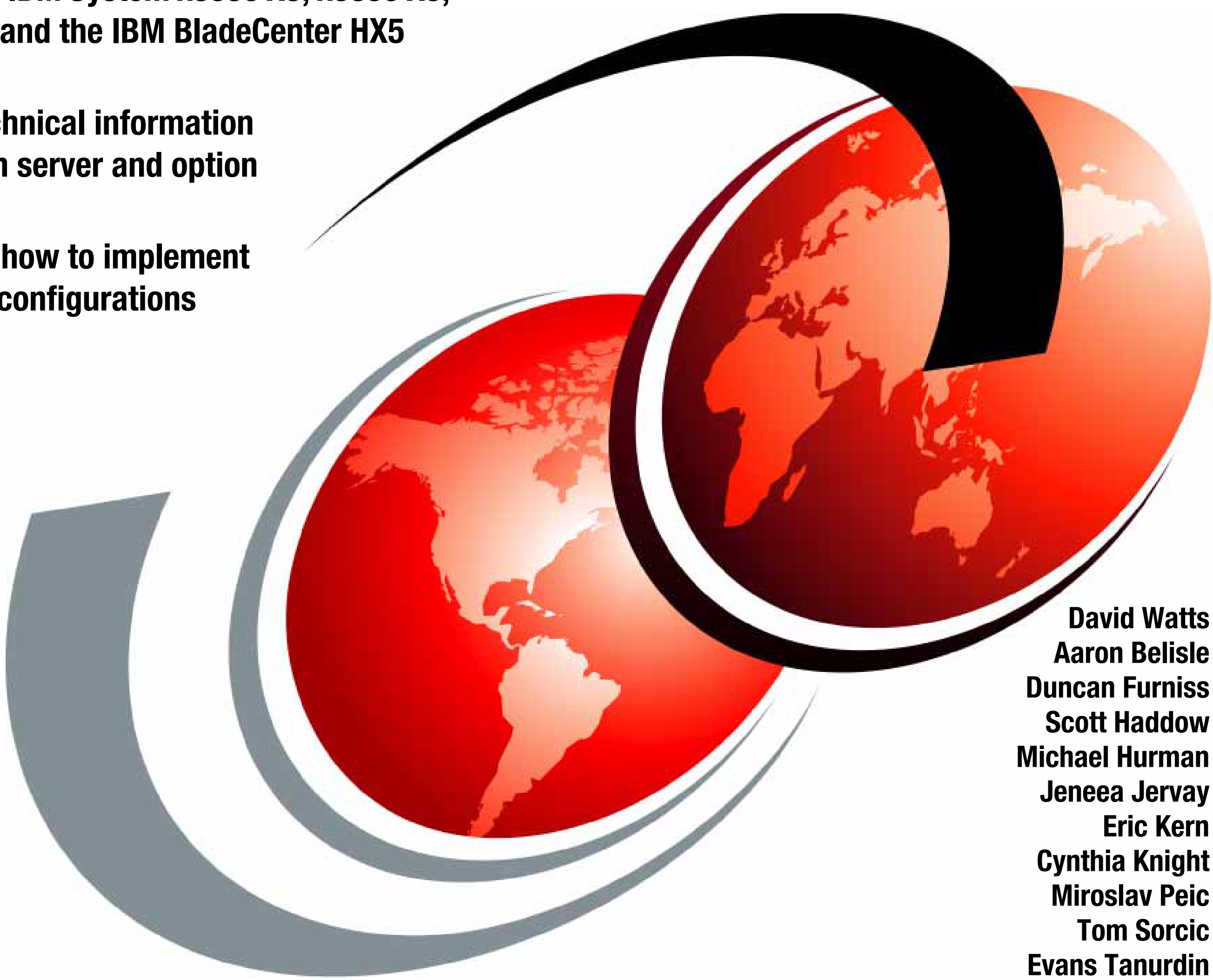
IBM

IBM eX5 Implementation Guide

Covers the IBM System x3950 X5, x3850 X5, x3690 X5, and the IBM BladeCenter HX5

Details technical information about each server and option

Describes how to implement two-node configurations



David Watts
Aaron Belisle
Duncan Furniss
Scott Haddow
Michael Hurman
Jeneea Jervay
Eric Kern
Cynthia Knight
Miroslav Peic
Tom Sorcic
Evans Tanurdin

Redbooks

ibm.com/redbooks

International Technical Support Organization

IBM eX5 Implementation Guide

May 2011

SG24-7909-00

Note: Before using this information and the product it supports, read the information in “Notices” on page xi.

First Edition (May 2011)

This edition applies to the following servers:

- IBM System x3850 X5, machine type 7145
- IBM System x3950 X5, machine type 7145
- IBM System x3690 X5, machine type 7148
- IBM BladeCenter HX5, machine type 7872

Contents

Notices

This information was developed for products and services offered in the U.S.A.

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

IBM 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:
IBM Director of Licensing, IBM Corporation, North Castle Drive, Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION 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 states 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. IBM 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 references in this information to non-IBM websites are provided for convenience only and do not in any manner serve as an endorsement of those websites. The materials at those websites are not part of the materials for this IBM product and use of those websites is at your own risk.

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

Information concerning non-IBM products was obtained from the suppliers of those products, their published announcements or other publicly available sources. IBM has not tested those products and cannot confirm the accuracy of performance, compatibility or any other claims related to non-IBM products. Questions on the capabilities of non-IBM products should be addressed to the suppliers of those products.

This information contains examples of data and reports used in daily business operations. To illustrate them as completely as possible, the examples include the names of individuals, companies, brands, and products. All of these names are fictitious and any similarity to the names and addresses used by an actual business enterprise is entirely coincidental.

COPYRIGHT LICENSE:

This information contains sample application programs in source language, which illustrate programming techniques on various operating platforms. You may copy, modify, and distribute these sample programs in any form without payment to IBM, for the purposes of developing, using, marketing or distributing application programs conforming to the application programming interface for the operating platform for which the sample programs are written. These examples have not been thoroughly tested under all conditions. IBM, therefore, cannot guarantee or imply reliability, serviceability, or function of these programs.

Trademarks

IBM, the IBM logo, and ibm.com are trademarks or registered trademarks of International Business Machines Corporation in the United States, other countries, or both. These and other IBM 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 IBM at the time this information was published. Such trademarks may also be registered or common law trademarks in other countries. A current list of IBM trademarks is available on the Web at http://www.ibm.com/legal/copytrade.shtml

The following terms are trademarks of the International Business Machines Corporation in the United States, other countries, or both:

AIX®
BladeCenter®
Calibrated Vectored Cooling™
DS4000®
Dynamic Infrastructure®
Electronic Service Agent™
eServer™
IBM Systems Director Active Energy Manager™
IBM®
iDataPlex™
Netfinity®
PowerPC®
POWER®
Redbooks®
Redpaper™
Redbooks (logo) ®
RETAIN®
ServerProven®
Smarter Planet™
System Storage®
System x®
System z®
Tivoli®
X-Architecture®
XIV®
xSeries®

The following terms are trademarks of other companies:

Java, and all Java-based trademarks are trademarks of Sun Microsystems, Inc. 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.

Intel Xeon, Intel, Itanium, 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.

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

Preface

High-end workloads drive ever-increasing and ever-changing constraints. In addition to requiring greater memory capacity, these workloads challenge you to do more with less and to find new ways to simplify deployment and ownership. And although higher system availability and comprehensive systems management have always been critical, they have become even more important in recent years.

Difficult challenges, such as these, create new opportunities for innovation. The IBM® eX5 portfolio delivers this innovation. This family of high-end computing introduces the fifth generation of IBM X-Architecture® technology. The family includes the IBM System x3850 X5, x3690 X5, and the IBM BladeCenter® HX5. These servers are the culmination of more than a decade of x86 innovation and firsts that have changed the expectations of the industry. With this latest generation, eX5 is again leading the way as the shift toward virtualization, platform management, and energy efficiency accelerates.

This book is divided into two parts. In the first part, we provide detailed technical information about the servers in the eX5 portfolio. This information is most useful in designing, configuring, and planning to order a server solution. In the second part of the book, we provide detailed configuration and setup information to get your servers operational. We focus particularly on setting up MAX5 configurations of all three eX5 servers as well as 2-node configurations of the x3850 X5 and HX5.

This book is aimed at clients, IBM Business Partners, and IBM employees that want to understand the features and capabilities of the IBM eX5 portfolio of servers and want to learn how to install and configure the servers for use in production.

The team who wrote this book

This book was produced by a team of specialists from around the world working at the International Technical Support Organization, Raleigh Center.

David Watts is a Consulting IT Specialist at the IBM ITSO Center in Raleigh. He manages residencies and produces IBM Redbooks® publications for hardware and software topics that are related to IBM System x® and IBM BladeCenter servers, and associated client platforms. He has authored over 80 books, papers, and web documents. He holds a Bachelor of Engineering degree from the University of Queensland (Australia) and has worked for IBM both in the US and Australia since 1989. David is an IBM Certified IT Specialist and a member of the IT Specialist Certification Review Board.

Aaron Belisle is a BladeCenter and System x Technical Support Specialist for IBM in Atlanta, Georgia. He has 12 years of experience working with servers and has worked at IBM for seven years. His areas of expertise include IBM BladeCenter, System x, and BladeCenter Fibre Channel fabrics.

Duncan Furniss is a Senior IT Specialist for IBM in Canada. He currently provides technical sales support for System x, BladeCenter, and IBM System Storage® products. He has co-authored six previous IBM Redbooks publications, the most recent being *Implementing an IBM System x iDataPlex Solution*, SG24-7629. He has helped clients design and implement x86 server solutions from the beginning of the IBM Enterprise X-Architecture initiative. He is an IBM Regional Designated Specialist for Linux®, High Performance Compute Clusters, and

Rack, Power and Cooling. He is an IBM Certified IT Specialist and member of the IT Specialist Certification Review Board.

Scott Haddow is a Presales Technical Support Specialist for IBM in the UK. He has 12 years of experience working with servers and storage. He has worked at IBM for six years, his experience spanning IBM Netfinity®, xSeries®, and now the System x brand. His areas of expertise include Fibre Channel fabrics.

Michael Hurman is a Senior IT Specialist for IBM STG Lab Services in South Africa. He has more than 12 years of international experience in IT and has co-authored previous IBM Redbooks publications including *Implementing the IBM BladeCenter S Chassis*, SG24-7682. His areas of expertise include assisting clients to design and implement System x, BladeCenter, IBM Systems Director, midrange storage and storage area networks (SAN_-based solutions. He started his career at IBM in 2006.

Jeneea Jervay (JJ) was a Technical Support Management Specialist in Raleigh at the time of writing this publication. She provided presales technical support to IBM Business Partners, clients, IBM Advanced Technical Support specialists, and IBM Field Technical Sales Support Specialists globally for the BladeCenter portfolio. She authored the *IBM BladeCenter Interoperability Guide* from 2007 to early 2010. She is a PMI Certified Project Manager and former System x and BladeCenter Top Gun instructor. She was the lead for the System x and BladeCenter Demand Acceleration Units (DAU) program. Previously, she was a member of the Americas System x and BladeCenter Brand team and the Sales Solution Center, which focused exclusively on IBM Business Partners. She started her career at IBM in 1995.

Eric Kern is a Senior Managing Consultant for IBM STG Lab Services. He currently provides technical consulting services for System x, BladeCenter, System Storage, and Systems Software. Since 2007, he has helped clients design and implement x86 server and systems management software solutions. Prior to joining Lab Services, he developed software for the BladeCenter's Advanced Management Module and for the Remote Supervisor Adapter II. He is a VMware Certified Professional and a Red Hat Certified Technician.

Cynthia Knight is an IBM Hardware Design Engineer in Raleigh and has worked for IBM for 11 years. She is currently a member of the IBM eX5 design team. Previous designs include the Ethernet add-in cards for the IBM Network Processor Reference Platform and the Chassis Management Module for BladeCenter T. She was also the lead designer for the IBM BladeCenter PCI Expansion Units.

Miroslav Peic is a System x Support Specialist in IBM Austria. He has a graduate degree in applied computer science and has many industry certifications, including the Microsoft® Certified Systems Administrator 2003. He trains other IBM professionals and provides technical support to them, as well as to IBM Business Partners and clients. He has 10 years experience in IT and has worked at IBM since 2008.

Tom Sorcic is an IT specialist and technical trainer for BladeCenter and System x support. He is part of Global Technology Enterprise Services at the Intel® Smart Center in Atlanta, Georgia, where he started working for IBM in 2001. He has 37 years of international experience with IT in banking, manufacturing, and technical support. An author of hundreds of web pages, he continues his original role as core team member for the Global System x Skills Exchange (GLOSSE) website, assisting in the site design and providing technical content on a wide variety of topics since 2008. He is a subject matter expert in all forms of IBM ServeRAID hardware, Ethernet networking, storage area networks, and Microsoft high availability clusters.

Evans Tanurdin is an IT Specialist at IBM Global Technology Services in Indonesia. He provides technical support and services on the IBM System x, BladeCenter, and System Storage product lines. His technology focus areas include the design, operation, and maintenance services of enterprise x86 server infrastructure. Other significant experiences include application development, system analysis, and database design. Evans holds a degree in Nuclear Engineering from Gadjah Mada University (Indonesia), and certifications from Microsoft, Red Hat, and Juniper.

The authors of this book were divided into two teams, Part 1 of the book is based on the IBM Redpaper™ *IBM eX5 Portfolio Overview: IBM System x3850 X5, x3950 X5, x3690 X5, and BladeCenter HX5*, REDP-4650, and written by one team of subject matter experts.

The team that wrote Part 1 (left to right): David, Duncan, JJ, Scott, Cynthia, and Eric

Part 2 of the book was written by a second team of subject matter experts. This team also provided updates to the first part of the book.

The team that wrote Part 2 (left to right): David, Evans, Aaron, Miro, Tom, and Mike

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

From IBM Marketing:

- Mark Chapman
- Michelle Gottschalk
- Harsh Kachhy

- Richard Mancini
- Tim Martin
- Kevin Powell
- Heather Richardson
- David Tareen

From IBM Development:

- Justin Bandholz
- Ralph Begun
- Jon Bitner
- Charles Clifton
- Candice Coletrane-Pagan
- David Drez
- Royce Espy
- Dustin Fredrickson
- Larry Grasso
- Dan Kelaher
- Randy Kolvick
- Chris LeBlanc
- Mike Schiskey
- Greg Sellman
- Mehul Shah
- Matthew Trzyna
- Matt Weber

From the IBM Redbooks organization:

- Mary Comianos
- Linda Robinson
- Stephen Smith

From other IBM employees throughout the world:

- Randall Davis, IBM Australia
- John Encizo, IBM U.S.
- Shannon Meier, IBM U.S.
- Keith Ott, IBM U.S.
- Andrew Spurgeon, IBM New Zealand
- Xiao Jun Wu, IBM China

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 client 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

Comments welcome

Your comments are important to us!

We want our books to be as helpful as possible. Send us your comments about this book or other IBM Redbooks publications in one of the following ways:

- Use the online **Contact us** review Redbooks form found at:

  ibm.com/redbooks

- Send your comments in an email to:

  redbooks@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. HYTD Mail Station P099
  2455 South Road
  Poughkeepsie, NY 12601-5400

Stay connected to IBM Redbooks

- Find us on Facebook:

  http://www.facebook.com/IBMRedbooks

- Follow us on Twitter:

  http://twitter.com/ibmredbooks

- Look for us on LinkedIn:

  http://www.linkedin.com/groups?home=&gid=2130806

- Explore new Redbooks publications, residencies, and workshops with the IBM Redbooks weekly newsletter:

  https://www.redbooks.ibm.com/Redbooks.nsf/subscribe?OpenForm

- Stay current on recent Redbooks publications with RSS Feeds:

  http://www.redbooks.ibm.com/rss.html

1

Introduction

The IBM eX5 product portfolio represents the fifth generation of servers built upon Enterprise X-Architecture. Enterprise X-Architecture is the culmination of bringing generations of IBM technology and innovation derived from our experience in high-end enterprise servers. Now with eX5, IBM scalable systems technology for Intel processor-based servers has also been delivered to blades. These servers can be expanded on demand and configured by using a building block approach that optimizes system design servers for your workload requirements.

As a part of the IBM Smarter Planet™ initiative, our Dynamic Infrastructure® charter guides us to provide servers that improve service, reduce cost, and manage risk. These servers scale to more CPU cores, memory, and I/O than previous systems, enabling them to handle greater workloads than the systems they supersede. Power efficiency and machine density are optimized, making them affordable to own and operate.

The ability to increase the memory capacity independently of the processors means that these systems can be highly utilized, yielding the best return from your application investment. These systems allow your enterprise to grow in processing, I/O, and memory dimensions, so that you can provision what you need now, and expand the system to meet future requirements. System redundancy and availability technologies are more advanced than the technologies that were previously available in the x86 systems.

This chapter contains the following topics:

1.1 eX5 systems

The four systems in the eX5 family are the x3850 X5, x3950 X5, x3690 X5, and the HX5 blade. The eX5 technology is primarily designed around three major workloads: database servers, server consolidation using virtualization services, and Enterprise Resource Planning (application and database) servers. Each system can scale with additional memory by adding an IBM MAX5 memory expansion unit to the server, and the x3850 X5, x3950 X5, and HX5 can also be scaled by connecting two systems to form a 2-node scale.

Figure 1-1 shows the IBM eX5 family.

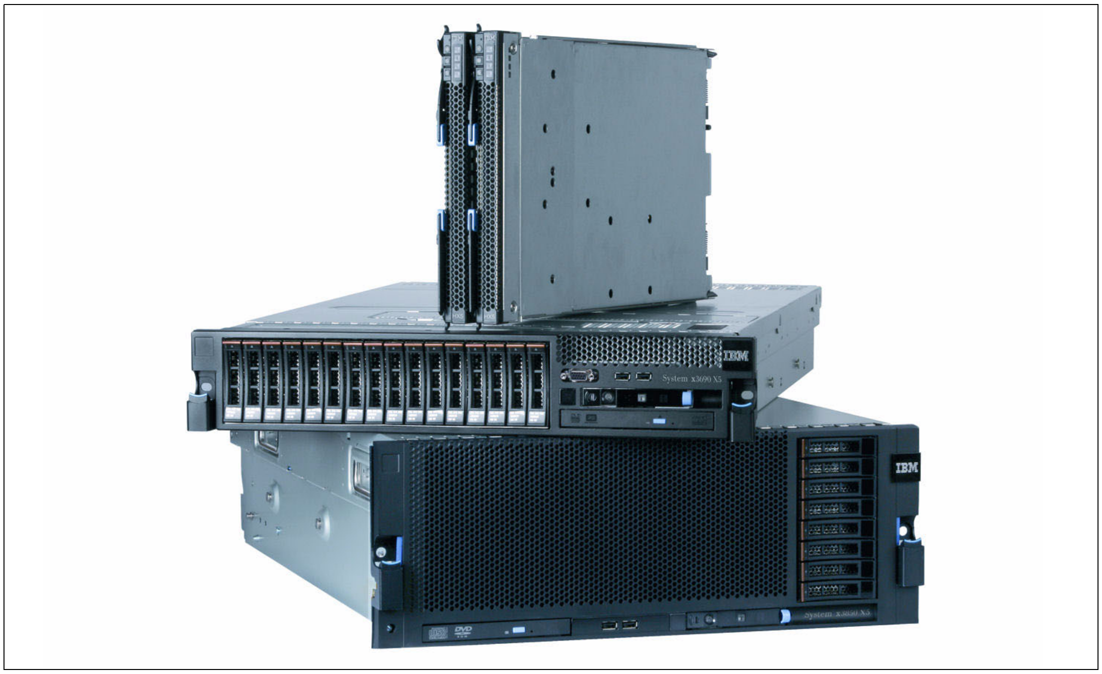


Figure 1-1 eX5 family (top to bottom): BladeCenter HX5 (2-node), System x3690 X5, and System x3850 X5 (the System x3950 X5 looks the same as the x3850 X5)

The IBM System x3850 X5 and x3950 X5 are 4U highly rack-optimized servers. The x3850 X5 and the workload-optimized x3950 X5 are the new flagship servers of the IBM x86 server family. These systems are designed for maximum utilization, reliability, and performance for computer-intensive and memory-intensive workloads.

The IBM System x3690 X5 is a new 2U rack-optimized server. This machine brings new features and performance to the middle tier, as well as a memory scalability option with MAX5.

The IBM BladeCenter HX5 is a single-wide (30 mm) blade server that follows the same design as all previous IBM blades. The HX5 brings unprecedented levels of capacity to high-density environments. The HX5 is expandable to form either a 2-node system with four processors, or a single-node system with the MAX5 memory expansion blade.

When compared to other machines in the System x portfolio, these systems represent the upper end of the spectrum, are suited for the most demanding x86 tasks, and can handle jobs which previously might have been run on other platforms. To assist with selecting the ideal system for a given workload, we have designed workload-specific models for virtualization and database needs.

1.2 Model summary

This section summarizes the models that are available for each of the eX5 systems.

1.2.1 IBM System x3850 X5 models

Table 1-1 lists the standard x3850 X5 models.

Table 1-1 Base models of the x3850 X5: Four socket-scalable server

Model[a]	Intel Xeon® processors (two standard; maximum of four)	Memory speed (MHz)	Standard memory (MAX5 is optional)	Memory cards (std/max)	ServeRAID BR10i std	10Gb Ethernet standard[b]	Power supplies (std/max)	Drive bays (std/max)
7145-ARx	E7520 4C 1.86 GHz, 18 MB L3, 95W [c]	800 MHz	2x 2 GB	1/8	No	No	1/2	None
7145-1Rx	E7520 4C 1.86 GHz, 18 MB L3, 95W [c]	800 MHz	4x 4 GB	2/8	Yes	Yes	2/2	4x 2.5"/8
7145-2Rx	E7530 6C 1.86 GHz,12 MB L3, 105W [c]	978 MHz	4x 4 GB	2/8	Yes	Yes	2/2	4x 2.5"/8
7145-3Rx	E7540 6C 2.0 GHz, 18 MB L3, 105W	1066 MHz	4x 4 GB	2/8	Yes	Yes	2/2	4x 2.5"/8
7145-4Rx	X7550 8C 2.0 GHz, 18 MB L3, 130W	1066 MHz	4x 4 GB	2/8	Yes	Yes	2/2	4x 2.5"/8
7145-5Rx	X7560 8C 2.27 GHz, 24 MB L3, 130W	1066 MHz	4x 4 GB	2/8	Yes	Yes	2/2	4x 2.5"/8

a. The x character in the seventh position of the machine model denotes the region-specific character. For example, U indicates US, and G indicates EMEA.
b. Emulex 10Gb Ethernet Adapter is installed in PCIe slot 7.
c. Any model using the E7520 or E7530 CPU cannot scale beyond single-node 4-way, even with the addition of MAX5.

1.2.2 Workload-optimized x3950 X5 models

Table 1-2 on page 4 lists the workload-optimized models of the x3950 X5 that have been announced. The MAX5 is optional on these models. (In the table, *std* is standard, and *max* is maximum.)

Model 5Dx

Model 5Dx is designed for database applications and uses solid-state drives (SSDs) for the best I/O performance. Backplane connections for eight 1.8-inch SSDs are standard and there is space for an additional eight SSDs. The SSDs themselves must be ordered separately. Because no SAS controllers are standard, you can select from the available cards as described in 3.9, "Storage" on page 90.

Model 4Dx

Model 4Dx is designed for virtualization and is fully populated with 4 GB memory dual inline memory modules (DIMMs), including in an attached MAX5 memory expansion unit, a total of 384 GB of memory. Backplane connections for four 2.5-inch serial-attached SCSI (SAS) hard disk drives (HDDs) are standard; however, the SAS HDDs themselves must be ordered separately. A ServeRAID BR10i SAS controller is standard in this model.

Table 1-2 Models of the x3950 X5: Workload-optimized models

Model[a]	Intel Xeon processors (two standard, maximum of four)	Memory speed	MAX5	Standard memory	Memory cards (std/max)	ServeRAID BR10i std	10Gb Ethernet standard[b]	Power supplies (std/max)	Drive bays (std/max)
Database workload-optimized models									
7145-5Dx	X7560 8C 2.27 GHz, 24 MB L3, 130W	1066 MHz	Opt	Server: 8x 4GB	4/8	No	Yes	2/2	8x 1.8"/16[c]
Virtualization workload-optimized models									
7145-4Dx	4x X7550 8C 2.0 GHz, 18 MB L3, 130W	1066 MHz	Std	Server: 64x 4GB MAX5: 32x 4GB	8/8	Yes	Yes	2/2	4x 2.5"/8

a. The x character in the seventh position of the machine model denotes the region-specific character. For example, U indicates US, and G indicates EMEA.
b. Emulex 10Gb Ethernet Adapter is installed in PCIe slot 7.
c. Includes, as standard, one 8-bay eXFlash SSD backplane; one additional eXFlash backplane is optional.

1.2.3 x3850 X5 models with MAX5

Table 1-3 lists the models that are standard with the 1U MAX5 memory expansion unit.

Table 1-3 Models of the x3850 X5 with the MAX5 standard

Model[a]	Intel Xeon processors (four standard and max)	Memory speed (MHz)	Standard memory (MAX5 is standard)	Memory cards (std/max)	ServeRAID BR10i std	10Gb Ethernet standard[b]	Power supplies (std/max)	Drive bays (std/max)
7145-2Sx	4x E7530 6C 1.86 GHz, 12 MB L3, 105W [c]	978 MHz	Server: 8x 4 GB MAX5: 2x 4 GB	4/8	Yes	Yes	2/2	4x 2.5"/8
7145-4Sx	4x X7550 8C 2.0 GHz, 18 MB L3, 130W	1066 MHz	Server: 8x 4 GB MAX5: 2x 4 GB	4/8	Yes	Yes	2/2	4x 2.5"/8
7145-5Sx	4x X7560 8C 2.27 GHz, 24 MB L3, 130W	1066 MHz	Server: 8x 4 GB MAX5: 2x 4 GB	4/8	Yes	Yes	2/2	4x 2.5"/8

a. The x character in the seventh position of the machine model denotes the region-specific character. For example, U indicates US, and G indicates EMEA.
b. Emulex 10Gb Ethernet Adapter is installed in PCIe slot 7.
c. Any model using the E7520 or E7530 CPU cannot scale beyond single-node 4-way, even with the addition of MAX5.

1.2.4 Base x3690 X5 models

Table 1-4 on page 5 provides the standard models of the x3690 X5. The MAX5 memory expansion unit is standard on specific models as indicated.

Table 1-4 x3690 X5 models

Model	Intel Xeon processors (two maximum)	Memory speed	MAX5	Standard memory[a]	Memory tray	ServeRAID M1015 standard	10Gb Ethernet standard[b]	Power supplies std/max	Drive bays std/max
7148-ARx	1x E7520 4C, 1.86 GHz, 95W	800 MHz	Opt	Server: 2x 4GB	Opt	Opt	Opt	1/4	None
7148-1Rx	1x E7520 4C, 1.86 GHz, 95W	800 MHz	Opt	Server: 2x 4GB	Opt	Std	Opt	1/4	4x 2.5”/16
7148-2Rx	1x E6540 6C, 2.00 GHz, 105W	1066 MHz	Opt	Server: 2x 4GB	Opt	Std	Opt	1/4	4x 2.5”/16
7148-3Rx	1x X6550 8C, 2.00 GHz, 130W	1066 MHz	Opt	Server: 2x 4GB	Opt	Std	Opt	1/4	4x 2.5”/16
7148-3Gx	1x X6550 8C, 2.00 GHz, 130W	1066 MHz	Opt	Server: 2x 4GB	Opt	Std	Std	1/4	4x 2.5”/16
7148-4Rx	1x X7560 8C, 2.26 GHz, 130W	1066 MHz	Opt	Server: 2x 4GB	Opt	Std	Opt	1/4	4x 2.5”/16
7148-3Sx	1x X7550 8C, 2.00GHz, 130W	1066 MHz	Std	Server: 2x 4GB MAX5: 2x 4GB	Opt	Std	Opt	Server: 2/4 MAX5: 1/2	4x 2.5”/16
7148-4Sx	1x X7560 8C, 2.26GHz, 130W	1066 MHz	Std	Server: 2x 4GB MAX5: 2x 4GB	Opt	Std	Opt	Server: 2/4 MAX5: 1/2	4x 2.5”/16

a. Up to 64 DIMM sockets: Each server has 16 DIMM sockets standard or 32 sockets with the addition of the internal memory tray (mezzanine). With the addition of the MAX5 memory expansion unit, 64 DIMM sockets in total are available.

b. Emulex 10Gb Ethernet Adapter.

1.2.5 Workload-optimized x3690 X5 models

Table 1-5 on page 6 lists the workload-optimized models.

Model 3Dx is designed for database applications and uses SSDs for the best I/O performance. Backplane connections for sixteen 1.8-inch solid-state drives are standard and there is space for an additional 16 solid-state drives. You must order the SSDs separately. No SAS controllers are standard, which lets you select from the available cards, as described in 4.9, “Storage” on page 145. The MAX5 is optional on this model.

Model 2Dx is designed for virtualization applications and includes VMware ESXi 4.1 on an integrated USB memory key. The server is fully populated with 4 GB memory DIMMs, including those in an attached MAX5 memory expansion unit, for a total of 256 GB of memory. Backplane connections for four 2.5-inch SAS drives are standard and there is space for an additional twelve 2.5-inch disk drives. You must order the drives separately. See 4.9, “Storage” on page 145.

Table 1-5 x3690 X5 workload-optimized models

Model	Intel Xeon processors (two maximum)	Memory speed	MAX5	Standard memory[a]	Memory tray	ServeRAID M1015 std	10Gb Eth standard[b]	Power supplies std/max	Drive bays std/max
Database workload-optimized models									
7148-3Dx	2x X6550 8C, 2.00 GHz, 130W	1066 MHz	Opt	Server: 4x 4 GB	Std	Opt	Opt	Server: 4/4	16x 1.8”/32
Virtualization workload-optimized models									
7148-2Dx	2x E6540 6C, 2.00 GHz, 105W	1066 MHz	Std	Server: 32x 4GB MAX5: 32x 4GB	Std	Opt	Std	Server: 4/4 MAX5: 2/2	4x 2.5”/16

a. Up to 64 DIMM sockets: Each server has 16 DIMM sockets standard or 32 sockets with the addition of the internal memory tray (mezzanine). With the addition of the MAX5 memory expansion unit, a total of 64 DIMM sockets are available.

b. Emulex 10Gb Ethernet Adapter.

1.2.6 BladeCenter HX5 models

Table 1-6 shows the base models of the BladeCenter HX5, with and without the MAX5 memory expansion blade. In the table, *Opt* indicates optional and *Std* indicates standard.

Table 1-6 Models of the HX5

Model[a]	Intel Xeon model and cores/max	Clock speed	TDP	HX5 max memory speed	MAX5 memory speed	MAX5	Scalable to four socket	10 GbE card[b]	Standard memory[c]
7872-42x	1x E7520 4C/2	1.86 GHz	95W	800 MHz	800 MHz	Opt	Yes	Opt	2x 4 GB
7872-82x	1x L7555 8C/2	1.86 GHz	95W	978 MHz	978 MHz	Opt	Yes	Opt	2x 4 GB
7872-61x	1x E7530 6C/2	1.86 GHz	105W	978 MHz	978 MHz	Opt	Yes	Opt	2x 4 GB
7872-64x	1x E7540 6C/2	2.00 GHz	105W	978 MHz	1066 MHz	Opt	Yes	Opt	2x 4 GB
7872-65x	1x E7540 6C/2	2.00 GHz	105W	978 MHz	1066 MHz	Opt	Yes	Std	2x 4 GB
7872-63x	2x E6540 6C/2	2.00 GHz	105W	978 MHz	1066 MHz	Std	No	Opt	HX5: 4x 4GB MAX5: None
7872-6Dx	2x E6540 6C/2	2.00 GHz	105W	978 MHz	1066 MHz	Std	No	Std	HX5: 4x 4GB MAX5: None
7872-83x	2x X6550 8C/2	2.00 GHz	130W	978 MHz	1066 MHz	Std	No	Opt	HX5: 4x 4GB MAX5: None
7872-84x	2x X7560 8C/2	2.26 GHz	130W	978 MHz	1066 MHz	Std	No	Opt	HX5: 4x 4GB MAX5: None
7872-86x	1x X7560 8C/2	2.26 GHz	130W	978 MHz	1066 MHz	Opt	Yes	Std	2x 4 GB

a. This column lists worldwide, generally available variant (GAV) model numbers. They are not orderable as listed and must be modified by country. The US GAV model numbers use the following nomenclature: *xx*U. For example, the US orderable part number for 7870-A2*x* is 7870-A2U. See the product-specific official IBM announcement letter for other country-specific GAV model numbers.

b. Emulex Virtual Fabric Adapter Expansion Card (CFFh)

c. The HX5 has 16 DIMM sockets and can hold 128 GB using 8 GB memory DIMMs. The MAX5 has 24 DIMM sockets and can hold 192 GB using 8 GB memory DIMMs. A 1-node HX5 + MAX5 supports 320 GB total using 8 GB DIMMs.

Also available is a virtualization workload-optimized model of these HX5s. This is a pre-configured, pre-tested model targeted at large-scale consolidation. Table 1-7 shows the model.

Table 1-7 Workload-optimized models of the HX5

Model	Intel Xeon model and cores/max	Clock speed	TDP	HX5 max memory speed[a]	MAX5	Scalable to four socket	10GbE card[b]	Standard memory (max 320 GB)[c]
Virtualization workload-optimized models (includes VMware ESXi 4.1 on a USB memory key)								
7872-68x	2x E6540 6C/2	2.00 GHz	105 W	978 MHz	Std	No	Std	160 GB HX5: 16x 4GB MAX5: 24x 4GB

a. Memory speed of the HX5 is dependent on the processor installed; however, the memory speed of the MAX5 is up to 1066 MHz irrespective of the processor installed in the attached HX5.
b. Emulex Virtual Fabric Adapter Expansion Card (CFFh).
c. HX5 has 16 DIMM sockets and can hold 128 GB using 8 GB memory DIMMs. MAX5 has 24 DIMM sockets and can hold 192 GB using 8 GB memory DIMMs. A 1-node HX5 + MAX5 supports 320 GB total using 8 GB DIMMs.

Model 7872-68x is a virtualization-optimized model and includes the following features in addition to standard HX5 and MAX5 features:

- Forty DIMM sockets, all containing 4 GB memory DIMMs for a total of 160 GB of available memory.
- VMware ESXi 4.1 on a USB memory key is installed internally in the server. See 5.15, “Integrated virtualization” on page 214 for details.
- Emulex Virtual Fabric Adapter Expansion Card (CFFh).

1.3 Positioning

Table 1-8 gives an overview of the features of the systems that are described in this book.

Table 1-8 Maximum configurations for the eX5 systems

Maximum configurations		x3850 X5/x3950 X5	x3690 X5	HX5
Processors	1-node	4	2	2
	2-node	8	Not available	4
Memory	1-node	1024 GB (64 DIMMs)[a]	512 GB (32 DIMMs)[b]	128 GB (16 DIMMs)
	1-node with MAX5	1536 GB (96 DIMMs)[a]	1024 GB (64 DIMMs)[b]	320 GB (40 DIMMs)
	2-node	2048 GB (128 DIMMs[a]	Not available	256 GB (32 DIMMs)
Disk drives (non-SSD)[c]	1-node	8	16	Not available
	2-node	16	Not available	Not available
SSDs	1-node	16	24	2
	2-node	32	Not available	4
Standard 1 Gb Ethernet interfaces	1-node	2[d]	2	2
	2-node	4	Not available	4

Maximum configurations		x3850 X5/x3950 X5	x3690 X5	HX5
Standard 10 Gb Ethernet interface	1-node	2	2	0
	2-node	4	Not available	0

a. Requires full processors in order to install and use all memory.
b. Requires that the memory mezzanine board is installed along with processor 2.
c. For the x3690 X5 and x3850 X5, additional backplanes might be needed to support these numbers of drives.
d. Depends on the model. See Table 3-2 on page 64 for the IBM System x3850 X5.

1.3.1 IBM System x3850 X5 and x3950 X5

The System x3850 X5 and the workload-optimized x3950 X5 are the logical successors to the x3850 M2 and x3950 M2. The x3850 X5 and x3950 X5 both support up to four processors and 1.024 TB (terabyte) of RAM in a single-node environment.

The x3850/x3950 X5 with the MAX5 memory expansion unit attached, as shown in Figure 1-2, can add up to an additional 512 GB of RAM for a total of 1.5 TB of memory.

Figure 1-2 IBM System x3850/x3950 X5 with the MAX5 memory expansion unit attached

Two x3850/x3950 X5 servers can be connected for a single system image with a max of eight processors and 2 TB of RAM.

Table 1-9 compares the number of processor sockets, cores, and memory capacity of the eX4 and eX5 systems.

Table 1-9 Comparing the x3850 M2 and x3950 M2 with the eX5 servers

	Processor sockets	Processor cores	Maximum memory
Previous generation servers (eX4)			
x3850 M2	4	24	256 GB
x3950 M2	4	24	256 GB

	Processor sockets	Processor cores	Maximum memory
x3950 M2 2-node	8	48	512 GB
Next generation server (eX5)			
x3850/x3950 X5	4	32	1024 GB
x3850/x3950 X5 2-node	8	64	2048 GB
x3850/x3950 X5 with MAX5	4	32	1536 GB

1.3.2 IBM System x3690 X5

The x3690 X5, as shown on Figure 1-3, is a 2-processor server that exceeds the capabilities of the current mid-tier server, the x3650 M3. You can configure the x3690 X5 with processors that have more cores and more cache than the x3650 M3. You can configure the x3690 X5 with up to 512 GB of RAM, whereas the x3650 M3 has a maximum memory capacity of 144 GB.

Figure 1-3 x3690 X5

Table 1-10 compares the processing and memory capacities.

Table 1-10 x3650 M3 compared to x3690 X5

	Processor sockets	Processor cores	Maximum memory
Previous generation server			
x3650 M3	2	12	144 GB
Next generation server (eX5)			
x3690 X5	2	16	512 GB[a]
x3690 with MAX5	2	16	1024 GB[a]

a. You must install two processors and the memory mezzazine to use the full memory capacity.

1.3.3 IBM BladeCenter HX5

The IBM Blade Center HX5, as shown in Figure 1-4 on page 10 with the second node attached, is a blade that exceeds the capabilities of the previous system HS22. The HS22V has more memory in a single-wide blade, but the HX5 can be scaled by adding another HX5 or by adding a MAX5 memory expansion blade.

Figure 1-4 Blade HX 5 dual scaled

Table 1-11 compares these blades.

Table 1-11 HS22, HS22V, and HX5 compared

	Processor sockets	Processor cores	Maximum memory
Comparative servers			
HS22 (30 mm)	2	12	192 GB
HS22V (30 mm)	2	12	288 GB
Next generation server (eX5)			
HX5 (30 mm)	2	16	128 GB
HX5 2-node (60 mm)	4	32	256 GB
HX5 with MAX5	2	16	320 GB

1.4 Energy efficiency

We put extensive engineering effort into keeping your energy bills low - from high-efficiency power supplies and fans to lower-draw processors, memory, and SSDs. We strive to reduce the power consumed by the systems to the extent that we include altimeters, which are capable of measuring the density of the atmosphere in the servers and then adjusting the fan speeds accordingly for optimal cooling efficiency.

Technologies, such as these altimeters, along with the Intel Xeon 7500/6500 series processors that intelligently adjust their voltage and frequency, help take costs out of IT:

- 95W 8-core processors use 27% less energy than 130W processors.
- 1.5V DDR3 DIMMs consume 10-15% less energy than the DDR2 DIMMs that were used in older servers.
- SSDs consume up to 80% less energy than 2.5-inch HDDs and up to 88% less energy than 3.5-inch HDDs.

- Dynamic fan speeds: In the event of a fan failure, the other fans run faster to compensate until the failing fan is replaced. Regular fans must run faster at all times, just in case, thereby wasting power.

Although these systems provide incremental gains at the individual server level, the eX5 systems can have an even greater green effect in your data center. The gain in computational power and memory capacity allows for application performance, application consolidation, and server virtualization at greater degrees than previously available in x86 servers.

1.5 Services offerings

The eX5 systems fit into the services offerings that are already available from IBM Global Technology Services for System x and BladeCenter. More information about these services is available at the following website:

http://www.ibm.com/systems/services/gts/systemxbcis.html

In addition to the existing offerings for asset management, information infrastructure, service management, security, virtualization and consolidation, and business and collaborative solutions; IBM Systems Lab Services and Training has six offerings specifically for eX5:

- Virtualization Enablement
- Database Enablement
- Enterprise Application Enablement
- Migration Study
- Virtualization Health Check
- Rapid! Migration Tool

IBM Systems Lab Services and Training consists of highly skilled consultants that are dedicated to help you accelerate the adoption of new products and technologies. The consultants use their relationships with the IBM development labs to build deep technical skills and use the expertise of our developers to help you maximize the performance of your IBM systems. The services offerings are designed around having the flexibility to be customized to meet your needs.

For more information, send email to this address:

mailto:stgls@us.ibm.com

Also, more information is available at the following website:

http://www.ibm.com/systems/services/labservices

1.6 What this book contains

In this book, readers get a general understanding of eX5 technology, what sets it apart from previous models, and the architecture that makes up this product line. This book is broken down into two main parts:

- Part One gives an in-depth look at specific components, such as memory, processors, storage, and a general breakdown for each model.
- Part Two describes implementing the servers, in particular the 2-node and MAX5 configurations. We also describe systems management, firmware update tools, and methods for performing system firmware updates. We also describe the detection of the most common failures and recovery scenarios for each situation.

Part 1

Product overview

In this first part of the book, we provide detailed technical information about the servers in the eX5 portfolio. This information is most useful in designing, configuring, and planning to order a server solution.

This part consists of the following chapters:

2

IBM eX5 technology

This chapter describes the technology that IBM brings to the IBM eX5 portfolio of servers. The chapter describes the fifth generation of IBM Enterprise X-Architecture (EXA) chip sets, called *eX5*. This chip set is the enabling technology for IBM to expand the memory subsystem independently of the remainder of the x86 system. Next, we describe the latest Intel Xeon 6500 and 7500 family of processors and give the features that are currently available. We then describe the current memory features, MAX5 memory expansion line, IBM exclusive system scaling and partitioning capabilities, and eXFlash. eXFlash can dramatically increase system disk I/O by using internal solid-state storage instead of traditional disk-based storage. We also describe integrated virtualization and implementation guidelines for installing a new server.

This chapter contains the following topics:

2.1 eX5 chip set

The members of the eX5 server family are defined by their ability to use IBM fifth-generation chip sets for Intel x86 server processors. IBM engineering, under the banner of Enterprise X-Architecture (EXA), brings advanced system features to the Intel server marketplace. Previous generations of EXA chip sets powered System x servers from IBM with scalability and performance beyond what was available with the chip sets from Intel.

The Intel QuickPath Interconnect (QPI) specification includes definitions for the following items:

- Processor-to-processor communications
- Processor-to-I/O hub communications
- Connections from processors to chip sets, such as eX5, referred to as *node controllers*

To fully utilize the increased computational ability of the new generation of Intel processors, eX5 provides additional memory capacity and additional scalable memory interconnects (SMIs), increasing bandwidth to memory. eX5 also provides these additional reliability, availability, and serviceability (RAS) capabilities for memory: Chipkill, Memory ProteXion, and Full Array Memory Mirroring.

QPI uses a source snoop protocol. This technique means that a CPU, even if it knows another processor has a cache line it wants (the cache line address is in the snoop filter, and it is in the shared state), must request a copy of the cache line and wait for the result to be returned from the source. The eX5 snoop filter contains the contents of the cache lines and can return them immediately. For more information about the source snoop protocol, see 2.2.4, “QuickPath Interconnect (QPI)” on page 18.

Memory that is directly controlled by a processor can be accessed faster than through the eX5 chip set, but because the eX5 chip set is connected to all processors, it provides less delay than accesses to memory controlled by another processor in the system.

2.2 Intel Xeon 6500 and 7500 family processors

The IBM eX5 servers use the Intel Xeon 6500 and Xeon 7500 family of processors to maximize performance. These processors are the latest in a long line of high-performance processors:

- The Xeon 6500 family is used in the x3690 X5 and BladeCenter HX5. These processors are only scalable to up to two processors. This processor does not support the ability to scale to multiple nodes; however, certain models support MAX5.
- The Xeon 7500 is the latest Intel scalable processor and can be used to scale to two or more processors. When used in the IBM x3850 and x3950 X5, these servers can scale up to eight processors. With the HX5 blade server, scaling up to two nodes with four processors is supported.

Table 2-1 on page 17 compares the Intel Xeon 6500 and 7500 with the Intel Xeon 5500 and 5600 processors that are available in other IBM servers.

Table 2-1 Two-socket, 2-socket scalable, and 4-socket scalable Intel processors minimum configuration

	Xeon 5500	Xeon 5600	Xeon 6500	Xeon 7500
Used in	x3400 M2 x3500 M2 x3550 M2 x3650 M2 HS22 HS22V	x3400 M3 x3500 M3 x3550 M3 x3650 M3 HS22 HS22V	x3690 X5 HX5	x3850 X5 x3950 X5 HX5
Intel development name	Nehalem-EP	Westmere-EP	Nehalem-EX	Nehalem-EX
Maximum processors per server	2	2	2	HX5: 2 x3850 X5: 4
CPU cores per processor	2 or 4	4 or 6	4, 6, or 8	4, 6, or 8
Last level cache (MB)	4 or 8 MB	8 or 12 MB	12 or 18 MB	18 or 24 MB
Memory DIMMs per processor (maximum)[a]	9	9	16[a]	16

a. Requires that the memory mezzanine board is installed along with processor two on x3690 X5

For more information about processor options and the installation order of the processors, see the following links:

- IBM System x3850 X5: 3.7, "Processor options" on page 74
- IBM System x3690 X5: 4.7, "Processor options" on page 130
- IBM BladeCenter HX5: 5.9, "Processor options" on page 192

2.2.1 Intel Virtualization Technology

Intel Virtualization Technology (Intel VT) is a suite of processor hardware enhancements that assists virtualization software to deliver more efficient virtualization solutions and greater capabilities, including 64-bit guest OS support.

Intel VT Flex Priority optimizes virtualization software efficiency by improving interrupt handling.

Intel VT Flex migration enables the Xeon 7500 series to be added to the existing virtualization pool with single, 2, 4, or 8-socket servers.

For more information about Intel Virtual Technology, go to the following website:

http://www.intel.com/technology/virtualization/

2.2.2 Hyper-Threading Technology

Intel Hyper-Threading Technology enables a single physical processor to execute two separate code streams (threads) concurrently. To the operating system, a processor core with Hyper-Threading is seen as two logical processors, each of which has its own architectural state, that is, its own data, segment, and control registers and its own advanced programmable interrupt controller (APIC).

Each logical processor can be individually halted, interrupted, or directed to execute a specified thread, independently from the other logical processor on the chip. The logical processors share the execution resources of the processor core, which include the execution engine, the caches, the system interface, and the firmware.

Hyper-Threading Technology is designed to improve server performance by exploiting the multi-threading capability of operating systems and server applications in such a way as to increase the use of the on-chip execution resources available on these processors. Applications types that make the best use of Hyper-Threading are virtualization, databases, email, Java™, and web servers.

For more information about Hyper-Threading Technology, go to the following website:

http://www.intel.com/technology/platform-technology/hyper-threading/

2.2.3 Turbo Boost Technology

Intel Turbo Boost Technology dynamically turns off unused processor cores and increases the clock speed of the cores in use. For example, with six cores active, a 2.26 GHz 8-core processor can run the cores at 2.53 GHz. With only three or four cores active, the same processor can run those cores at 2.67 GHz. When the cores are needed again, they are dynamically turned back on and the processor frequency is adjusted accordingly.

Turbo Boost Technology is available on a per-processor number basis for the eX5 systems. For ACPI-aware operating systems, no changes are required to take advantage of it. Turbo Boost Technology can be engaged with any number of cores enabled and active, resulting in increased performance of both multi-threaded and single-threaded workloads.

Frequency steps are in 133 MHz increments, and they depend on the number of active cores. For the 8-core processors, the number of frequency increments is expressed as four numbers separated by slashes: the first two for when seven or eight cores are active, the next for when five or six cores are active, the next for when three or four cores are active, and the last for when one or two cores are active, for example, 1/2/4/5 or 0/1/3/5.

When temperature, power, or current exceeds factory-configured limits and the processor is running above the base operating frequency, the processor automatically steps the core frequency back down to reduce temperature, power, and current. The processor then monitors temperature, power, and current and re-evaluates. At any given time, all active cores run at the same frequency.

For more information about Turbo Boost Technology, go to the following website:

http://www.intel.com/technology/turboboost/

2.2.4 QuickPath Interconnect (QPI)

Early Intel Xeon multiprocessor systems used a shared front-side bus, over which all processors connect to a core chip set, and which provides access to the memory and I/O subsystems, as shown in Figure 2-1 on page 19. Servers that implemented this design include the IBM eServer™ xSeries 440 and the xSeries 445.

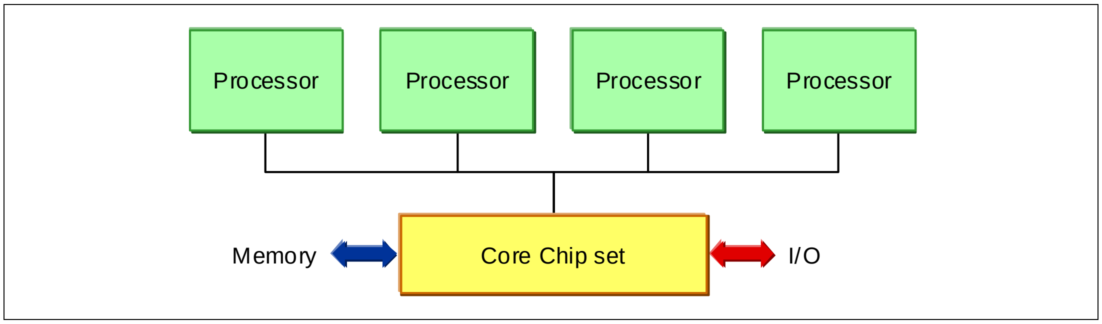


Figure 2-1 Shared front-side bus, in the IBM x360 and x440; with snoop filter in the x365 and x445

The front-side bus carries all reads and writes to the I/O devices, and all reads and writes to memory. Also, before a processor can use the contents of its own cache, it must know whether another processor has the same data stored in its cache. This process is described as *snooping* the other processor's caches, and it puts a lot of traffic on the front-side bus.

To reduce the amount of cache snooping on the front-side bus, the core chip set can include a *snoop filter*, which is also referred to as a *cache coherency filter*. This filter is a table that keeps track of the starting memory locations of the 64-byte chunks of data that are read into cache, called *cache lines*, or the actual cache line itself, and one of four states: modified, exclusive, shared, or invalid (MESI).

The next step in the evolution was to divide the load between a pair of front-side buses, as shown in Figure 2-2. Servers that implemented this design include the IBM System x3850 and x3950 (the *M1* version).

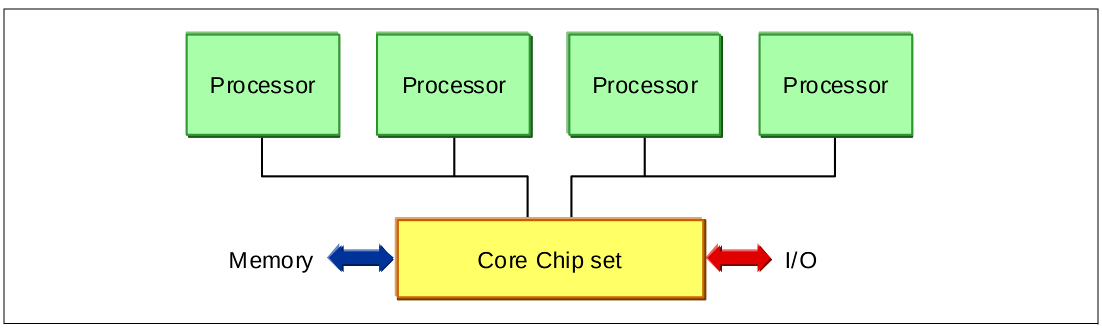


Figure 2-2 Dual independent buses, as in the x366 and x460 (later called the x3850 and x3950)

This approach had the effect of reducing congestion on each front-side bus, when used with a snoop filter. It was followed by independent processor buses, shown in Figure 2-3 on page 20. Servers implementing this design included the IBM System x3850 M2 and x3950 M2.

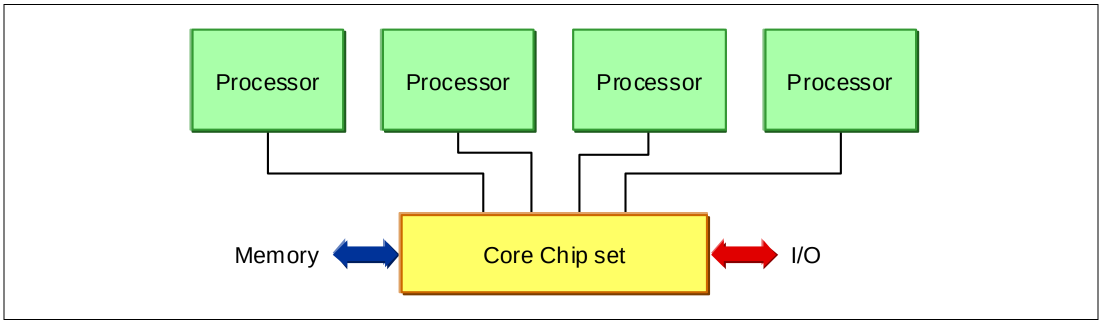


Figure 2-3 Independent processor buses, as in the x3850 M2 and x3950 M2

Instead of a parallel bus connecting the processors to a core chip set, which functions as both a memory and I/O controller, the Xeon 6500 and 7500 family processors implemented in IBM eX5 servers include a separate memory controller to each processor. Processor-to-processor communications are carried over shared-clock, or *coherent* QPI links, and I/O is transported over *non-coherent* QPI links through I/O hubs. Figure 2-4 shows this information.

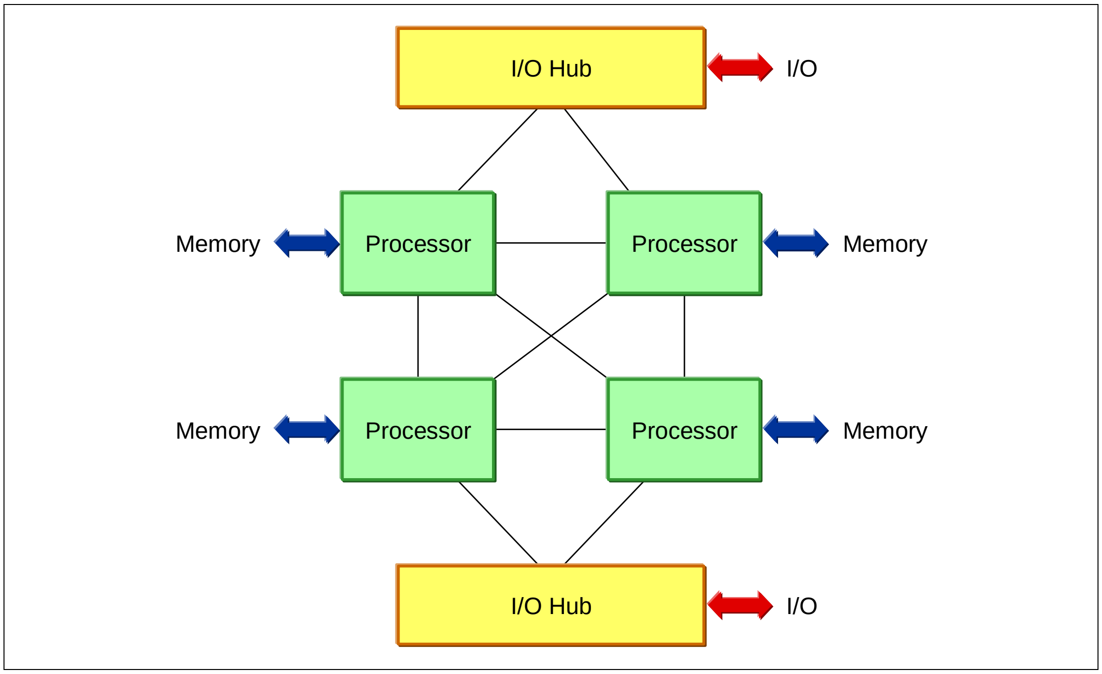


Figure 2-4 Figure 2-4 QPI, as used in the eX5 portfolio

In previous designs, the entire range of memory was accessible through the core chip set by each processor, a shared memory architecture. This design creates a non-uniform memory access (NUMA) system, in which part of the memory is directly connected to the processor where a given thread is running, and the rest must be accessed over a QPI link through another processor. Similarly, I/O can be local to a processor, or remote through another processor.

For QPI use, Intel has modified the MESI cache coherence protocol to include a forwarding state, so when a processor asks to copy a shared cache line, only one other processor responds.

For more information about QPI, go to the following website:

http://www.intel.com/technology/quickpath/

2.2.5 Processor performance in a green world

All eX5 servers from the factory are designed to use power in the most efficient means possible. Controlling how much power the server is going to use is managed by controlling the core frequency and power applied to the processors, controlling the frequency and power applied to the memory, and by reducing fan speeds to fit the cooling needs of the server. For most server configurations, these functions are ideal to provide the best performance possible without wasting energy during off-peak usage.

Servers that are used in virtualized clusters of host computers often have the same attempts being made to manage power consumption at the operating system level. In this environment, the operating system makes decisions about moving and balancing virtual servers across an array of host servers. The operating system, running on multiple hosts, reports to a single cluster controller about the resources that remain on the host and the resource demands of any virtual servers running on that host. The cluster controller makes decisions to move virtual servers from one host to another host to completely power down hosts that are no longer needed during off-peak hours.

It is a common occurrence to have virtual servers moving back and forth across the same set of host servers, because the host servers are themselves changing their own processor performance to save power. The result is an inefficient system that is both slow to respond and actually consumes more power.

The solution for virtual server clusters is to turn off the power management features of the host servers. The process to change the hardware-controlled power management in the F1-Setup, offered during power-on self test (POST), is to select **System Settings** → **Operating Modes** → **Choose Operating Mode**. Figure 2-5 shows the available options and the selection to choose to configure the server for Performance Mode.

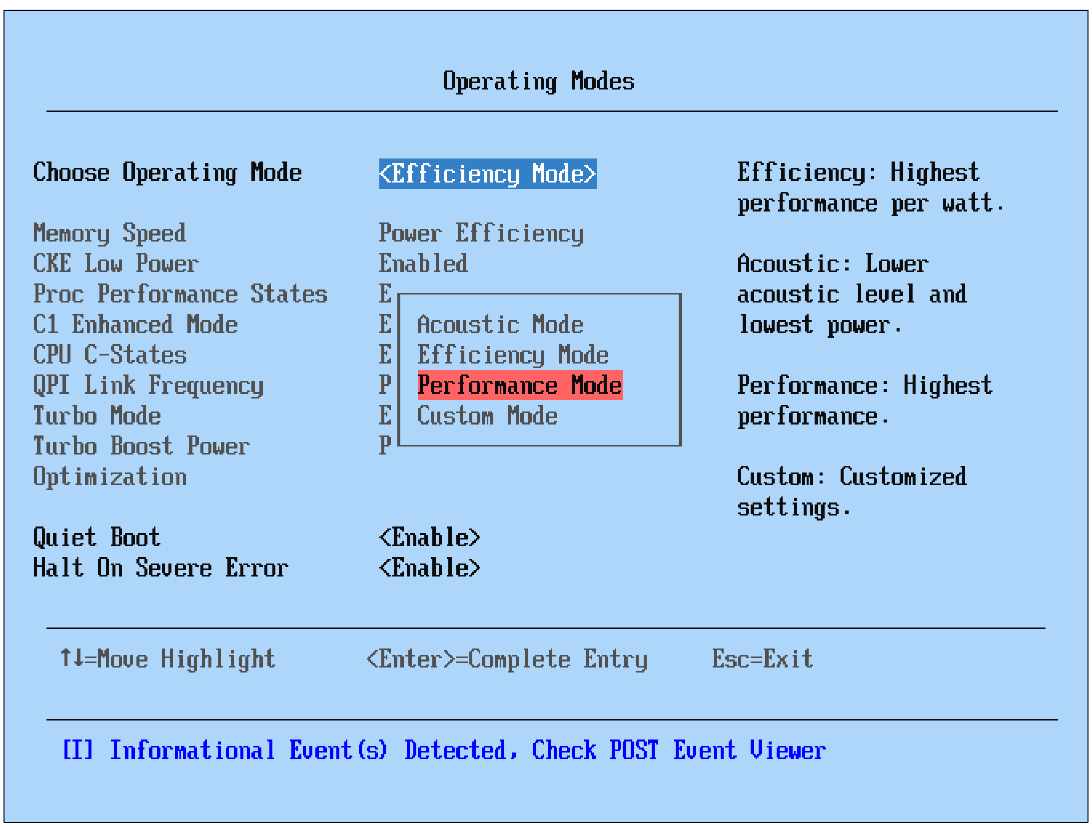


Figure 2-5 Setup (F1) → System Settings → Operating Modes to set Performance Mode

2.3 Memory

In this section, we describe the major features of the memory subsystem in eX5 systems. We describe the following topics in this section:

- 2.3.1, “Memory speed” on page 22
- 2.3.2, “Memory DIMM placement” on page 23
- 2.3.3, “Memory ranking” on page 24
- 2.3.4, “Nonuniform memory architecture (NUMA)” on page 26
- 2.3.5, “Hemisphere Mode” on page 26
- 2.3.6, “Reliability, availability, and serviceability (RAS) features” on page 28
- 2.3.7, “I/O hubs” on page 30

2.3.1 Memory speed

As with Intel Xeon 5500 processor (Nehalem-EP), the speed at which the memory that is connected to the Xeon 7500 and 6500 processors (Nehalem-EX) runs depends on the capabilities of the specific processor. With Nehalem-EX, the scalable memory interconnect (SMI) link runs from the memory controller integrated in the processor to the memory buffers on the memory cards.

The SMI link speed is derived from the QPI link speed:

- 6.4 gigatransfers per second (GT/s) QPI link speed capable of running memory speeds up to 1066 MHz
- 5.86 GT/s QPI link speed capable of running memory speeds up to 978 MHz
- 4.8 GT/s QPI link speed capable of running memory speeds up to 800 MHz

Gigatransfers: Gigatransfers per second (GT/s) or 1,000,000,000 transfers per second is a way to measure bandwidth. The actual data that is transferred depends on the width of the connection (that is, the transaction size). To translate a given value of GT/s to a theoretical maximum throughput, multiply the transaction size by the GT/s value. In most circumstances, the transaction size is the width of the bus in bits. For example, the SMI links are 13 bits to the processor and 10 bits from the processor.

Because the memory controller is on the CPU, the memory slots for a CPU can only be used if a CPU is in that slot. If a CPU fails when the system reboots, it is brought back online without the failed CPU and without the memory associated with that CPU slot.

QPI bus speeds are listed in the processor offerings of each system, which equates to the SMI bus speed. The QPI speed is listed as `x4.8` or similar, as shown in the following example:

```
2x 4 Core 1.86GHz,18MB x4.8 95W (4x4GB), 2 Mem Cards
2x 8 Core 2.27GHz,24MB x6.4 130W (4x4GB), 2 Mem Cards
```

The value `x4.8` corresponds to an SMI link speed of 4.8 GT/s, which corresponds to a memory bus speed of 800 MHz. The value `x6.4` corresponds to an SMI link speed of 6.4 GT/s, which corresponds to a memory bus speed of 1066 MHz.

The processor controls the maximum speed of the memory bus. Even if the memory dual inline memory modules (DIMMs) are rated at 1066 MHz, if the processor supports only 800 MHz, the memory bus speed is 800 MHz.

What about 1333 MHz? The maximum memory speed that is supported by Xeon 7500 and 6500 processors is 1066 MHz (1333 MHz is not supported). Although the 1333 MHz DIMMs are still supported, they can operate at a maximum speed of 1066 MHz.

Memory performance test on various memory speeds

Based on benchmarks using an IBM internal load generator run on an x3850 X5 system configured with four X7560 processors and 64x 4 GB quad-rank DIMMs, the following results were observed:

- Peak throughput per processor observed at 1066 MHz: 27.1 gigabytes per second (GBps)
- Peak throughput per processor observed at 978 MHz: 25.6 GBps
- Peak throughput per processor observed at 800 MHz: 23.0 GBps

Stated another way, an 11% throughput increase exists when frequency is increased from 800 MHz to 978 MHz; a 6% throughput increase exists when frequency is increased from 978 MHz to 1066 MHz.

Key points regarding these benchmark results:

- Use these results only as a guide to the relative performance between the various memory speeds, not the absolute speeds.
- The benchmarking tool that is used accesses only local memory, and there were no remote memory accesses.
- Given the nature of the benchmarking tool, these results might not be achievable in a production environment.

2.3.2 Memory DIMM placement

The eX5 servers support a variety of ways to install memory DIMMs, which we describe in detail in later chapters. However, it is important to understand that because of the layout of the SMI links, memory buffers, and memory channels, you must install the DIMMs in the correct locations to maximize performance.

Figure 2-6 on page 24 shows eight possible memory configurations for the two memory cards and 16 DIMMs connected to each processor socket in an x3850 X5. Similar configurations apply to the x3690 X5 and HX5. Each configuration has a relative performance score. The following key information from this chart is important:

- The best performance is achieved by populating all memory DIMMs in the server (configuration 1 in Figure 2-6 on page 24).
- Populating only one memory card per socket can result in approximately a 50% performance degradation (compare configuration 1 with 5).
- Memory performance is better if you install DIMMs on all memory channels than if you leave any memory channels empty (compare configuration 2 with 3).
- Two DIMMs per channel result in better performance that one DIMM per channel (compare configuration 1 with 2, and compare 5 with 6).

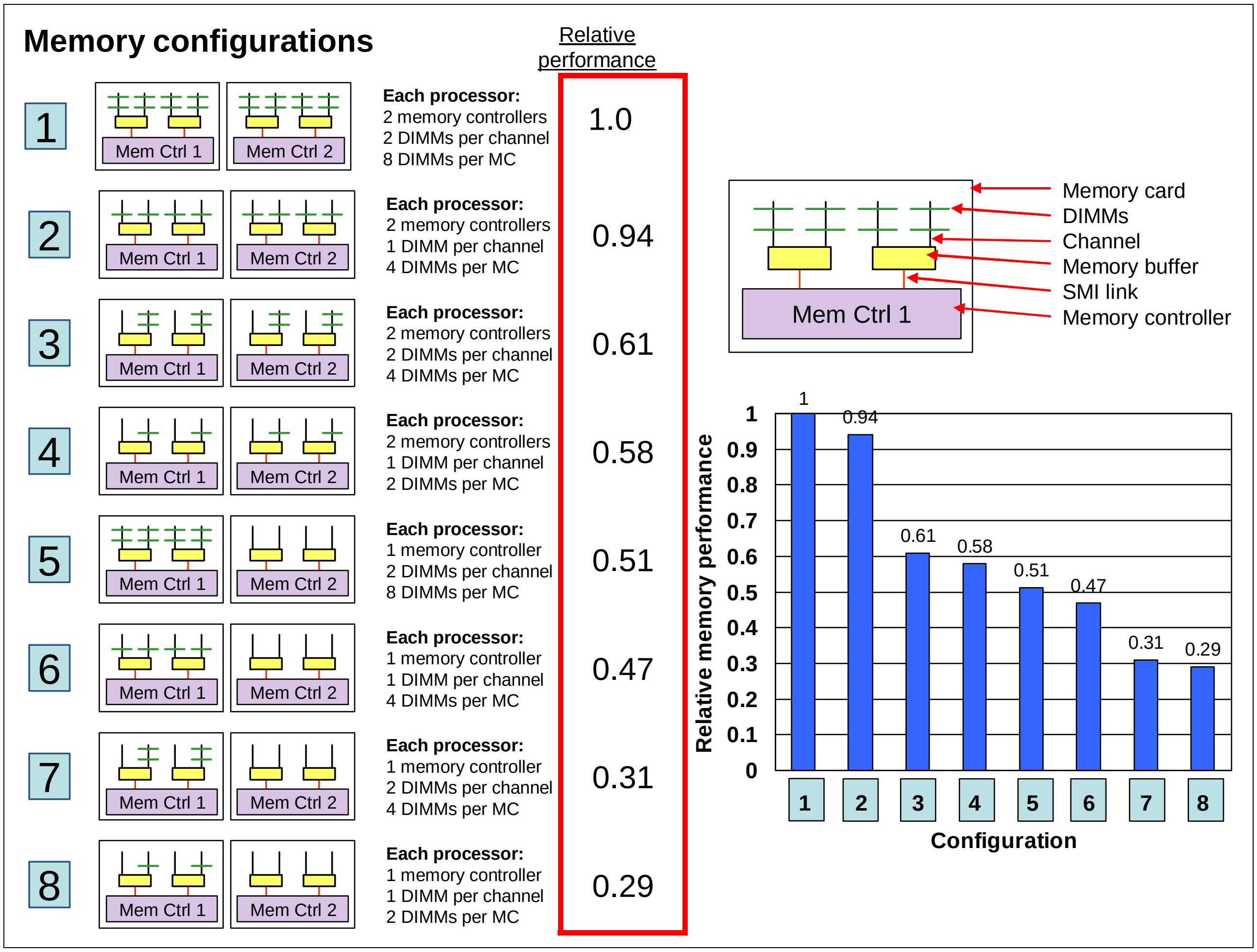


Figure 2-6 Relative memory performance based on DIMM placement (one processor and two memory cards shown)

2.3.3 Memory ranking

The underlying speed of the memory as measured in MHz is *not* sensitive to memory population. (In Intel Xeon 5500 processor-based systems, such as the x3650 M2, if rules regarding optimal memory population are not followed, the system BIOS clocks the memory subsystem down to a slower speed. This situation is not the case with the x3850 X5.)

Unlike Intel 5500 processor-based systems, more ranks are *better* for performance in the x3850 X5. Therefore, quad-rank memory is better than dual-rank memory, and dual-rank memory is better than single-rank memory. Again, the frequency of the memory as measured in MHz does not change depending on the number of ranks used. (Intel 5500-based systems, such as the x3650 M2, are sensitive to the number of ranks installed. Quad-rank memory in those systems always triggers a stepping down of memory speed as enforced by the BIOS, which is not the case with the eX5 series.)

Performance test between ranks

With the Xeon 7500 and 6500 processors, having more ranks gives better performance. The better performance is the result of the addressing scheme. The addressing scheme can extend the pages across ranks, thereby making the pages effectively larger and therefore creating more page-hit cycles.

We used three types of memory DIMMs for this analysis:

- Four GB 4Rx8 (four ranks using x8 DRAM technology)
- Two GB 2Rx8 (two ranks)
- One GB 1Rx8 (one rank)

We used the following memory configurations:

- Fully populated memory:
  - Two DIMMs on each memory channel
  - Eight DIMMs per memory card
- Half-populated memory:
  - One DIMM on each memory channel
  - Four DIMMs per memory card (slots 1, 3, 6, and 8; see Figure 3-16 on page 76)
- Quarter-populated memory:
  - One DIMM on just half of the memory channels
  - Two DIMMs per memory card

Although several benchmarks were conducted, this section focuses on the results gathered using the industry-standard STREAM benchmark, as shown in Figure 2-7.

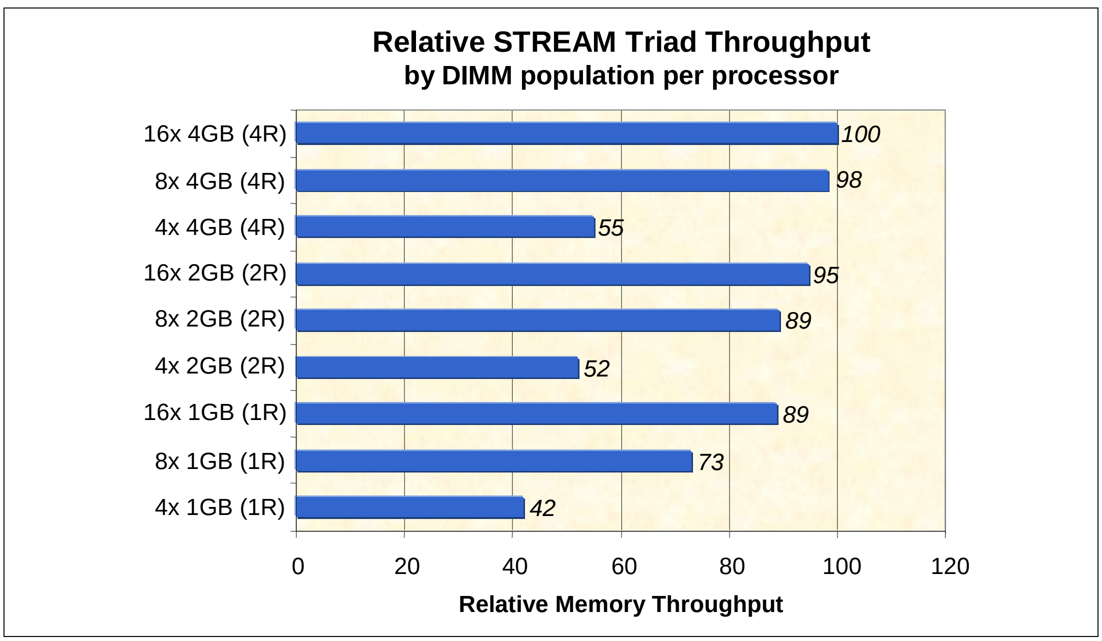


Figure 2-7 Comparing the performance of memory DIMM configurations using STREAM

Taking the top performance result of 16x 4 GB quad-rank DIMMs as the baseline, we see how the performance drops to 95% of the top performance with 16x 2 GB dual-rank DIMMs, and 89% of the top performance with 16x 1 GB single-rank DIMMs.

You can see similar effects across the three configurations based on eight DIMMs per processor and four DIMMs per processor. These results also emphasize the same effect that is shown in 3.8.3, “Maximizing memory performance” on page 84 for the x3850 X5, where performance drops away dramatically when all eight memory channels per CPU are not used.

Tip: Additional ranks increase the memory bus loading, which is why on Xeon 5500 (Nehalem EP) platforms, the opposite effect can occur: memory slows down if too many rank loads are attached. The use of scalable memory buffers in the x3850 X5 with Xeon 7500/6500 processors avoids this slowdown.

2.3.4 Nonuniform memory architecture (NUMA)

Nonuniform memory architecture (NUMA) is an important consideration when configuring memory, because a processor can access its own local memory faster than non-local memory. Not all configurations use 64 DIMMs spread across 32 channels. Certain configurations might have a more modest capacity and performance requirement. For these configurations, another principle to consider when configuring memory is that of *balance*. A balanced configuration has all of the memory cards configured with the same *amount* of memory, even if the quantity and size of the DIMMs differ from card to card. This principle helps to keep remote memory access to a minimum. DIMMs must always be installed in matched pairs.

A server with a NUMA, such as the servers in the eX5 family, has local and remote memory. For a given thread running in a processor core, *local memory* refers to the DIMMs that are directly connected to that particular processor. *Remote memory* refers to the DIMMs that are not connected to the processor where the thread is running currently.

Remote memory is attached to another processor in the system and must be accessed through a QPI link. However, using remote memory adds latency. The more such latencies add up in a server, the more performance can degrade. Starting with a memory configuration where each CPU has the same local RAM capacity is a logical step toward keeping remote memory accesses to a minimum.

For more information about NUMA installation options, see the following sections:

- IBM System x3850 X5: 3.8.2, “DIMM population sequence” on page 79
- IBM System x3690 X5: “Two processors with memory mezzanine installed” on page 135
- IBM BladeCenter HX5: 5.10.2, “DIMM population order” on page 196

2.3.5 Hemisphere Mode

Hemisphere Mode is an important performance optimization of the Xeon 6500 and 7500 processors. Hemisphere Mode is automatically enabled by the system if the memory configuration allows it. This mode interleaves memory requests between the two memory controllers within each processor, enabling reduced latency and increased throughput. It also allows the processor to optimize its internal buffers to maximize memory throughput.

Two-node configurations: A memory configuration that enables Hemisphere Mode is *required* for 2-node configurations on x3850 X5.

Hemisphere Mode is a global parameter that is set at the system level. This setting means that if even one processor’s memory is incorrectly configured, the entire system can lose the performance benefits of this optimization. Stated another way, either all processors in the system use Hemisphere Mode, or all do not.

Hemisphere Mode is enabled only when the memory configuration behind each memory controller on a processor is identical. Because the Xeon 7500 memory population rules dictate that a minimum of two DIMMs are installed on each memory controller at a time (one on each of the attached memory buffers), DIMMs must be installed in quantities of four per processor to enable Hemisphere Mode.

In addition, because eight DIMMs per processor are required for using all memory channels, eight DIMMs per processor need to be installed at a time for optimized memory performance. Failure to populate all eight channels on a processor can result in a performance reduction of approximately 50%.

Hemisphere Mode does not require that the memory configuration of each CPU is identical. For example, Hemisphere Mode is still enabled if CPU 0 is configured with 8x 4 GB DIMMs, and processor 1 is configured with 8x 2 GB DIMMs. Depending on the application characteristics, however, an unbalanced memory configuration can cause reduced performance because it forces a larger number of remote memory requests over the inter-CPU QPI links to the processors with more memory.

We summarize these points:

- There are two memory buffers per memory channel, two channels per memory controller, and two controllers per processor. Each memory channel must contain *at least* one DIMM to enable Hemisphere Mode.
- Within a processor, both memory controllers need to contain identical DIMM configurations to enable Hemisphere Mode. Therefore, for best results, install at least eight DIMMs per processor.

Industry-standard tests run on one Xeon 7500 processor with various memory configurations have shown that there are performance implications if Hemisphere Mode is not enabled. For example, for a configuration with eight DIMMs installed and spread across both memory controllers in a processor and all memory buffers (see Figure 2-8), there is a drop in performance of 16% if Hemisphere Mode is not enabled.

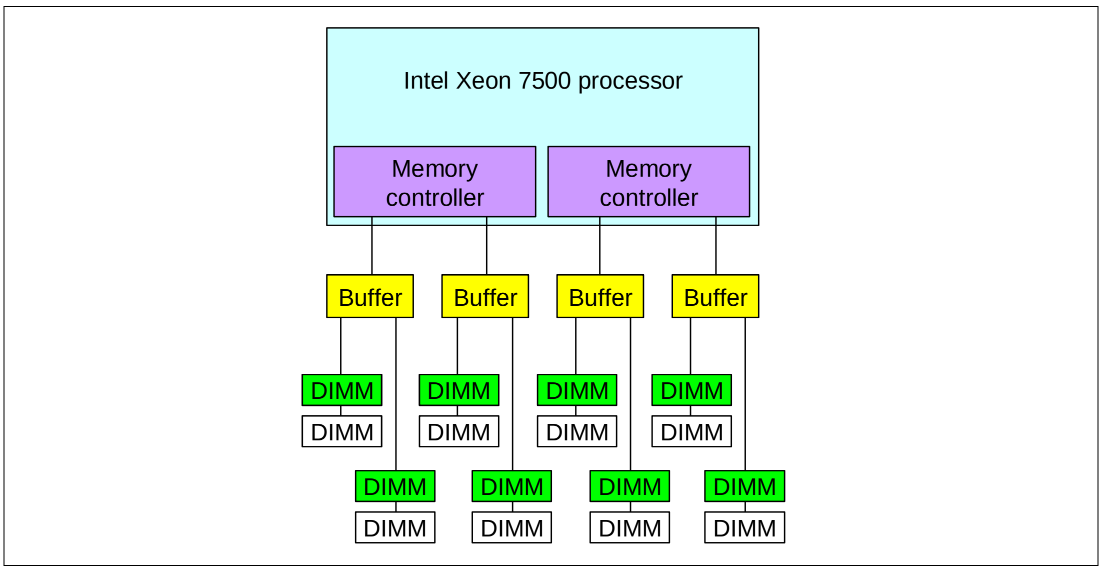


Figure 2-8 Example memory configuration

For more information about Hemisphere Mode installation options, see the following sections:

- IBM System x3850 X5: 3.8.2, “DIMM population sequence” on page 79
- IBM System x3690 X5: “Two processors with memory mezzanine installed” on page 135
- IBM BladeCenter HX5: 5.10.2, “DIMM population order” on page 196

2.3.6 Reliability, availability, and serviceability (RAS) features

In addition to Hemisphere Mode, DIMM balance and memory size, memory performance is also affected by the various memory reliability, availability, and serviceability (RAS) features that can be enabled from the Unified Extensible Firmware Interface (UEFI) shell. These settings can increase the reliability of the system; however, there are performance trade-offs when these features are enabled.

The available memory RAS settings are normal, mirroring, and sparing. On the X5 platforms, you can access these settings under the Memory option menu in System Settings.

This section is not meant to provide a comprehensive overview of the memory RAS features that are available in the Xeon 7500 processor, but rather it provides a brief introduction to each mode and its corresponding performance effects. For more information about memory RAS features and platform-specific requirements, see the following sections:

- System x3850 X5: 6.9, “UEFI settings” on page 259
- System x3690 X5: 7.8, “UEFI settings” on page 337
- BladeCenter HX5: 8.5, “UEFI settings” on page 396

The following sections provide a brief description of each memory RAS setting.

Memory mirroring

To further improve memory reliability and availability beyond error checking and correcting (ECC) and Chipkill, the chip set can mirror memory data to two memory ports. To successfully enable mirroring, you must have both memory cards per processor installed, and populate the same amount of memory in both memory cards. Partial mirroring (mirroring of part but not all of the installed memory) is not supported.

Memory mirroring, or *full array memory mirroring* (FAMM) *redundancy,* provides the user with a redundant copy of all code and data addressable in the configured memory map. Memory mirroring works within the chip set by writing data to two memory ports on every memory-write cycle. Two copies of the data are kept, similar to the way RAID-1 writes to disk. Reads are interleaved between memory ports. The system automatically uses the most reliable memory port as determined by error logging and monitoring.

If errors occur, only the alternate memory port is used, until bad memory is replaced. Because a redundant copy is kept, mirroring results in having only half the installed memory available to the operating system. FAMM does not support asymmetrical memory configurations and requires that each port is populated in identical fashion. For example, you must install 2 GB of identical memory equally and symmetrically across the two memory ports to achieve 1 GB of mirrored memory.

FAMM enables other enhanced memory features, such as Unrecoverable Error (UE) recovery, and is required for support of memory hot replace.

Memory mirroring is independent of the operating system.

For more information about system-specific memory mirroring installation options, see the following sections:

- x3850 X5: 3.8.4, “Memory mirroring” on page 87
- x3690 X5: 4.8.6, “Memory mirroring” on page 141
- BladeCenter HX5: 5.10.4, “Memory mirroring” on page 200

Memory sparing

Sparing provides a degree of redundancy in the memory subsystem, but not to the extent of mirroring. In contrast to mirroring, sparing leaves more memory for the operating system. In sparing mode, the trigger for failover is a preset threshold of correctable errors. Depending on the type of sparing (DIMM or rank), when this threshold is reached, the content is copied to its spare. The failed DIMM or rank is then taken offline, with the spare counterpart activated for use. There are two sparing options:

- DIMM sparing

  Two unused DIMMs are spared per memory card. These DIMMs must have the same rank and capacity as the largest DIMMs that we are sparing. The size of the two unused DIMMs for sparing is subtracted from the usable capacity that is presented to the operating system. DIMM sparing is applied on all memory cards in the system.

- Rank sparing

  Two ranks per memory card are configured as spares. The ranks have to be as large as the rank relative to the highest capacity DIMM that we are sparing. The size of the two unused ranks for sparing is subtracted from the usable capacity that is presented to the operating system. Rank sparing is applied on all memory cards in the system.

You configure these options by using the UEFI during start-up.

For more information about system-specific memory sparing installation options, see the following sections:

- IBM System x3850 X5: 3.8.5, “Memory sparing” on page 89
- IBM System x3690 X5: 4.8.7, “Memory sparing” on page 143
- IBM BladeCenter HX5: 5.10.5, “Memory sparing” on page 202

Chipkill

Chipkill memory technology, an advanced form of error checking and correcting (ECC) from IBM, is available for the eX5 blade. Chipkill protects the memory in the system from any single memory chip failure. It also protects against multi-bit errors from any portion of a single memory chip.

Redundant bit steering

Redundant bit steering (RBS) provides the equivalent of a hot-spare drive in a RAID array. It is based in the memory controller, and it senses when a chip on a DIMM has failed and when to route the data around the failed chip.

The eX5 servers do not currently support redundant bit steering, because the integrated memory controller of the Intel Xeon 6500 and 7500 processors do not support the feature. However, the MAX5 memory expansion unit supports RBS but only when x4 memory DIMMs are used. The x8 DIMMs do not support RBS.

RBS is automatically enabled in the MAX5 memory port, if all DIMMs installed to that memory port are x4 DIMMs.

RBS uses the ECC coding scheme that provides Chipkill coverage for x4 DRAMs. This coding scheme leaves the equivalent of one x4 DRAM spare in every pair of DIMMs. In the event that a chip failure on the DIMM is detected by memory scrubbing, the memory controller can reroute data around that failed chip through these spare bits. DIMMs using x8 DRAM technology use a separate ECC coding scheme that does not leave spare bits, which is why RBS is not available on x8 DIMMs.

RBS operates automatically without issuing a Predictive Failure Analysis (PFA) or light path diagnostics alert to the administrator, although an event is logged to the service processor log. After the second DIMM failure, PFA and light path diagnostics alerts occur on that DIMM normally.

Lock step

IBM eX5 memory operates in lock step mode. Lock step is a memory protection feature that involves the pairing of two memory DIMMs. The paired DIMMs can perform the same operations and the results are compared. If any discrepancies exist between the results, a memory error is signaled. Lock step mode gives a maximum of 64 GB of usable memory with one CPU installed, and 128 GB of usable memory with two CPUs installed (using 8 GB DIMMs).

Memory must be installed in pairs of two identical DIMMs per processor. Although the size of the DIMM pairs installed can differ, the pairs must be of the same speed.

Machine Check Architecture (MCA)

MCA is a RAS feature that has previously only been available for other processor architectures, such as Intel Itanium®, IBM POWER®, and other reduced instruction set computing (RISC) processors, and mainframes. Implementation of the MCA requires hardware support, firmware support, such as the UEFI, and operating system support.

The MCA enables system-error handling that otherwise requires stopping the operating system. For example, if a memory location in a DIMM no longer functions properly and it cannot be recovered by the DIMM or memory controller logic, MCA logs the failure and prevents that memory location from being used. If the memory location was in use by a thread at the time, the process that owns the thread is terminated.

Microsoft, Novell, Red Hat, VMware, and other operating system vendors have announced support for the Intel MCA on the Xeon processors.

Scalable memory buffers

Unlike the Xeon 5500 and 5600 series, which use unbuffered memory channels, the Xeon 6500 and 7500 processors use scalable memory buffers in the systems design. This approach reflects the various workloads for which these processors were intended. The 6500 and 7500 family processors are designed for workloads requiring more memory, such as virtualization and databases. The use of scalable memory buffers allows more memory per processor, and prevents memory bandwidth reductions when more memory is added per processor.

2.3.7 I/O hubs

The connection to I/O devices (such as keyboard, mouse, and USB) and to I/O adapters (such as hard disk drive controllers, Ethernet network interfaces, and Fibre Channel host bus adapters) is handled by I/O hubs, which then connect to the processors through QPI links.

Figure 2-4 on page 20 shows the I/O hub connectivity. Connections to the I/O devices are fault tolerant, because data can be routed over either of the two QPI links to each I/O hub. For optimal system performance in the four processor systems (with two I/O hubs), balance the high-throughput adapters across the I/O hubs.

For more information regarding each of the eX5 systems and the available I/O adapters, see the following sections:

- IBM System x3850 X5: 3.12, “I/O cards” on page 104.
- IBM System x3690 X5: 4.10.4, “I/O adapters” on page 168.
- IBM BladeCenter HX5: 5.13, “I/O expansion cards” on page 209.

2.4 MAX5

Memory Access for eX5 (*MAX5*) is the name give to the memory and scalability subsystem that can be added to eX5 servers. In the Intel QPI specification, the MAX5 is a node controller. MAX5 for the rack-mounted systems (x3850 X5, x3950 X5, and x3690 X5) is in the form of a 1U device that attaches beneath the server. For the BladeCenter HX5, MAX5 is implemented in the form of an expansion blade that adds 30 mm to the width of the blade (the width of one extra blade bay). Figure 2-9 shows the HX5 with the MAX5 attached.

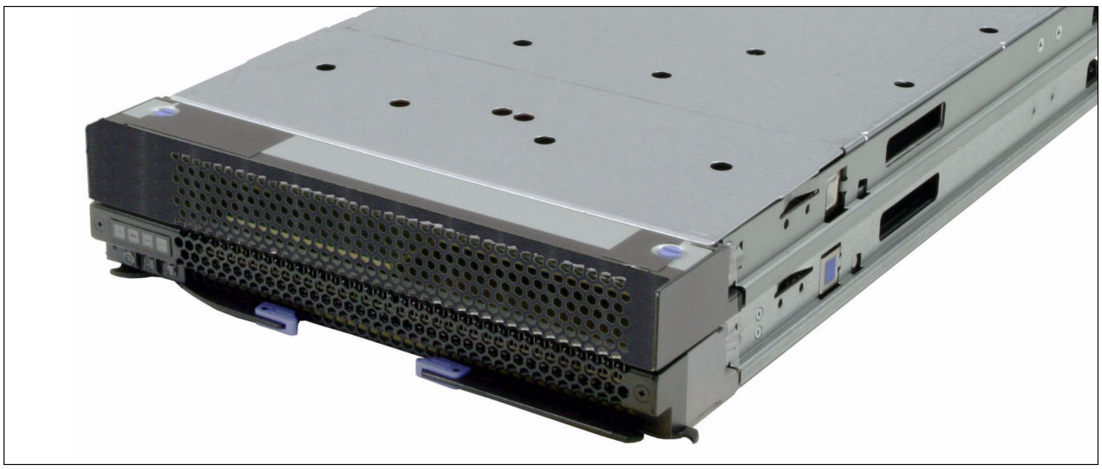


Figure 2-9 Single-node HX5 and MAX5

Figure 2-10 shows the x3850 X5 with the MAX5 attached.

Figure 2-10 IBM System x3850 X5 with MAX5 (the MAX5 is the 1U unit beneath the main system)

Figure 2-11 shows the MAX5 removed from the housing.

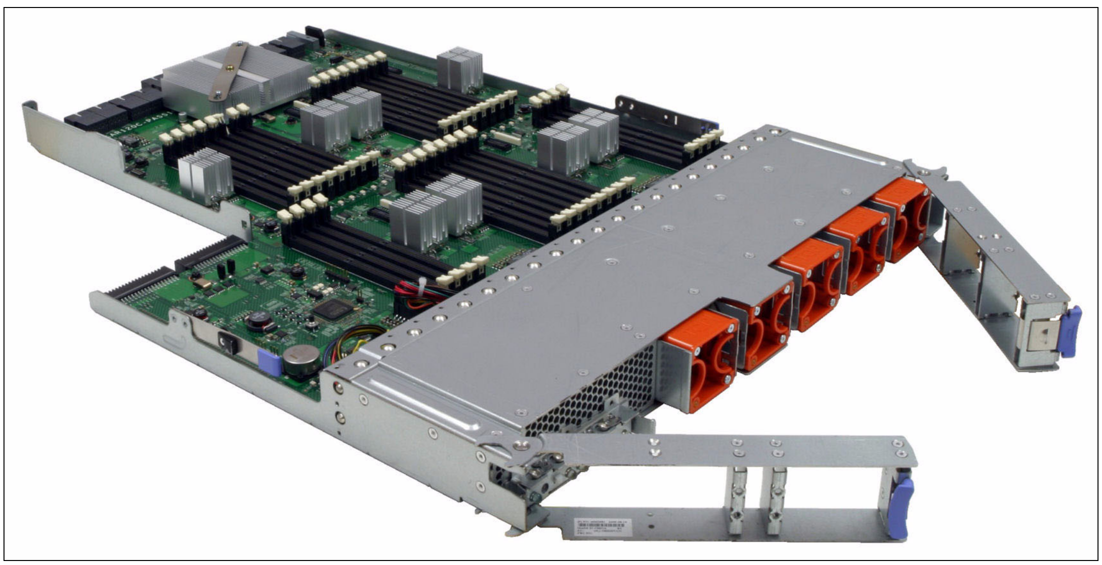


Figure 2-11 IBM MAX5 for the x3850 X5 and x3690 X5

MAX5 connects to these systems through QPI links and provides the EXA scalability interfaces. The eX5 chip set, which is described in 2.1, “eX5 chip set” on page 16, is contained in the MAX5 units.

Table 2-2 through Table 2-4 show the memory capacity and bandwidth increases that are possible with MAX5 for the HX5, x3690 X5, and x3850 X5.

Table 2-2 HX5 compared to HX5 with MAX5

	HX5	HX5 with MAX5
Memory bandwidth	16 DDR3 channels at 978 MHz	16 DDR3 channels at 978 MHz + 12 DDR3 channels at 1066 MHz
Memory capacity	128 GB using 8 GB DIMMs	320 GB using 8 GB DIMMs

Table 2-3 x3690 X5 compared to x3690 X5 with MAX5

	x3690 X5	x3690 with MAX5
Memory bandwidth	32 DDR3 channels at 1066 MHz[a]	64 DDR3 channels at 1066 MHz
Memory capacity	512 GB using 16 GB DIMMs	1 TB using 16 GB DIMMs

a. Must install optional mezzanine board

Table 2-4 x3850 X5 compared to x3850 X5 with MAX5

	x3850 X5	x3850 X5 with MAX5
Memory bandwidth	32 DDR3 channels at 1066 MHz	48 DDR3 channels at 1066 MHz
Memory capacity	1 TB using 16 GB DIMMs	1.5 TB using 16 GB DIMMs

For more information about system-specific MAX5 installation options, see the following sections:

- IBM System x3850 X5: “MAX5 memory” on page 79
- IBM System x3690 X5: 4.8.3, “MAX5 memory” on page 136
- IBM BladeCenter HX5: “MAX5 memory population order” on page 198

2.5 Scalability

The architecture of the eX5 servers permits system scaling of up to two nodes on HX5 and x3850 X5. The architecture also supports memory scaling. Figure 2-12 shows these types of scaling.

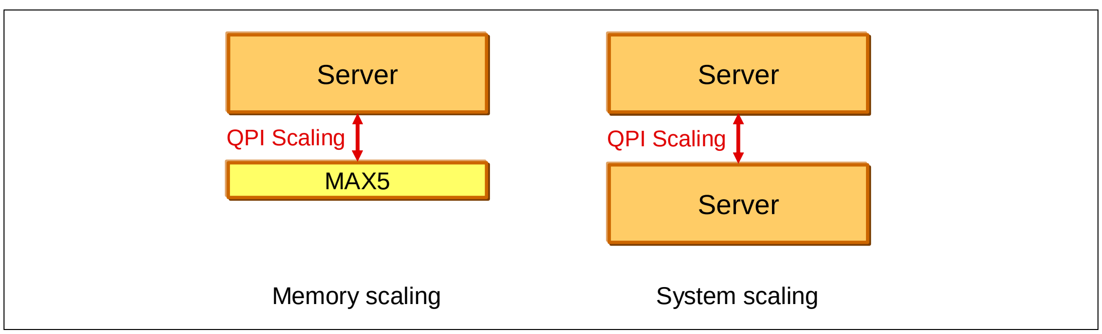


Figure 2-12 Types of scaling with eX5 systems

The x3850 X5 and HX5 both support 2-node system scaling or memory scaling. The x3690 X5 supports memory scaling only.

As shown in Figure 2-12 on page 33, the following scalability is possible:

- Memory scaling: A MAX5 unit can attach to an eX5 server through QPI link cables. This method provides the server with additional memory DIMM slots. We refer to this combination as a *memory-enhanced system*. All eX5 systems support this scaling.
- System scaling: Two servers can connect to form a single system image. The connections are formed by using QPI link cables. The x3850 X5 and HX5 support this type of scaling.

For more information about system-specific scaling options, see the following sections:

- IBM System x3850 X5: 3.6, “Scalability” on page 70
- IBM System x3690 X5: 4.6, “Scalability” on page 128
- BladeCenter HX5: 5.8, “Scalability” on page 188

2.6 Partitioning

You can operate the HX5 scaled system as two independent systems or as a single system, without removing the blade and taking off the side-scale connector. This capability is called *partitioning* and is referred to as IBM FlexNode technology. You partition by using the Advanced Management Module (AMM) in the IBM BladeCenter chassis for the HX5. Figure 2-13 depicts an HX5 system that is scaled to two nodes and an HX5 system that is partitioned into two independent servers.

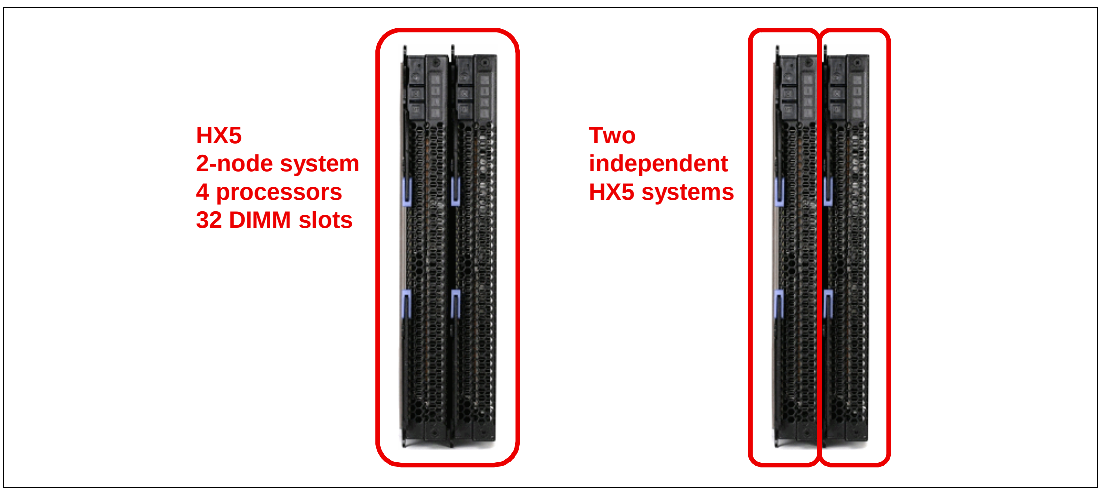


Figure 2-13 HX5 scaling and partitioning

x3690 X5 and x3850 X5: The x3690 X5 and x3850 X5 do not support partitioning.

Figure 2-14 on page 35 and Figure 2-15 on page 35 show the scalable complex configuration options for stand-alone mode through the Advanced Management Module of the BladeCenter chassis.

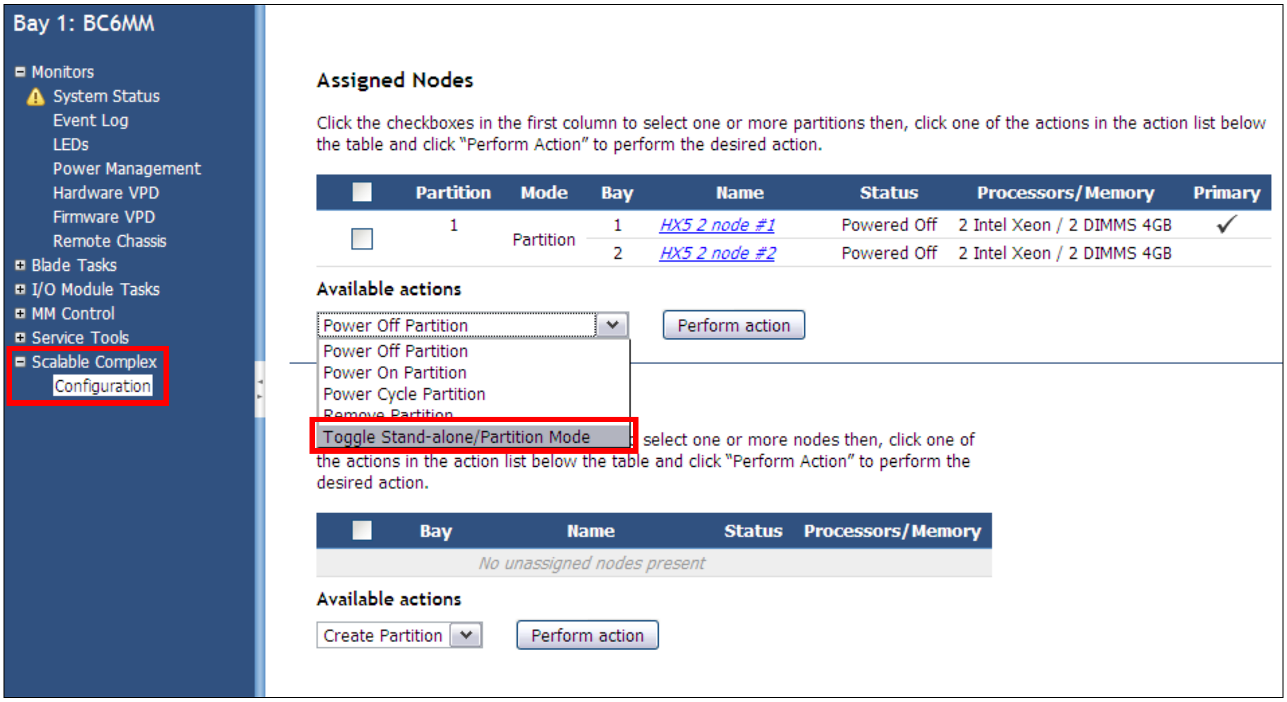


Figure 2-14 Shows option for putting a partition into stand-alone mode

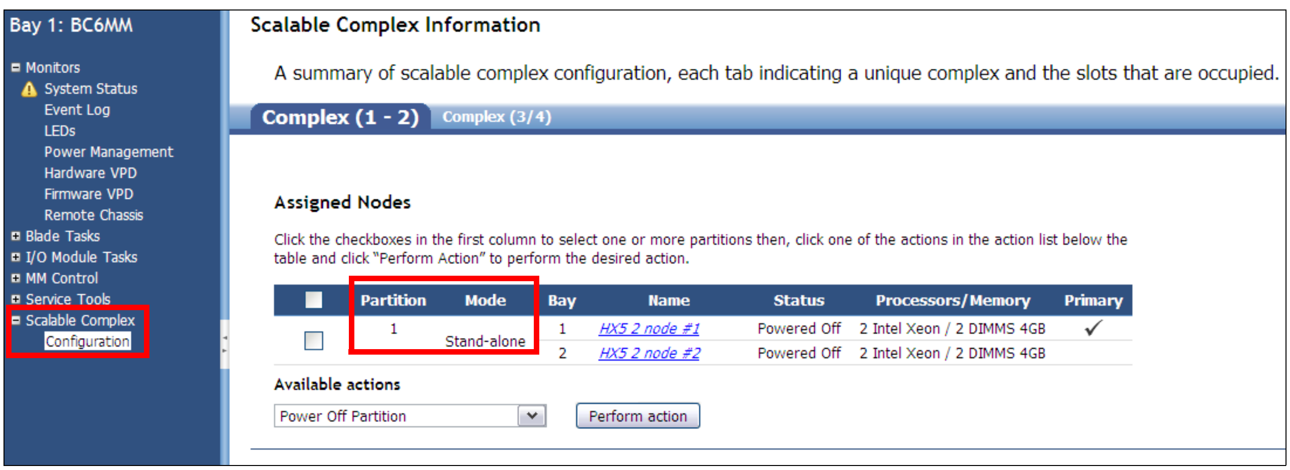


Figure 2-15 HX5 partition in stand-alone mode

The AMM can be accessed remotely, so partitioning can be done without physically touching the systems. Partitioning can allow you to qualify two system types with little additional work, and it allows you more flexibility in system types for better workload optimization.

The HX5 blade, when scaled as a 2-node (4-socket) system, supports FlexNode partitioning as standard.

Before a 2-node HX5 solution can be used, you must create a partition. When the scalability card is added, the two blades still act as single nodes until a partition is made.

For more information about creating a scalable complex, see 8.6, "Creating an HX5 scalable complex" on page 402.

2.7 UEFI system settings

The Unified Extensible Firmware Interface (UEFI) is a pre-boot environment that provides an interface between server firmware and operating system. It replaces BIOS as the software that manages the interface between server firmware, operating system, and hardware initialization, and eliminates the 16-bit, real-mode limitation that BIOS had.

Obtain more information about UEFI at the following website:

http://www.uefi.org/home/

Many of the advanced technology options that are available in the eX5 systems are controlled in the UEFI system settings. They affect processor and memory subsystem performance with regard to the power consumption.

Access the UEFI page by pressing F1 during the system initialization process, as shown in Figure 2-16.

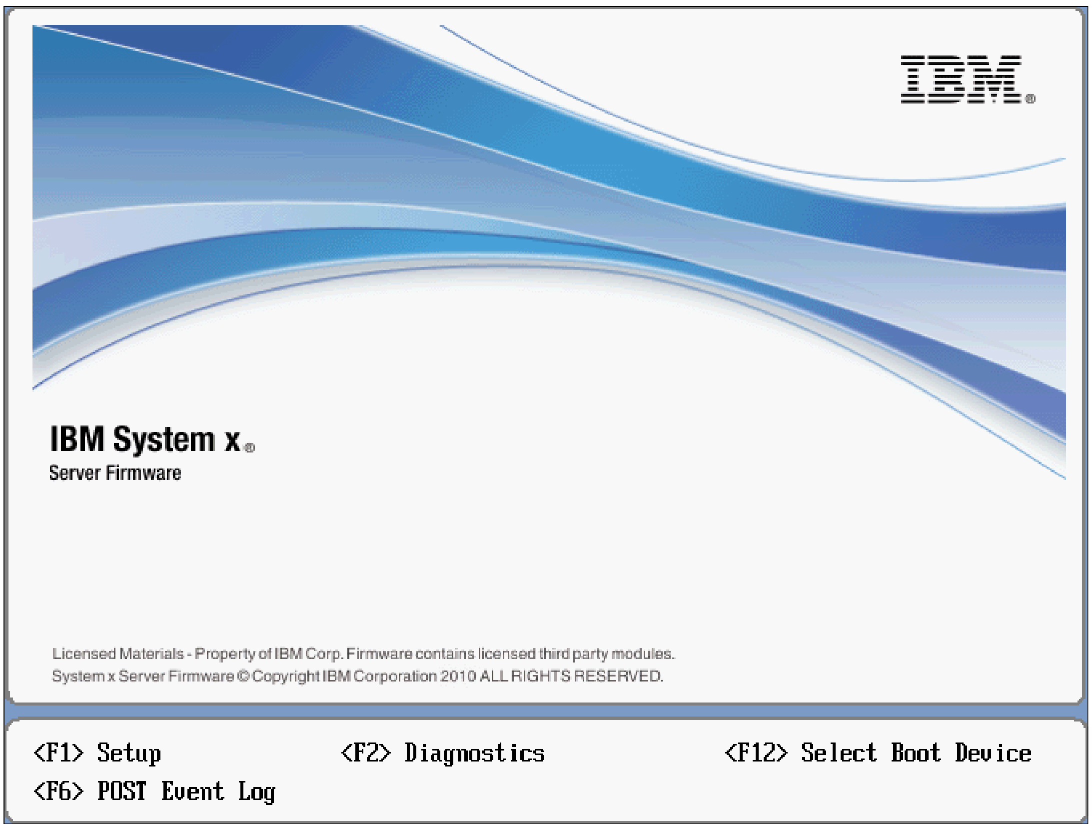


Figure 2-16 UEFI panel on system start-up

Figure 2-17 on page 37 shows the UEFI System Configuration and Boot Management window.

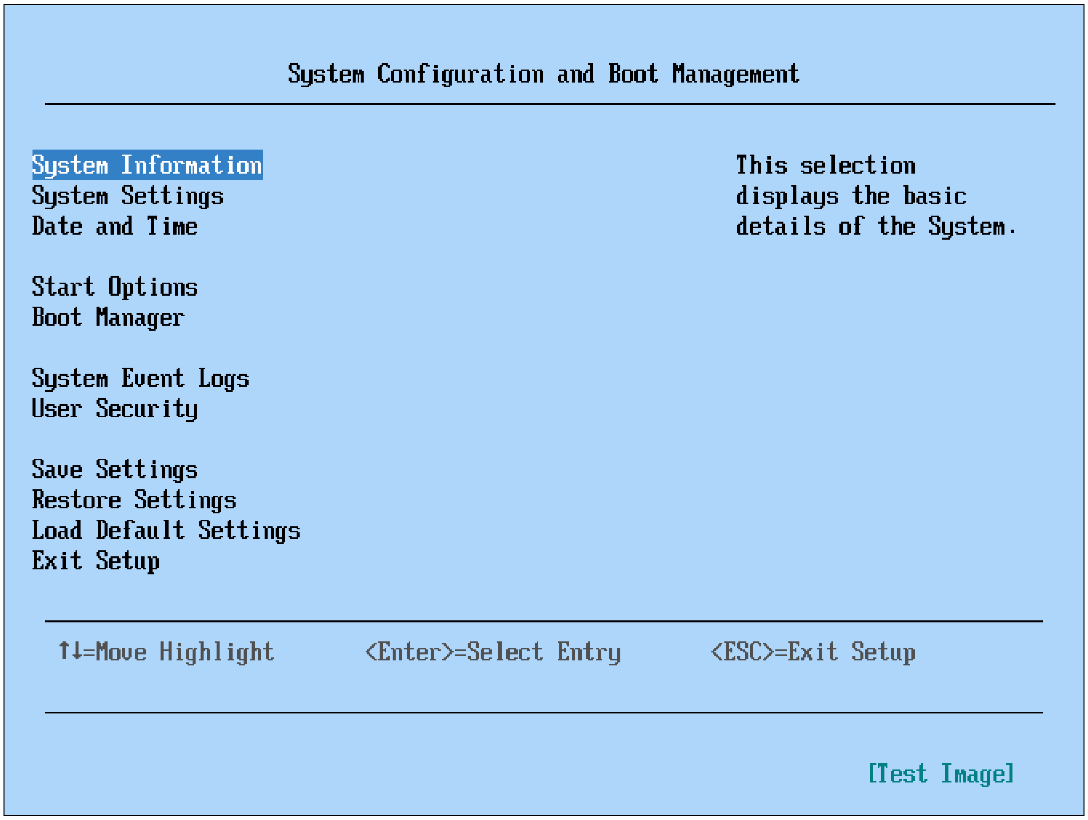


Figure 2-17 UEFI settings main panel

Choose **System Settings** to access the system settings options that we will describe here, as shown in Figure 2-18 on page 38.

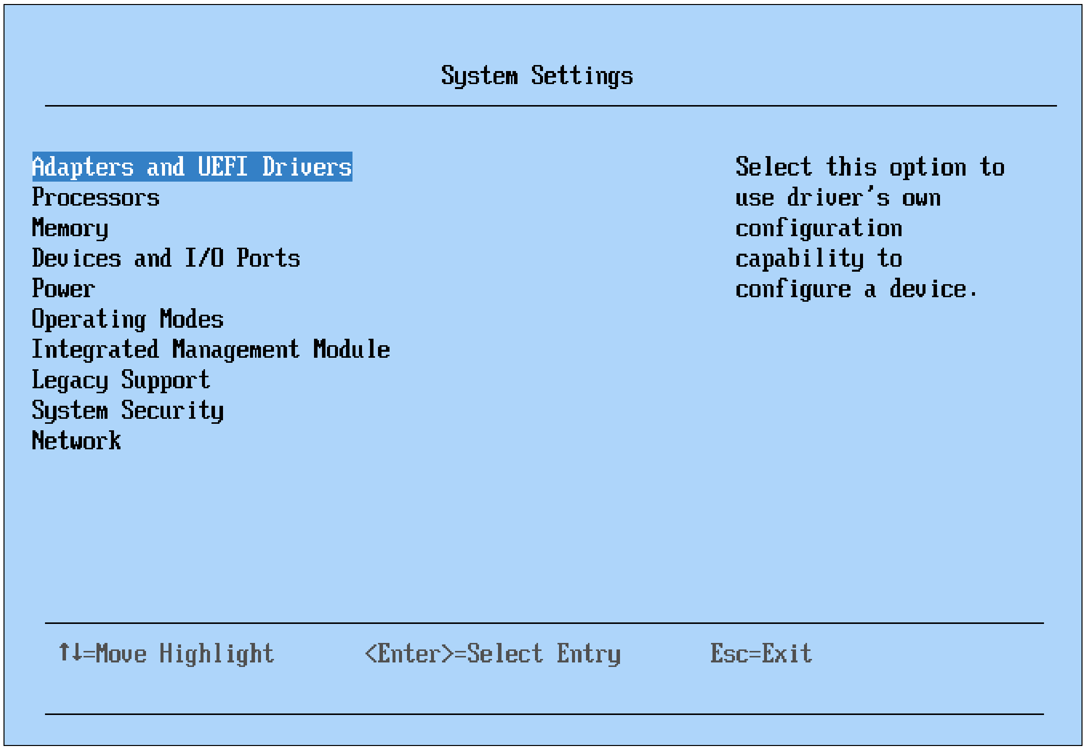


Figure 2-18 UEFI System Settings panel

For more information about system-specific UEFI options, see the following sections:

- IBM System x3850 X5: 6.9, “UEFI settings” on page 259
- IBM System x3690 X5: 7.8, “UEFI settings” on page 337
- IBM BladeCenter HX5: 8.5, “UEFI settings” on page 396

2.7.1 System power operating modes

IBM eX5 servers are designed to provide optimal performance with reasonable power consumption, which depends on the operating frequency and voltage of the processors and memory subsystem. The operating frequency and voltage of the processors and memory subsystem affect the system fan speed that adjusts to the current cooling requirement of the server.

In most operating conditions, the default settings are ideal to provide the best performance possible without wasting energy during off-peak usage. However, for certain workloads, it might be appropriate to change these settings to meet specific power to performance requirements.

The UEFI provides several predefined setups for commonly desired operation conditions. This section describes the conditions for which these setups can be configured.

These predefined values are referred to as *operating modes* and are similar across the entire line of eX5 servers. Access the menu in UEFI by selecting **System Settings** → **Operating Modes** → **Choose Operating Mode**. You see the four operating modes from which to choose, as shown in Figure 2-19 on page 39. When a mode is chosen, the affected settings change to the shown predetermined values.

We describe these modes in the following sections.

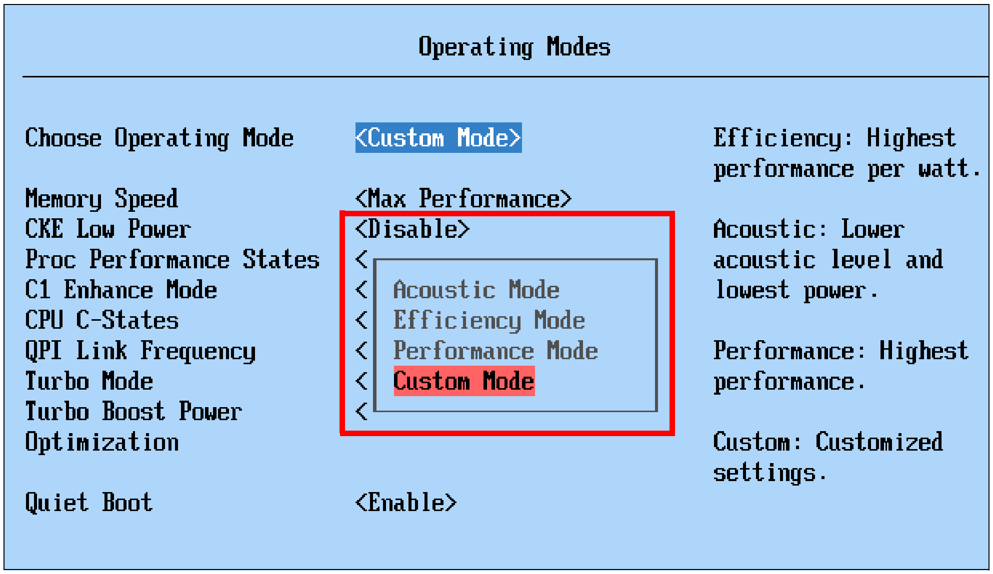


Figure 2-19 Operating modes in UEFI

Acoustic Mode

Figure 2-20 shows the Acoustic Mode predetermined values. They emphasize power-saving server operation for generating less heat and noise. In turn, the system is able to lower the fan speed of the power supplies and the blowers by setting the processors, QPI link, and memory subsystem to a lower working frequency. Acoustic Mode provides lower system acoustics, less heat, and the lowest power consumption at the expense of performance.

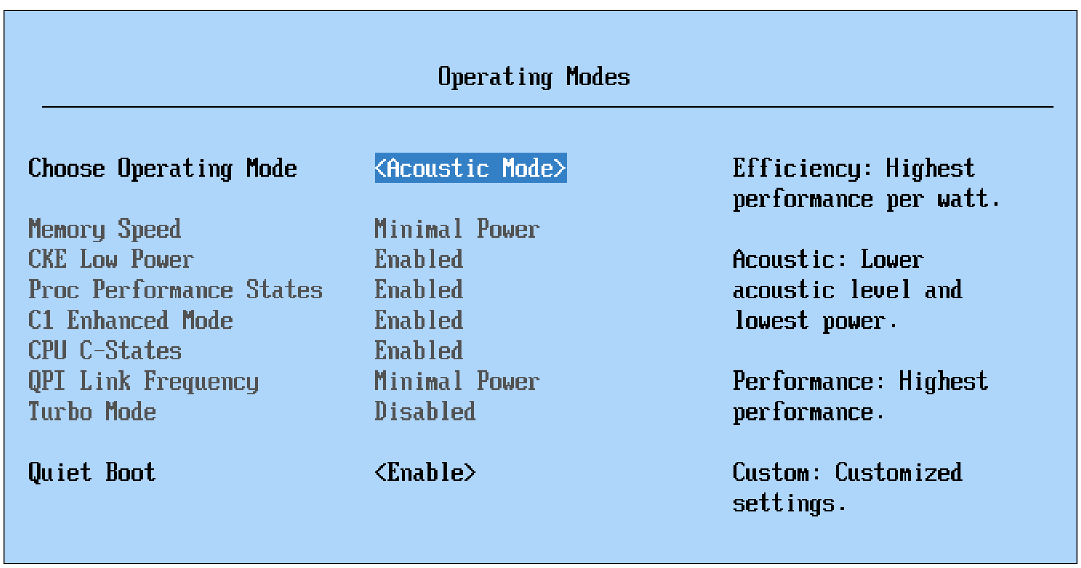


Figure 2-20 Acoustic Mode predetermined values

Efficiency Mode

Figure 2-21 on page 40 shows the Efficiency Mode predetermined values. This operating mode provides the best balance between server performance and power consumption. In short, Efficiency Mode gives the highest performance per watt ratio.

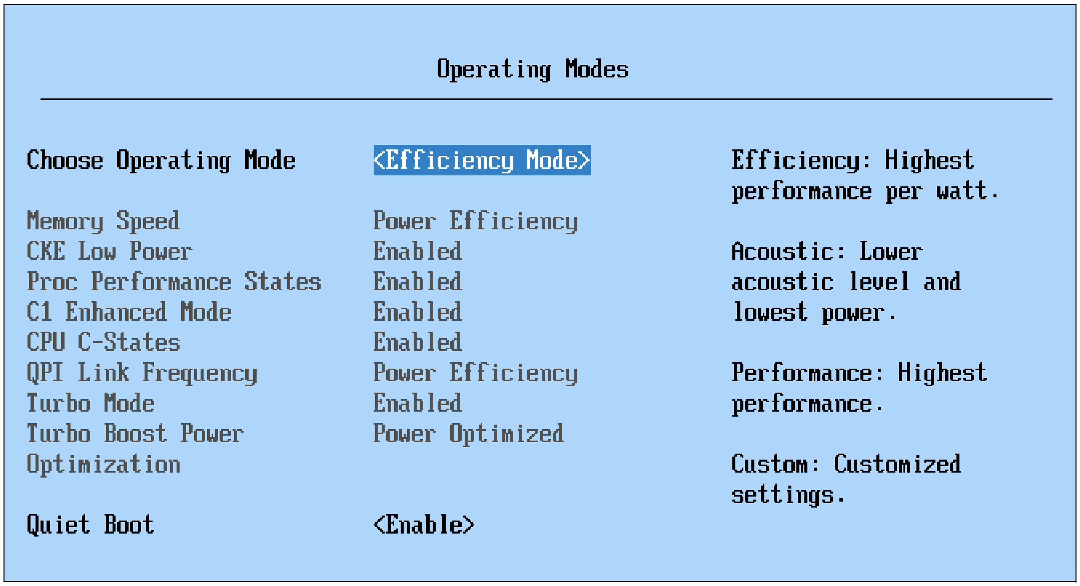


Figure 2-21 Efficiency Mode predetermined values

Performance Mode

Figure 2-22 shows the Performance Mode predetermined values. The server is set to use the maximum performance limits within UEFI. These values include turning off several power management features of the processor to provide the maximum performance from the processors and memory subsystem. Performance Mode provides the best system performance at the expense of power efficiency.

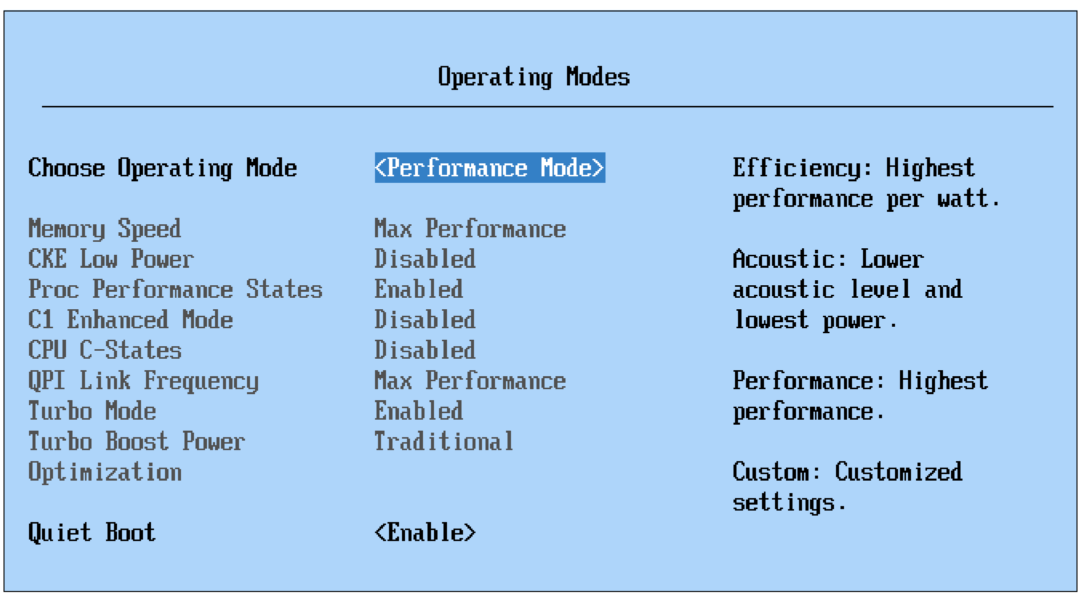


Figure 2-22 Performance Mode predetermined values

Performance Mode is also a good choice when the server runs virtualization workloads. Servers that are used as physical hosts in virtualization clusters often have similar power consumption management at the operating system level. In this environment, the operating system makes decisions about moving and balancing virtual servers across an array of physical host servers. Each virtualized guest operating system reports to a single cluster controller about the resources usage and demand on that physical server. The cluster controller makes decisions to move virtual servers between physical hosts to cater to each guest OS resource requirement and, when possible, shuts down unneeded physical hosts to save power.

Aggressive power management at the hardware level can interfere with the OS-level power management, resulting in a common occurrence where virtual servers move back and forth across the same set of physical host servers. This situation results in an inefficient virtualization environment that responds slowly and consumes more power than necessary. Running the server in Performance Mode prevents this occurrence, in most cases.

Custom Mode

The default value that is set in new eX5 systems is Custom Mode, as shown in Figure 2-23. It is the recommended factory default setting. The values are set to provide optimal performance with reasonable power consumption. However, this mode allows the user to individually set the power-related and performance-related options.

See 2.7.3, “Performance-related individual system settings” on page 43 for a description of individual settings.

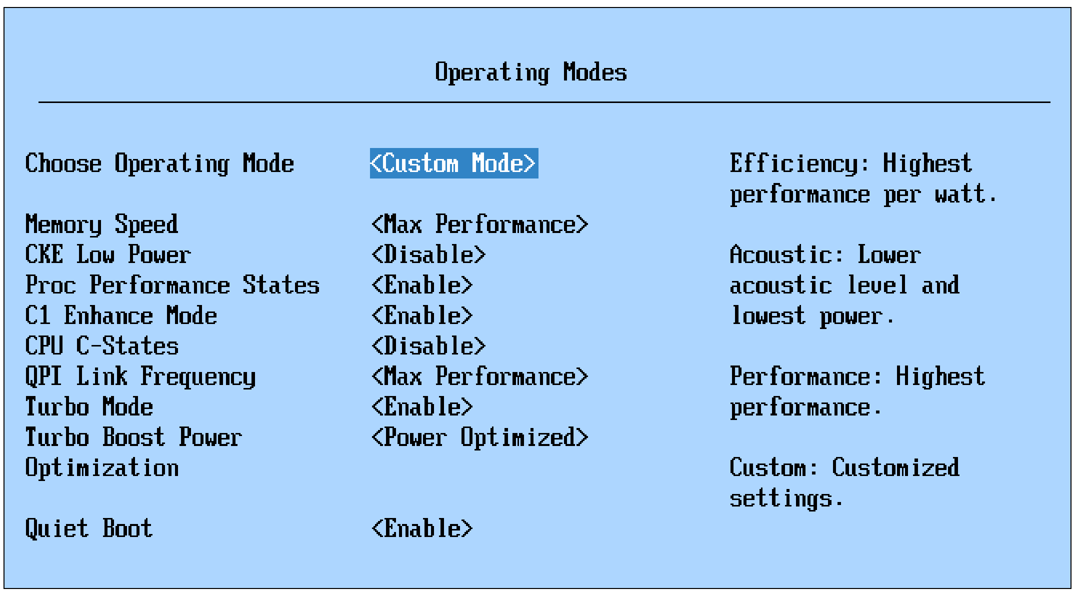


Figure 2-23 Custom Mode factory default values

Table 2-5 shows comparisons of the available operating modes of IBM eX5 servers. Using the Custom Mode, it is possible to run the system using properties that are in-between the predetermined operating modes.

Table 2-5 Operating modes comparison

Settings	Efficiency	Acoustics	Performance	Custom (Default)
Memory Speed	Power Efficiency	Minimal Power	Max Performance	Max Performance
CKE Low Power	Enabled	Enabled	Disabled	Disable
Proc Performance States	Enabled	Enabled	Enabled	Enable
C1 Enhanced Mode	Enabled	Enabled	Disabled	Enable
CPU C-States	Enabled	Enabled	Disabled	Enable

Settings	Efficiency	Acoustics	Performance	Custom (Default)
QPI Link Frequency	Power Efficiency	Minimal Power	Max Performance	Max Performance
Turbo Mode	Enabled	Disabled	Enabled	Enable
Turbo Boost Power Optimization	Power Optimized	-	Traditional	Power Optimized

Additional Settings

In addition to the Operating Mode selection, the UEFI settings under Operating Modes include these additional settings:

- Quiet Boot (Default: *Enable*)

  This mode enables system booting with less information displayed.

- Halt On Severe Error (Default: *Disable,* only available in System x3690 X5)

  This mode enables system boot halt when a severe error event is logged.

2.7.2 System power settings

Power settings include basic power-related configuration options:

- IBM Systems Director Active Energy Manager™ (Default: Capping Enabled)

  The Active Energy Manager option enables the server to use the power capping feature of Active Energy Manager, an extension of IBM Systems Director.

  Active Energy Manager measures, monitors, and manages the energy and thermal components of IBM systems, enabling a cross-platform management solution and simplifying the energy management of IBM servers, storage, and networking equipment. In addition, Active Energy Manager extends the scope of energy management to include non-IBM systems, facility providers, facility management applications, protocol data units (PDUs), and equipment supporting the IPv6 protocol. With Active Energy Manager, you can accurately understand the effect of the power and cooling infrastructure on servers, storage, and networking equipment. One of its features is to set caps for how much power the server can draw.

  Learn more about IBM Systems Director Active Energy Manager at the following website:

  `http://www.ibm.com/systems/software/director/aem/`

- Power Restore Policy (Default: Restore)

  This option defines system behavior after a power loss.

Figure 2-24 on page 43 shows the available options in the UEFI system Power settings.

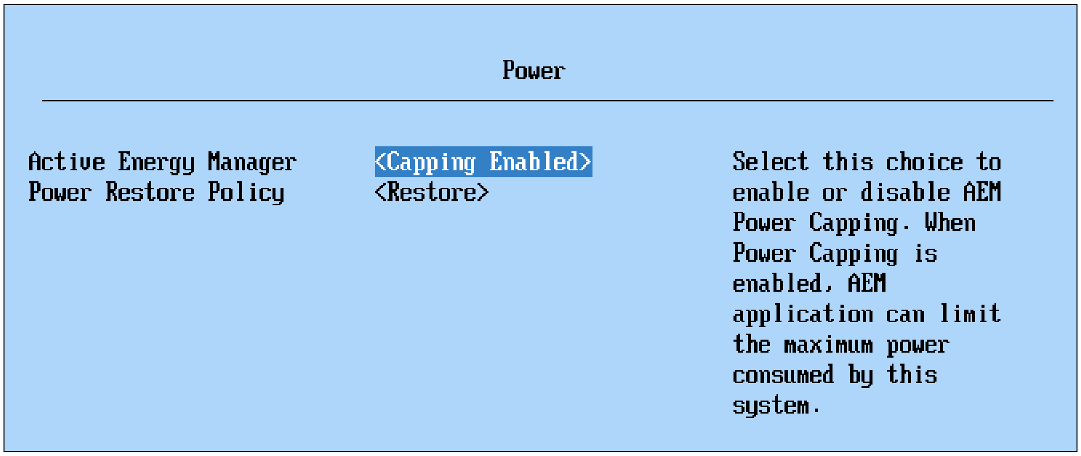


Figure 2-24 UEFI Power settings window

2.7.3 Performance-related individual system settings

The UEFI default settings are configured to provide optimal performance with reasonable power consumption. Other operating modes are also available to meet various power and performance requirements. However, individual system settings enable users to fine-tune the desired characteristics of the IBM eX5 servers.

This section describes the UEFI settings that are related to system performance. Remember that, in most cases, increasing system performance increases the power consumption of the system.

Processors

Processor settings control the various performance and power features that are available on the installed Xeon processor.

Figure 2-25 on page 44 shows the UEFI Processor system settings window with the default values.

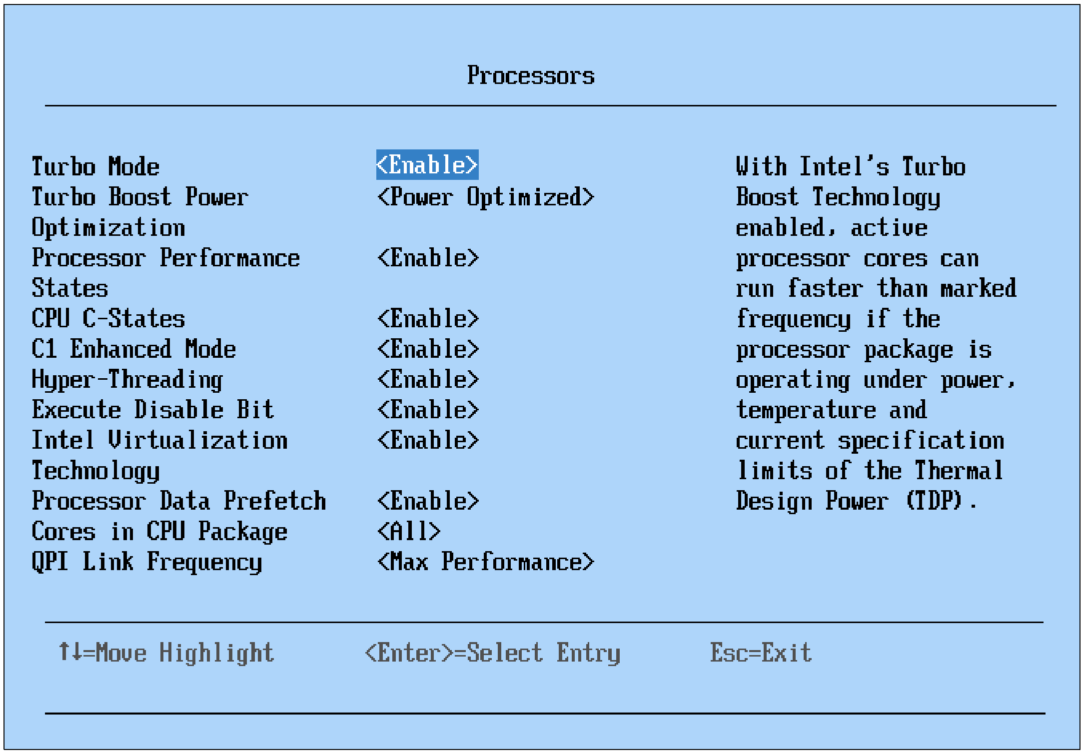


Figure 2-25 UEFI Processor system settings panel

The following processor feature options are available:

- Turbo Mode (Default: *Enable*)

  This mode enables the processor to increase its clock speed dynamically as long as the CPU does not exceed the *Thermal Design Power (TDP)* for which it was designed. See 2.2.3, "Turbo Boost Technology" on page 18 for more information.

- Turbo Boost Power Optimization (Default: *Power Optimized*)

  This option specifies which algorithm to use when determining whether to overclock the processor cores in Turbo Mode:

  – *Power Optimized* provides reasonable Turbo Mode in relation to power consumption.

    Turbo Mode does not engage unless additional performance has been requested by the operating system for a period of 2 seconds.

  – *Traditional* provides a more aggressive Turbo Mode operation.

    Turbo Mode engages as more performance is requested by the operating system.

- Processor Performance States (Default: *Enable*)

  This option enables Intel Enhanced SpeedStep Technology that controls dynamic processor frequency and voltage changes, depending on operation.

- CPU C-States (Default: *Enable*)

  This option enables dynamic processor frequency and voltage changes in the idle state, providing potentially better power savings.

- C1 Enhanced Mode (Default: *Enable*)

  This option enables processor cores to enter an enhanced halt state to lower the voltage requirement, and it provides better power savings.

- Hyper-Threading (Default: *Enable*)

  This option enables logical multithreading in the processor, so that the operating system can execute two threads simultaneously for each physical core.

- Execute Disable Bit (Default: *Enable*)

  This option enables the processor to disable the execution of certain memory areas, therefore preventing buffer overflow attacks.

- Intel Virtualization Technology (Default: *Enable*)

  This option enables the processor hardware acceleration feature for virtualization.

- Processor Data Prefetch (Default: *Enable*)

  This option enables the memory data access prediction feature to be stored in the processor cache.

- Cores in CPU Package (Default: *All*)

  This option sets the number of processor cores to be activated.

- QPI Link Frequency (Default: *Max Performance*)

  This option sets the operating frequency of the processor's QPI link:

  – Minimal Power provides less performance for better power savings.

    The QPI link operates at the lowest frequency, which, in the eX5 systems, is 4.8 GT/s.

  – Power Efficiency provides best performance per watt ratio.

    The QPI link operates 1 step under the rated frequency, that is, 5.86 GT/s for processors rated at 6.4 GT/s.

  – *Max Performance* provides the best system performance.

    The QPI link operates at the rated frequency, that is, 6.4 GT/s for processors rated at 6.4 GT/s.

Memory

The Memory settings window provides the available memory operation options, as shown in Figure 2-26 on page 46.

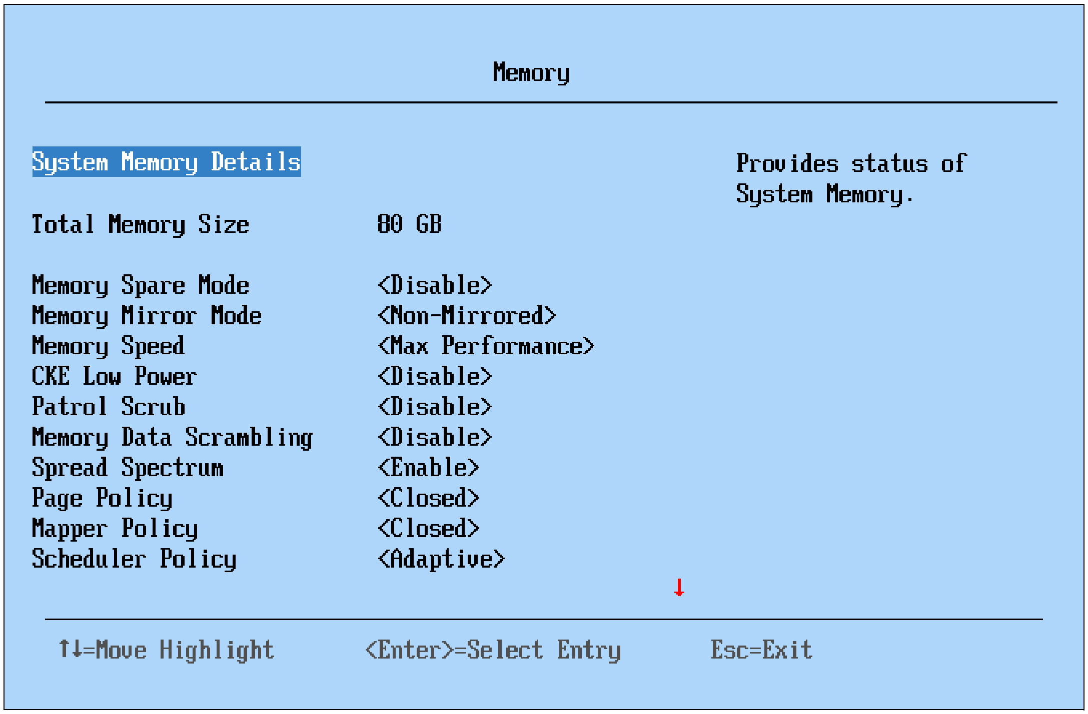


Figure 2-26 UEFI Memory system settings panel

The following memory feature options are available:

- Memory Spare Mode (Default: *Disable*)

  This option enables memory sparing mode, as described in "Memory sparing" on page 29.

- Memory Mirror Mode (Default: *Non-Mirrored*)

  This option enables memory mirroring mode, as described in "Memory mirroring" on page 28.

  > **Memory Mirror Mode:** Memory Mirror Mode cannot be used in conjunction with Memory Spare Mode.

- Memory Speed (Default: *Max Performance*)

  This option sets the operating frequency of the installed DIMMs:

  – Minimal Power provides less performance for better power savings.

    The memory operates at the lowest supported frequency, which, in the eX5 systems, is 800 MHz.

  – Power Efficiency provides the best performance per watt ratio.

    The memory operates one step under the rated frequency, that is, 977 MHz for DIMMs that are rated at 1066 MHz or higher.

  – *Max Performance* provides the best system performance.

    The memory operates at the rated frequency, that is, 1066 MHz for DIMMs rated at 1066 MHz or higher.

    > **Tip:** Although memory DIMMs rated at 1333MHz are supported on eX5 servers, the currently supported maximum memory operating frequency is 1066 MHz.

- CKE Low Power (Default: *Disable*)

  This option enables the memory to enter a low-power state for power savings by reducing the signal frequency.

- Patrol Scrub (Default: *Disable*)

  This option enables scheduled background memory scrubbing before any error is reported, as opposed to default demand scrubbing on an error event. This option provides better memory subsystem resiliency at the expense of a small performance loss.

- Memory Data Scrambling (Default: *Disable*)

  This option enables a memory data scrambling feature to further minimize bit-data errors.

- Spread Spectrum (Default: *Enable*)

  This option enables the memory spread spectrum feature to minimize electromagnetic signal interference in the system.

- Page Policy (Default: *Closed*)

  This option determines the Page Manager Policy in evaluating memory access:

  – *Closed:* Memory pages are closed immediately after each transaction.

  – Open: Memory pages are left open for a finite time after each transaction for possible recurring access.

  – Adaptive: Use Adaptive Page Policy to decide the memory page state.

  – Multi-CAS Widget: The widget allows multiple consecutive column address strobes (CAS) to the same memory ranks and banks in the Open Page Policy.

- Mapper Policy (Default: *Closed*)

  This option determines how memory pages are mapped to the DIMM subsystem.

  – *Closed:* Memory is mapped closed to prevent DIMMs from being excessively addressed.

  – Open: Memory is mapped open to decrease latency.

- Scheduler Policy (Default: *Adaptive*)

  This option determines the scheduling mode optimization based on memory operation:

  – Static Trade Off: Equal trade-off between read/write operation latency
  – Static Read Primary: Minimize read latency and consider reads as primary operation
  – Static Write Primary: Minimize write latency and consider writes as primary operation
  – *Adaptive:* Memory scheduling adaptive to system operation

- MAX5 Memory Scaling Affinity (Default: *Non-Pooled*)

  – The *Non-Pooled* option splits the memory in the MAX5 and assigns it to each of the installed processors.

  – The Pooled option presents the additional memory in the MAX5 as a pool of memory that is not assigned to any particular processor.

2.8 IBM eXFlash

IBM eXFlash is the name given to the eight 1.8-inch solid-state drives (SSDs), the backplanes, SSD hot-swap carriers, and indicator lights that are available for the x3690 X5, x3850 X5, and x3950 X5.

Each eXFlash can replace four 2.5-inch serial-attached SCSIs (SAS) or (SSDs). You can install the following number of eXFlash units:

- The x3850 X5 can have either of the following configurations:
  - Up to four SAS or Serial Advanced Technology Attachment (SATA) drives, plus the eight SSDs in one eXFlash unit
  - Sixteen SSDs in two eXFlash units
- The x3950 X5 database-optimized models have one eXFlash unit standard with space for eight SSDs, and a second eXFlash is optional.

  The x3690 X5 can have up to 24 SSDs in three eXFlash units.

Spinning disks, although an excellent choice for cost per megabyte, are not always the best choice when considered for their cost per I/O operation per second (IOPS).

In a production environment where the tier-one capacity requirement can be met by IBM eXFlash, the total cost per IOPS can be lower than any solution requiring attachment to external storage. Host bus adapters (HBAs), switches, controller shelves, disk shelves, cabling, and the actual disks all carry a cost. They might even require an upgrade to the machine room infrastructure, for example, a new rack or racks, additional power lines, or perhaps additional cooling infrastructure.

Also, remember that the storage acquisition cost is only a part of the total cost of ownership (TCO). TCO takes the ongoing cost of management, power, and cooling for the additional storage infrastructure detailed previously. SSDs use only a fraction of the power, generate only a fraction of the heat that spinning disks generate, and, because they fit in the chassis, are managed by the server administrator.

IBM eXFlash is optimized for a heavy mix of read and write operations, such as transaction processing, media streaming, surveillance, file copy, logging, backup and recovery, and business intelligence. In addition to its superior performance, eXFlash offers superior uptime with three times the reliability of mechanical disk drives. SSDs have no moving parts to fail. They use Enterprise Wear-Leveling to extend their use even longer. All operating systems that are listed in ServerProven® for each machine are supported for use with eXFlash.

The eXFlash SSD backplane uses two long SAS cables, which are included with the backplane option. If two eXFlash backplanes are installed, four cables are required. You can connect the eXFlash backplane to the dedicated RAID slot if desired.

In a system that has two eXFlash backplanes installed, two controllers are required in PCIe slots 1 - 4 to control the drives; however, up to four controllers can be used. In environments where RAID protection is required, use two RAID controllers per backplane to ensure that peak IOPS can be reached. Although use of a single RAID controller results in a functioning solution, peak IOPS can be reduced by a factor of approximately 50%. Remember that each RAID controller controls only its own disks. With four B5015 controllers, each controller controls four disks. The effect of RAID-5 is that four disks (one per array) are used for parity.

You can use both RAID and non-RAID controllers. The IBM 6Gb SSD Host Bus Adapter (HBA) is optimized for read-intensive environments, and you can achieve maximum performance with only a single 6Gb SSD HBA. RAID controllers are a better choice for environments with a mix of read and write activity.

The eXFlash units can connect to the same types of ServeRAID disk controllers as the SAS and SATA disks. For higher performance, connect them to the IBM 6Gb SAS HBA or the ServeRAID B5015 SSD Controller.

In addition to using less power than rotating magnetic media, the SSDs are more reliable, and they can service many more IOPS. These attributes make them well suited to I/O intensive applications, such as complex queries of databases.

Figure 2-27 shows an eXFlash unit, with the status lights assembly on the left side.

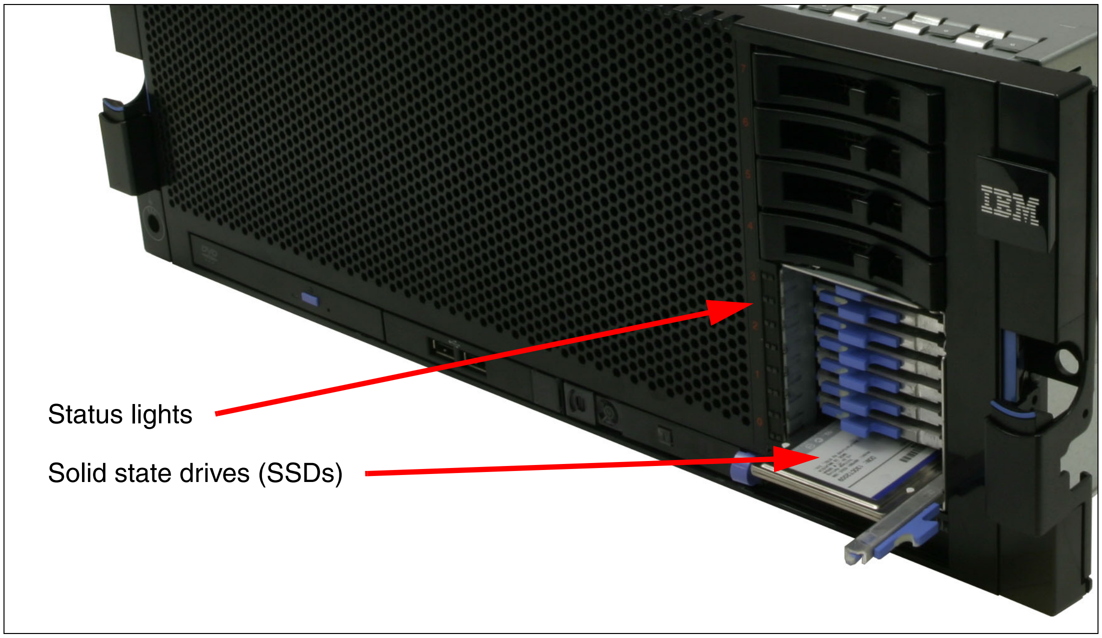


Figure 2-27 x3850 X5 with one eXFlash

For more information about system-specific memory eXFlash options, see the following sections:

- IBM System x3850 X5: 3.9.3, "IBM eXFlash and 1.8-inch SSD support" on page 93
- IBM System x3690 X5: 4.9.2, "IBM eXFlash and SSD disk support" on page 149.

2.8.1 IBM eXFlash price-performance

The information in this section gives an idea of the *relative* performance of spinning disks when compared with the SSDs in IBM eXFlash. This section does not guarantee that these data rates are achievable in a production environment because of the number of variables involved. However, in most circumstances, we expect the scale of the performance differential between these two product types to remain constant.

If we take a 146 GB, 15K RPM 2.5-inch disk drive as a baseline and assume that it can perform 300 IOPS, we can also state that eight disks can provide 2400 IOPS. At a current US list price per drive of USD579 (multiplied by eight = USD4,632), that works out to USD1.93 per IOP and USD4 per GB.

I/O operations per second (IOPS): IOPS is used predominantly as a measure for database performance. Workloads measured in IOPS are typically sized by taking the realistically achievable IOPS of a single disk and multiplying the number of disks until the anticipated (or measured) IOPS in the target environment is reached. Additional factors, such as the RAID level, number of HBAs, and storage ports can also affect the performance. The key point is that IOPS-driven environments traditionally require significant disk. When sizing, exceeding the requested capacity to reach the required number of IOPS is often necessary.

Under similar optimized benchmarking conditions, eight of the 50 GB, 1.8-inch SSDs are able to sustain 48,000 read IOPS and, in a separate benchmark, 16,000 write IOPS. The cost of USD12,000 for the SSDs works out at approximately USD0.25 per IOP and USD60 per gigabyte.

Additional spinning disks create additional costs in terms of shelves, rack space, and power and cooling, none of which are applicable for the SSDs, driving their total TCO down even further. The initial cost per GB is higher for the SSDs, but view it in the context of TCO over time.

For more information regarding each of the eX5 systems, see the following sections:

- 3.9, “Storage” on page 90
- 5.11, “Storage” on page 203

2.9 Integrated virtualization

This section contains a list of virtualization options that are optional within the eX5 series.

2.9.1 VMware ESXi

ESXi is an embedded version of VMware ESX. The footprint of ESXi is small (approximately 32 MB) because it does not use the Linux-based Service Console. Instead, it uses management tools, such Virtual Center, Remote Command-Line Interface (CLI), and Common Information Model (CIM) for standards-based and agentless hardware monitoring. VMware ESXi includes full VMware File System (VMFS) support across Fibre Channel and iSCSI SAN, and network-attached storage (NAS). It supports 4-way Virtual Symmetrical Multiprocessor Systems (SMP) (VSMP). ESXi 4.0 supports 64 CPU threads, for example, eight x 8-core CPUs, and can address 1 TB of RAM.

The VMware ESXi 4.0 and 4.1 embedded virtualization keys for the x3850 X5, x3690 X5, and HX5 are orderable, as listed in Table 2-6.

Table 2-6 VMware ESXi 4.x memory key

Part number	Feature code	Description
41Y8278	1776	IBM USB Memory Key for VMware ESXi 4.0
41Y8287	2420	IBM USB Memory Key for VMware ESXi 4.1 with MAX5

2.9.2 Red Hat RHEV-H (KVM)

The Kernel-based Virtualization Machine (KVM hypervisor) that is supported with Red Hat Enterprise Linux (RHEL) 5.4 and later is available on the x3850 X5. RHEL-H (KVM) is standard with the purchase of RHEL 5.4 and later. All hardware components that have been tested with RHEL 5.*x* are also supported running RHEL 5.4 (and later), and they are supported to run RHEV-H (KVM). IBM Support Line and Remote Technical Support (RTS) for Linux support RHEV-H (KVM).

RHEV-H (KVM) supports 96 CPU threads (an 8-core processor with Hyper-Threading enabled has 16 threads) and can address 1 TB RAM.

KVM includes the following features:

- Advanced memory management support
- Robust and scalable Linux virtual memory manager
- Support for large memory systems with greater than 1 TB RAM
- Support for nonuniform memory access (NUMA)
- Transparent memory page sharing
- Memory overcommit

KVM also provides the following advanced features:

- Live migration
- Snapshots
- Memory Page sharing
- SELinux for high security and isolation
- Thin provisioning
- Storage overlays

2.9.3 Windows 2008 R2 Hyper-V

Windows® 2008 R2 Hyper-V is also supported to run on the eX5 servers. You can confirm Hyper-V support in ServerProven.

2.10 Changes in technology demand changes in implementation

This section introduces new implementation concepts that are now available due to the new technology that has been made available in the IBM eX5 servers.

2.10.1 Using swap files

With the introduction of large amounts of addressable memory, when using an UEFI-aware 64-bit operating system, the question that comes to mind with a non-virtualized operating system is, “Do I continue to use a swap file to increase the amount of usable memory that an operating system can use?” The answer is no.

Using a swap file introduces memory page swaps that take milliseconds to perform as opposed to possible remote memory access on a MAX5 that will take nanoseconds to perform. Not using a swap file improves the performance of the single 64-bit operating system.

Note, however, that when using SSD drives as your primary storage for the operating system, remember that it is better to not have an active swap file on this type of storage. SSD drives are designed to support a large but finite number of write operations to any single 4k storage cell on the drive (to the order of 1 million write operations). After that limit has been reached, the storage cell is no longer usable. As storage cells begin to die, the drive automatically maps around them, but when enough cells fail, the drive first reports a Predictive Failure Analysis (PFA) and then eventually fails. Therefore, you must be careful determining how dynamic the data is that is being stored on SSD storage. Memory swap file space must never be assigned to SSD storage. When you must use memory swap files, assign the swap file space to conventional SAS or SATA hard drives.

2.10.2 SSD drives and battery backup cache on RAID controllers

When using conventional SAS or SATA hard drives on a ServeRAID controller, it is common practice to enable writeback cache for the logical drive to prevent data corruption if a loss of power occurs.

With SATA SSD drives, writes to the drives are immediately stored in the memory of the SSD drive. The potential for loss of data is dramatically reduced. Writing to writeback cache first and then to the SSD drives actually increases the latency time for writing the data to the SSD device.

Today's SSD optimized controllers have neither read nor writeback cache. If you are in a solid SSD environment, the best practice is to not install a RAID battery and to not enable cache. When your storage uses a mixed media environment, the best practice is to use a ServeRAID-5*xxx* controller with the IBM ServeRAID 5000 Performance Accelerator Key.

We describe this topic in detail in the following sections:

- IBM System x3690 X5: 4.9.1, "2.5-inch SAS drive support" on page 145
- IBM System x3850 X5: "ServeRAID M5000 Series Performance Accelerator Key" on page 95

2.10.3 Increased resources for virtualization

The huge jump in processing capacity and memory allows for the consolidation of services, while still maintaining fault tolerance using scalable clustered host solutions. As your servers approach peak demand, additional hosts can be automatically powered on and activated to spread the computing demand to additional virtual servers as demand increases. As the peak demand subsides, the same environment can automatically consolidate virtual servers to a smaller group of active hosts, saving power while still maintaining true fault tolerance.

By using larger servers with built-in redundancy for power, fans, storage access, and network access, it is now possible to combine the functional requirements of a dozen or more servers into a dual-hosted virtual server environment that can withstand the possible failure of a complete host. As demand increases, the number of hosts can be increased to maintain the same virtual servers, with no noticeable changes or programming costs to allow the same virtual server to function in the new array of hosts.

With this capability, the server becomes an intelligent switch in the network. Instead of trying to balance network traffic through various network adapters on various servers, you can now create a virtual network switch inside a cluster of host servers to which the virtual servers logically attach. All of the physical network ports of the server, provided that they are the same type of link, can be aggregated into a single IEEE 802.3ad load balance link to maximize link utilization between the server and the external network switch. Two scaled x3850 X5s running in a clustered virtualized environment can replace an entire 42U rack of conventional 1U servers, and their associated top rack network and SAN switch.

2.10.4 Virtualized Memcached distributed memory caching

Many web content providers and light provisioning providers use servers designed for speed, and not fault tolerance, to store the results from database or API calls so that clients can be redirected from the main database server to a MemCached device for all future pages within the original database lookup. This capability allows the database or web content server to off-load the processing time that is needed to maintain those client sessions. You can define the same physical servers as virtual servers with access to a collection of SSD drives. The number of virtual servers can be dynamically adjusted to fit the demands of the database or web content server.

3

IBM System x3850 X5 and x3950 X5

In this chapter, we introduce the IBM System x3850 X5 and the IBM System x3950 X5. The x3850 X5 and x3950 X5 are the follow-on products to the eX4-based x3850 M2, and like their predecessor, are 4-socket systems. The x3950 X5 models are optimized for specific workloads, such as virtualization and database workloads.

The MAX5 memory expansion unit is a 1U device, which you connect to the x3850 X5 or x3950 X5, and provides the server with an additional 32 DIMM sockets. It is ideal for applications that can take advantage of as much memory as is available.

This chapter contains the following topics:

3.1 Product features

The IBM System x3850 X5 and x3950 X5 servers address the following requirements that many IBM enterprise clients need:

- The ability to have increased performance on a smaller IT budget
- The ability to increase database and virtualization performance without having to add more CPUs, especially when software is licensed on a per-socket basis
- The ability to add memory capacity on top of existing processing power, so that the overall performance goes up while software licensing costs remain static
- The flexibility to achieve the desired memory capacity with larger capacity single DIMMs
- The ability to pay for the system they need today, with the capability to grow both memory capacity and processing power when necessary in the future

The base building blocks of the solution are the x3850 X5 server and the MAX5 memory expansion drawer. The x3850 X5 is a 4U system with four processor sockets and up to 64 DIMM sockets. The MAX5 memory expansion drawer is a 1U device that adds 32 DIMM sockets to the server.

The x3950 X5 is the name for the preconfigured IBM models, for specific workloads. The announced x3950 X5 models are optimized for database applications. Future x3950 X5 models will include models that are optimized for virtualization.

> **Referring to the models:** Throughout this chapter, where a feature is not unique to either the x3850 X5 or the x3950 X5 but is common to both models, the term *x3850 X5* is used.

3.1.1 IBM System x3850 X5 product features

IBM System x3850 X5, machine type 7145, is the follow-on product to the IBM System x3850 M2 and x3950 M2. It is a 4U 4-socket Intel 7500-based (Nehalem-EX) platform with 64 DIMM sockets. It can be scaled up to eight processor sockets depending on model and 128 DIMM sockets by connecting a second server to form a single system image and maximize performance, reliability, and scalability.

The x3850 X5 is targeted at enterprise clients looking for increased consolidation opportunities with expanded memory capacity.

See Table 3-1 on page 62 for a comparison of eX4 x3850 M2 and eX5 x3850 X5.

The x3850 X5 offers the following key features:

- Four Xeon 7500 series CPUs (4 core, 6 core, and 8 core)
- Scalable to eight sockets by connecting two x3850 X5 servers
- 64 DDR3 DIMM sockets
- Up to eight memory cards can be installed, each with eight DIMM slots
- Seven PCIe 2.0 slots (one slot contains the Emulex 10Gb Ethernet dual-port adapter)
- Up to eight 2.5-inch hard disk drives (HDDs) or sixteen 1.8-inch solid-state drives (SSDs)
- RAID-0 and RAID-1 standard; optional RAID-5 and 50, RAID-6 and 60, and encryption
- Two 1 Gb Ethernet ports
- One Emulex 10Gb Ethernet dual-port adapter (standard on all models, except ARx)
- Internal USB for embedded hypervisor (VMware and Linux hypervisors)
- Integrated management module

The x3850 X5 has the following physical specifications:

- Width: 440 mm (17.3 in.)
- Depth: 712 mm (28.0 in.)
- Height: 173 mm (6.8 in.) or 4 rack units (4U)
- Minimum configuration: 35.4 kg (78 lb.)
- Maximum configuration: 49.9 kg (110 lb.)

Figure 3-1 shows the x3850 X5.

Figure 3-1 Front view of the x3850 X5 showing eight 2.5-inch SAS drives

In Figure 3-1, two serial-attached SCSI (SAS) backplanes have been installed (at the right of the server). Each backplane supports four 2.5-inch SAS disks (eight disks in total).

Notice the orange colored bar on each disk drive. This bar denotes that the disks are hot-swappable. The color coding used throughout the system is orange for hot-swap and blue for non-hot-swap. Changing a hot-swappable component requires no downtime; changing a non-hot-swappable component requires that the server is powered off before removing that component.

Figure 3-2 on page 58 shows the major components inside the server and on the front panel of the server.

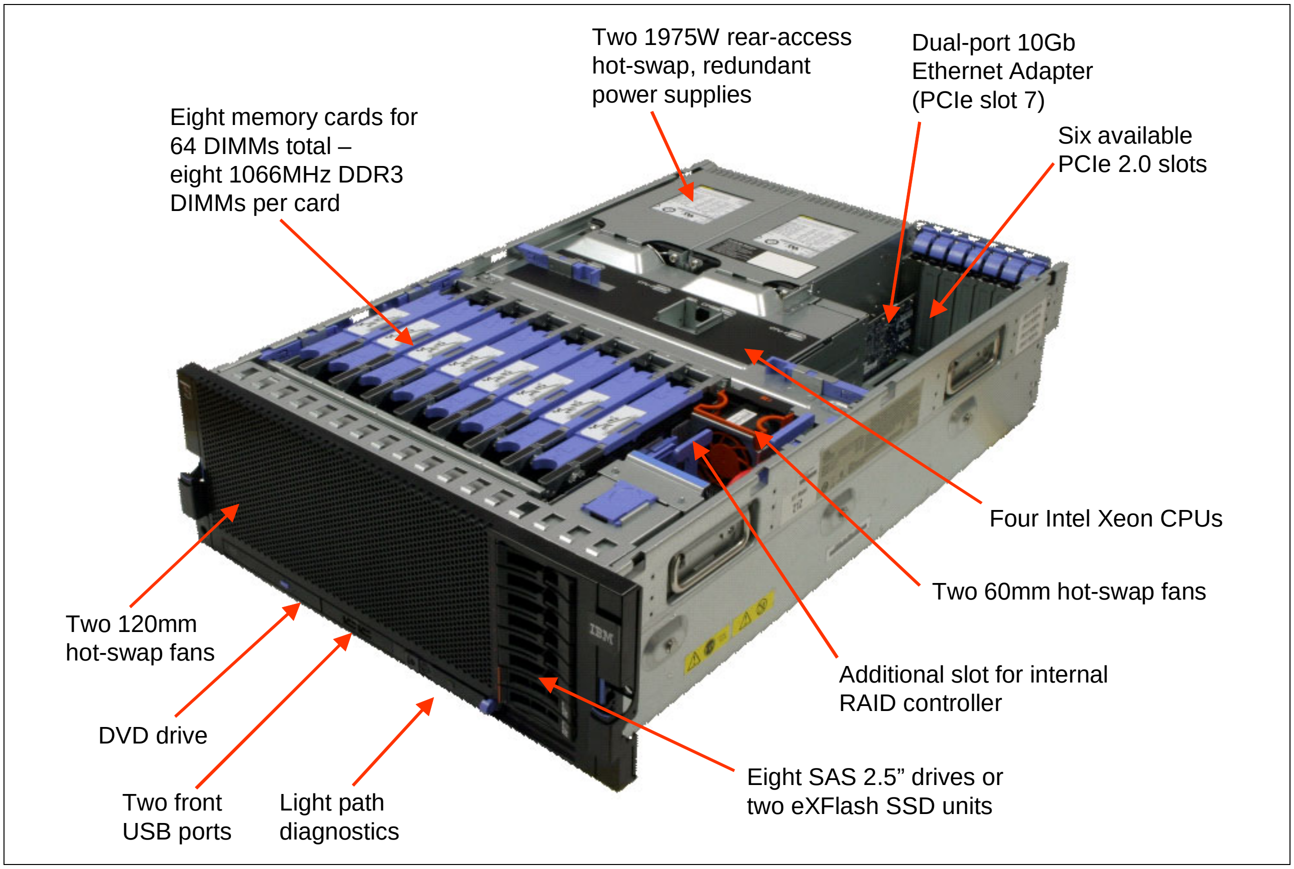


Figure 3-2 x3850 X5 internals

Figure 3-3 shows the connectors at the back of the server.

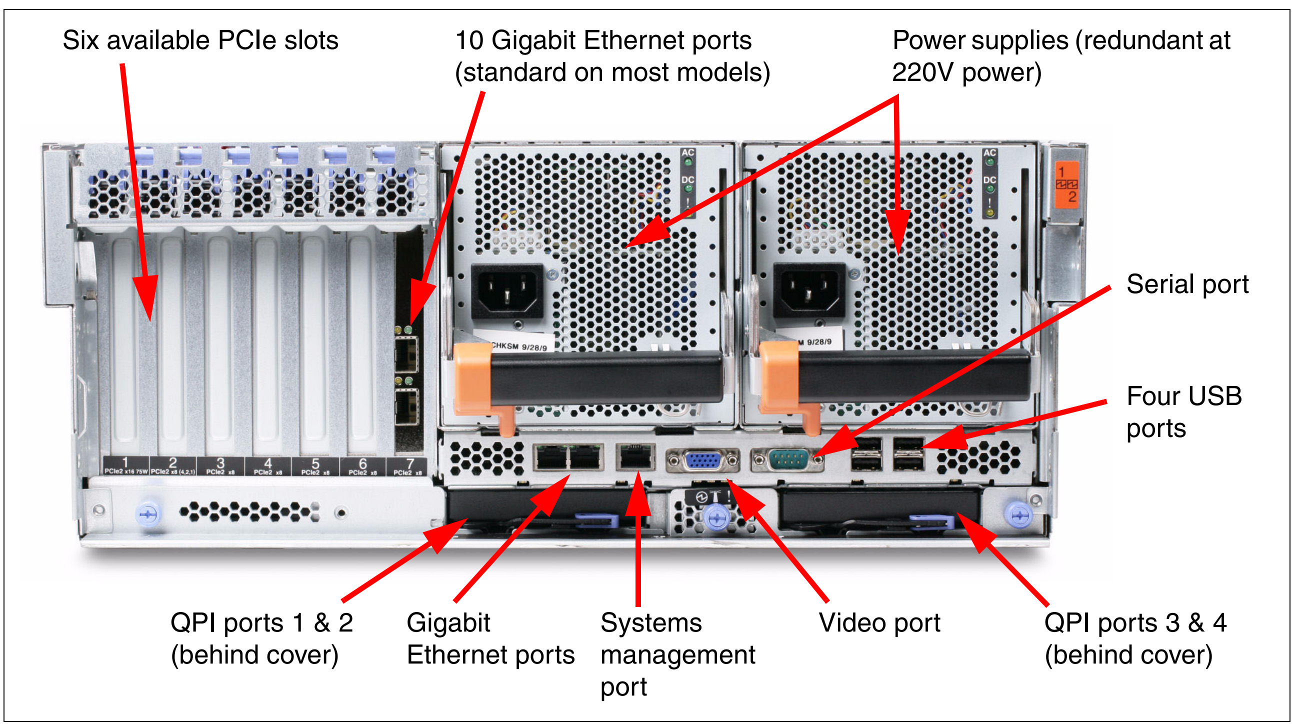


Figure 3-3 Rear of the x3850 X5

3.1.2 IBM System x3950 X5 product features

For certain enterprise workloads, IBM offers preconfigured models under the product name *x3950 X5*. These models do not differ from standard x3850 X5 models in terms of the machine type or the options used to configure them, but because they are configured with components that make them optimized for specific workloads, they are differentiated by this naming convention.

No model of x3850 X5 or x3950 X5 requires a scalability key for 8-socket operation (as was the case with the x3950 M2). Also, because the x3850 X5 and x3950 X5 use the same machine type, they can be scaled together into an 8-socket solution, assuming that each model uses four identical CPUs and that memory is set as a valid Hemisphere configuration. For more information about Hemisphere Mode, see 2.3.5, “Hemisphere Mode” on page 26.

The IBM x3950 X5 is optimized for database workloads and virtualization workloads. Virtualization-optimized models of the x3950 X5 include a MAX5 as standard. Database-optimized models include eXFlash as standard. See 3.3, “Models” on page 64 for more information.

3.1.3 IBM MAX5 memory expansion unit

The IBM MAX5 for System x (MAX5) memory expansion unit has 32 DDR3 dual inline memory module (DIMM) sockets, one or two 675-watt power supplies, and five 40 mm hot-swap speed-controlled fans. It provides added memory and multinode scaling support for the x3850 X5 server.

The MAX5 expansion module is based on eX5, the next generation of Enterprise X-Architecture. The MAX5 expansion module is designed for performance, expandability, and scalability. Its fans and power supplies use hot-swap technology for easy replacement without requiring the expansion module to be turned off.

Figure 3-4 shows the x3850 X5 with the attached MAX5.

Figure 3-4 x3850 X5 with the attached MAX5 memory expansion unit

The MAX5 has the following specifications:

- IBM EXA5 chip set.
- Intel memory controller with eight memory ports (four DIMMs on each port).
- Intel QuickPath Interconnect (QPI) architecture technology to connect the MAX5 to the x3850 X5. Four QPI links operate at up to 6.4 gigatransfers per second (GT/s).
- Scalability:
  - Connects to an x3850 X5 server using QPI cables.
- Memory DIMMs:
  - Minimum: 2 DIMMs, 4 GB.
  - Maximum: 32 DIMM connectors (up to 512 GB of memory using 16 GB DIMMs).
  - Type of DIMMs: PC3-10600, 1067 MHz, ECC, and DDR3 registered SDRAM DIMMs.
  - Supports 2 GB, 4 GB, 8 GB, and 16 GB DIMMs.

  All DIMM sockets in the MAX5 are accessible regardless of the number of processors installed on the host system.
- Five hot-swap 40 mm fans.
- Power supply:
  - Hot-swap power supplies with built-in fans for redundancy support.
  - 675-watt (110 - 220 V ac auto-sensing).
  - One power supply standard, two maximum (second power supply is for redundancy).
- Light path diagnostics LEDs:
  - Board LED
  - Configuration LED
  - Fan LEDs
  - Link LED (for QPI and EXA5 links)
  - Locate LED
  - Memory LEDs
  - Power-on LED
  - Power supply LEDs
- Physical specifications:
  - Width: 483 mm (19.0 in.)
  - Depth: 724 mm (28.5 in.)
  - Height: 44 mm (1.73 in.) (1U rack unit)
  - Basic configuration: 12.8 kg (28.2 lb.)
  - Maximum configuration: 15.4 kg (33.9 lb.)

With the addition of the MAX5 memory expansion unit, the x3850 X5 gains an additional 32 DIMM sockets for a total of 96 DIMM sockets. Using 16 GB DIMMs means that a total of 1.5 TB of RAM can be installed.

All DIMM sockets in the MAX5 are accessible, regardless of the number of processors installed on the host system.

Figure 3-5 on page 61 shows the ports at the rear of the MAX5 memory expansion unit. The QPI ports on the MAX5 are used to connect to a single x3850 X5. The EXA ports are reserved for future use.

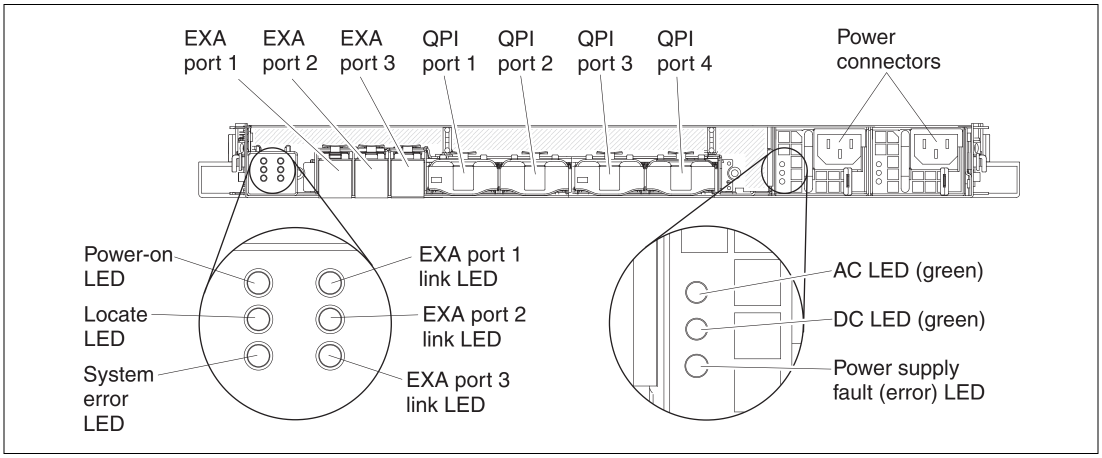


Figure 3-5 MAX5 connectors and LEDs

Figure 3-6 shows the internals of the MAX5 including the IBM EXA chip, which acts as the interface to the QPI links from the x3850 X5.

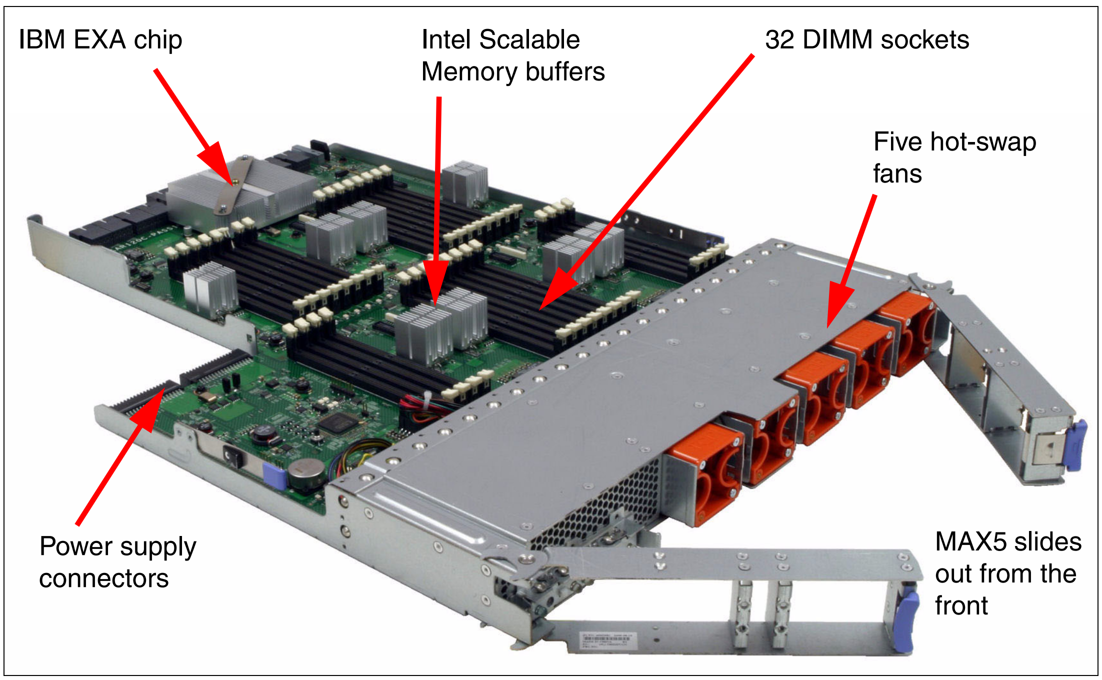


Figure 3-6 MAX5 memory expansion unit internals

For an in-depth look at the MAX5 offering, see 3.5, “MAX5” on page 68.

3.1.4 Comparing the x3850 X5 to the x3850 M2

Table 3-1 on page 62 shows a high-level comparison between the eX4-based x3850 M2 and the eX5-based x3850 X5.

Table 3-1 Comparison of the x3850 M2 to the x3850 X5

Subsystem	x3850 X5	x3850 M2
CPU card	► No Voltage Regulator Modules (VRMs), 4 Voltage Regulator Down (VRDs) ► Top access to CPUs and CPU card	► No Voltage Regulator Down (VRDs), 4 Voltage Regulator Modules (VRMs) ► Top access to CPU/VRM and CPU card
Memory	► Eight memory cards ► DDR3 PC3-10600 running at up to 1066 MHz (processor dependent) ► Eight DIMMs per memory card ► 64 DIMMs per chassis, maximum ► With the MAX5, 96 DIMMs per chassis	► Four memory cards ► DDR2 PC2-5300 running at 533 MHz ► Eight DIMMs per memory card ► 32 DIMMs per chassis, maximum
PCIe subsystem	► Intel 7500 “Boxboro” chip set ► All slots PCIe 2.0 ► Seven slots total at 5 Gb, 5 GHz, 500 MBps per lane ► Slot 1 PCIe x16, Slot2 x4 (x8 mechanical), Slots 3 - 7 x8 ► All slots non-hot-swap	► IBM CalIOC2 2.0 chip set ► All slots PCIe 1.1 ► Seven slots total at 2.5 GHz, 2.5 Gb, 250 MBps per lane ► Slot 1 x16, slot 2 x8 (x4), slots 3 - 7 x8 ► Slots 6 - 7 are hot-swap
SAS controller	► Standard ServeRAID BR10i with RAID 0 and 1 (most models) ► Optional ServeRAID M5015 with RAID 0, 1, and 5 ► Upgrade to RAID-6 and encryption ► No external SAS port	► LSI Logic 1078 with RAID-1 ► Upgrade key for RAID-5 ► SAS 4x external port for EXP3000 attach
Ethernet controller	► BCM 5709 dual-port Gigabit Ethernet, ► PCIe 2.0 x4 ► Dual-port Emulex 10Gb Ethernet adapter in PCIe slot 7 on all models except ARx	► BCM 5709 dual-port Gigabit Ethernet ► PCIe 1.1 x4
Video controller	► Matrox G200 in IMM ► 16 MB VRAM	► ATI RN50 on Remote Supervisor Adapter (RSA2) ► 16 MB VRAM
Service processor	► Maxim VSC452 Integrated BMC (IMM) ► Remote presence feature is standard	► RSA2 standard ► Remote presence feature is optional
Disk drive support	► Eight 2.5-inch internal drive bays or 16 1.8-inch solid-state drive bays ► Support for SATA and SSD	► Four 2.5-inch internal drive bays
USB, SuperIO design	► ICH10 chip set ► USB: Six external ports, two internal ► No SuperIO ► No PS/2 keyboard/mouse connectors ► No diskette drive controller ► Optional optical drive	► ICH7 chip set ► USB: Five external ports, one internal ► No SuperIO ► No PS/2 keyboard/mouse connectors ► No diskette drive controller
Fans	► 2x 120 mm ► 2x 60 mm ► 2x 120 mm in power supplies	► 4x 120 mm ► 2x 92 mm ► 2x 80 mm in power supplies
Power supply units	► 1975 W hot-swap, full redundancy high voltage, 875 W low voltage ► Rear access ► Two power supplies standard; two maximum (most models)[a]	► 1440 W hot-swap, full redundancy high voltage, 720 W low voltage ► Rear access ► Two power supplies standard; two maximum

a. Configuration restrictions at 110 V

3.2 Target workloads

This solution includes the following target workloads:

- Virtualization

  The following features address this workload:

  - Integrated USB key: All x3850 X5 models support the addition of an internal USB key that is preloaded with VMware ESXi 4.0 or ESXi 4.1 and that allows clients to set up and run a virtualized environment simply and quickly.
  - MAX5 expansion drawer: The average consolidated workload benefits from increased memory capacity per socket.

    As a general guideline, virtualization is a workload that is memory-intensive and I/O-intensive. A single-node x3850 X5 with MAX5 has a total of 96 available DIMM slots. The Intel 7500 series 8-core processors are an ideal choice for a VMware environment because the software is licensed by socket. The more cores per CPU, the more performance you get for the same single socket license.

    > **VMware ESXi support:** If you use a MAX5 unit, you must use VMware ESXi 4.1 or later. VMware ESXi 4.0 does not have support for MAX5.
    >
    > For more information, see the following website:
    >
    > `http://www.vmware.com/resources/compatibility/detail.php?device_cat=server&device_id=5317&release_id=144#notes`

  - Virtualization optimized models: One virtualization workload-optimized model of the x3950 X5 is announced. See 3.3, “Models” on page 64 for more information.
  - Processor support: The Intel 7500 series processors support VT FlexMigration Assist and VMware Enhanced VMotion.

- Database

  Database workloads require powerful CPUs and disk subsystems that are configured to deliver high I/O per second (IOPS) from sheer memory capacity (although the importance of sufficient low-latency, high-throughput memory must not be underestimated). IBM predefined database models use 8-core CPUs and use the power of eXFlash (high-IOPS SSDs). For more information about eXFlash, see 3.9.3, “IBM eXFlash and 1.8-inch SSD support” on page 93.

- Compute-intensive

  The x3850 X5 supports Windows HPC Server 2008 R2, an operating system designed for high-end applications that require high-performance computing (HPC) clusters. Features include a new high-speed NetworkDirect Remote Direct Memory Access (RDMA), highly efficient and scalable cluster management tools, a service-oriented architecture (SOA) job scheduler, and cluster interoperability through standards, such as the High Performance Computing Basic Profile (HPCBP) specification, which is produced by the Open Grid Forum (OGF).

For the workload-specific model details, see 3 3.3, “Models” on page 64.

3.3 Models

This section lists the currently available models. The x3850 X5 and x3950 X5 (both models are machine type 7145) have a three-year warranty.

For information about the recent models, consult tools, such as the Configurations and Options Guide (COG) or Standalone Solutions Configuration Tool (SSCT). These tools are available at the Configuration tools website:

http://www.ibm.com/systems/x/hardware/configtools.html

x3850 X5 base models without MAX5

Table 3-2 lists the base models of the x3850 X5 that do not include the MAX5 memory expansion unit as a standard. The MAX5 is optional. In the table, *std* is standard, *max* is maximum, and *C* is core (such as 4C is 4-core).

Table 3-2 Base models of the x3850 X5: Four-socket scalable server

Model[a]	Intel Xeon processors (two standard; maximum of four)	Memory speed (MHz)	Standard memory (MAX5 is optional)	Memory cards (std/max)	ServeRAID BR10i std	10Gb Ethernet standard[b]	Power supplies (std/max)	Drive bays (std)
7145-ARx	E7520 4C 1.86 GHz, 18 MB L3, 95W [c]	800 MHz	2x 2 GB	1/8	No	No	1/2	None
7145-1Rx	E7520 4C 1.86 GHz, 18 MB L3, 95W [c]	800 MHz	4x 4 GB	2/8	Yes	Yes	2/2	None
7145-2Rx	E7530 6C 1.86 GHz, 12 MB L3, 105W [c]	978 MHz	4x 4 GB	2/8	Yes	Yes	2/2	None
7145-3Rx	E7540 6C 2.0 GHz, 18 MB L3, 105W	1066 MHz	4x 4 GB	2/8	Yes	Yes	2/2	None
7145-4Rx	X7550 8C 2.0 GHz, 18 MB L3, 130W	1066 MHz	4x 4 GB	2/8	Yes	Yes	2/2	None
7145-5Rx	X7560 8C 2.26 GHz, 24 MB L3, 130W	1066 MHz	4x 4 GB	2/8	Yes	Yes	2/2	None

a. The x character in the seventh position of the machine model denotes the region-specific character. For example, U indicates US, and G indicates EMEA.
b. The Emulex 10Gb Ethernet Adapter is installed in PCIe slot 7.
c. Any model using the E7520 or E7530 CPU cannot scale beyond single node 4-way.

Workload-optimized x3950 X5 models

Table 3-3 on page 65 lists the workload-optimized models of the x3950 X5 that have been announced. The MAX5 is optional on these models. (In the table, *std* is standard, and *max* is maximum.)

Model 5Dx

Model 5Dx is designed for database applications and uses SSDs for the best I/O performance. Backplane connections for eight 1.8-inch SSDs are standard and there is space for an additional eight SSDs. You must order the SSDs separately. Because no SAS controllers are standard, you can select from the available cards that are described in 3.9, “Storage” on page 90.

Model 4Dx

Model 4Dx is designed for virtualization and is fully populated with 4 GB memory DIMMs, including in an attached MAX5 memory expansion unit, for a total of 384 GB of memory.

Backplane connections for four 2.5-inch SAS HDDs are standard; however, you must order the SAS HDDs separately. A ServeRAID BR10i SAS controller is standard in this model.

Table 3-3 Models of the x3950 X5: Workload-optimized models

Model[a]	Intel Xeon processors (two standard, maximum of four)	Memory speed	MAX5	Standard memory	Memory cards (std/max)	ServeRAID BR10i std	10Gb Ethernet standard[b]	Power supplies (std/max)	Drive bays (std/max)
Database workload-optimized models									
7145-5Dx	X7560 8C 2.27 GHz, 24 MB L3, 130W	1066 MHz	Opt	Server: 8x 4GB	4/8	No	Yes	2/2	None
Virtualization workload optimized models									
7145-4Dx	4x X7550 8C 2.0 GHz, 18 MB L3, 130W	1066 MHz	Std	Server: 64x 4GB MAX5: 32x 4GB	8/8	Yes	Yes	2/2	None

a. The x character in the seventh position of the machine model denotes the region-specific character. For example, U indicates US, and G indicates EMEA.
b. The Emulex 10Gb Ethernet Adapter is installed in PCIe slot 7.

x3850 X5 models with MAX5

Table 3-4 lists the models that are standard with the 1U MAX5 memory expansion unit.

Table 3-4 Models of the x3850 X5 with the MAX5 standard

Model[a]	Intel Xeon processors (four standard and max)	Memory speed (MHz)	Standard memory (MAX5 is standard)	Memory cards (std/max)	ServeRAID BR10i std	10Gb Ethernet standard[b]	Power supplies (std/max)	Drive bays (std/max)
7145-2Sx	4x E7530 6C 1.86 GHz, 12 MB L3, 105W [c]	978 MHz	Server: 8x 4 GB MAX5: 2x 4 GB	4/8	Yes	Yes	2/2	None
7145-4Sx	4x X7550 8C 2.0 GHz, 18 MB L3, 130W	1066 MHz	Server: 8x 4 GB MAX5: 2x 4 GB	4/8	Yes	Yes	2/2	None
7145-5Sx	4x X7560 8C 2.27 GHz, 24 MB L3, 130W	1066 MHz	Server: 8x 4 GB MAX5: 2x 4 GB	4/8	Yes	Yes	2/2	None

a. The x character in the seventh position of the machine model denotes the region-specific character. For example, U indicates US, and G indicates EMEA.
b. The Emulex 10Gb Ethernet Adapter is installed in PCIe slot 7.
c. Any model using the E7520 or E7530 CPU cannot scale beyond single node 4-way.

3.4 System architecture

This section explains the system board architecture and the use of the QPI wrap card.

3.4.1 System board

Figure 3-7 shows the system board layout of a single-node 4-way system.

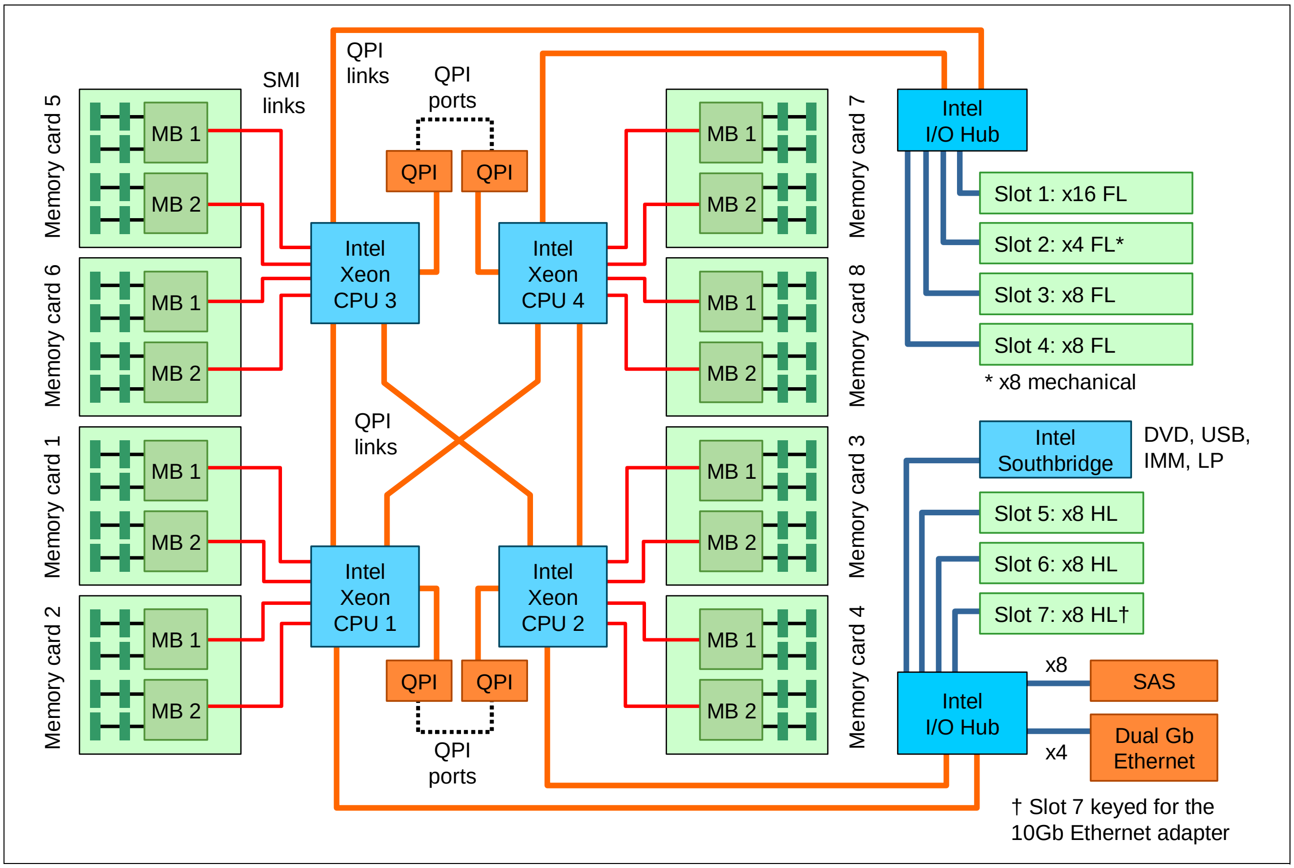


Figure 3-7 Block diagram for single-node x3850 X5

In Figure 3-7, the dotted lines indicate where the QPI Wrap Cards are installed in a 4-processor configuration. These wrap cards complete the full QPI mesh to allow all four processors to connect to each other. The QPI Wrap Cards are not needed in 2-processor configurations and are removed when a MAX5 is connected.

Figure 3-12 on page 70 is a block diagram of the x3850 X5 connected to a MAX5.

3.4.2 QPI Wrap Card

In the x3850 X5, QPI links are used for interprocessor communication, both in a single-node system and in a 2-node system. They are also used to connect the system to a MAX5 memory expansion drawer. In a single-node x3850 X5, the QPI links connect in a full mesh between all CPUs. To complete this mesh, the QPI Wrap Card is used.

Tip: The QPI Wrap Cards are only for single-node configurations with three or four processors installed. They are *not necessary* for any of the following items:

- Single-node configurations with two processors
- Configurations with MAX5 memory expansion units
- Two-node configurations

Figure 3-8 shows the QPI Wrap Card.

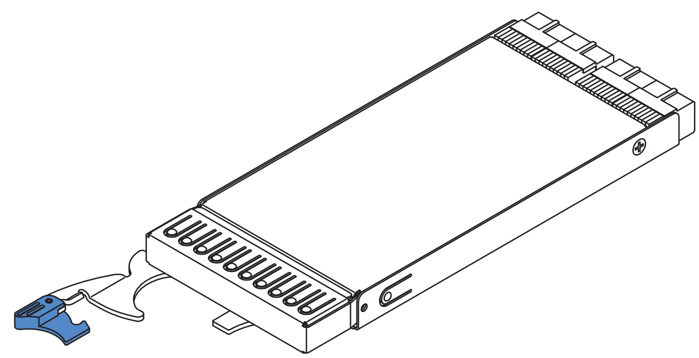


Figure 3-8 QPI Wrap Card

For single-node systems with three or four processors installed, but without the MAX5 memory expansion unit connected, install two QPI Wrap Cards. Figure 3-9 shows a diagram of how the QPI Wrap Cards are used to complete the QPI mesh. Although the QPI Wrap Cards are not mandatory, they provide a performance boost by ensuring that all CPUs are only one *hop* away from each other, as shown in Figure 3-9.

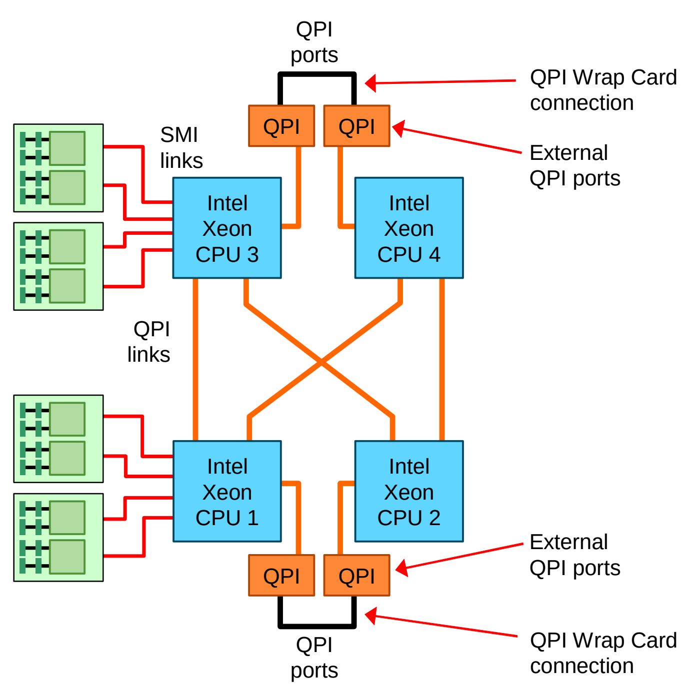


Figure 3-9 Location of QPI Wrap Cards

The QPI Wrap Cards are not included with standard server models and must be ordered separately. See Table 3-5.

Table 3-5 Ordering information for the QPI Wrap Card

Part number	Feature code	Description
49Y4379	Not applicable	IBM x3850 X5 and x3950 X5 QPI Wrap Card Kit (quantity 2)

Tips:

- Part number 49Y4379 includes two QPI Wrap Cards. You order only one of these parts per server.
- QPI Wrap Cards cannot be ordered individually.

The QPI Wrap Cards are installed in the QPI bays at the back of the server, as shown in Figure 3-10.

QPI Wrap Cards are not needed in a 2-node configuration and not needed in a MAX5 configuration. When the QPI Wrap Cards are installed, no external QPI ports are available. If you later want to attach a MAX5 expansion unit or connect a second node, you must first remove the QPI Wrap Cards.

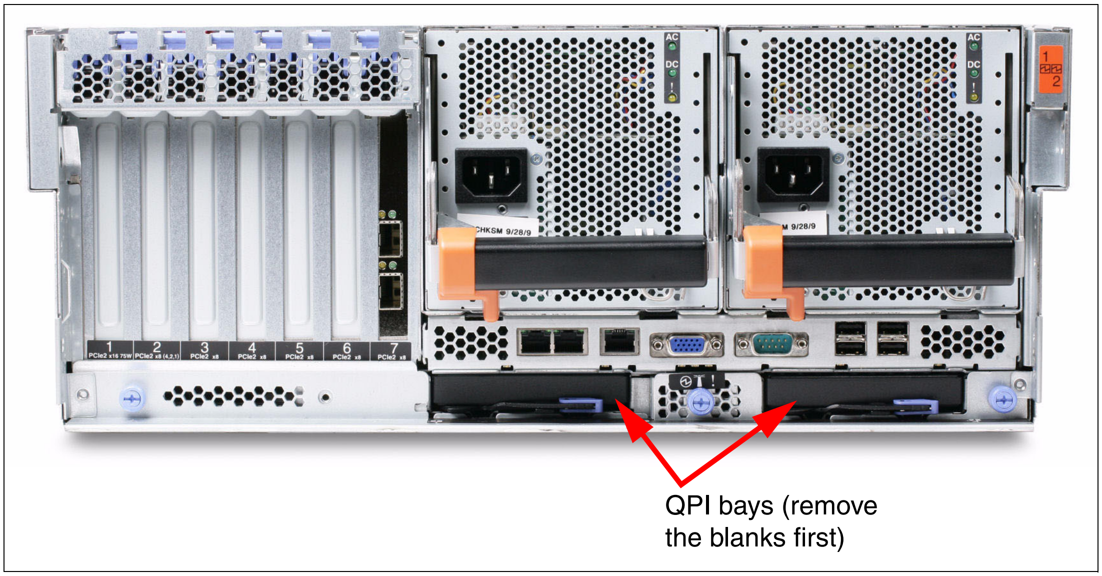


Figure 3-10 Rear of the x3850 X5

3.5 MAX5

As introduced in 3.1.3, “IBM MAX5 memory expansion unit” on page 59, the MAX5 memory expansion drawer is available for both the x3850 X5 and the x3950 X5. Models of the x3850 X5 and x3950 X5 are available that include the MAX5, as described in 3.3, “Models” on page 64, Also, you can order the MAX5 separately, as listed in Table 3-6. When ordering a MAX5, remember to order the cable kit as well. For power supply fault redundancy, order the optional power supply.

Table 3-6 Ordering information for the IBM MAX5 for System x

Part number	Feature code	Description
59Y6265	4199	IBM MAX5 for System x
60Y0332	4782	IBM 675W HE Redundant Power Supply
59Y6267	4192	IBM MAX5 to x3850 X5 Cable Kit

The eX5 chip set in the MAX5 is an IBM unique design that attaches to the QPI links as a node controller, giving it direct access to all CPU bus transactions. It increases the number of DIMMs supported in a system by a total of 32, and it also adds another 16 channels of memory bandwidth, boosting overall throughput. Therefore, the MAX5 adds additional memory and performance.

The eX5 chip connects directly through QPI links to all of the CPUs in the x3850 X5, and it maintains a directory of each CPU's last-level cache. Therefore, when a CPU requests content stored in the cache of another CPU, the MAX5 not only has that same data stored in its own cache, it is able to return the acknowledgement of the snoop *and* the data to the requesting CPU in the same transaction. For more information about QPI links and snooping, see 2.2.4, "QuickPath Interconnect (QPI)" on page 18.

The MAX5 also has EXA scalability ports used in an EXA-scaled configuration (that is, a 2-node and MAX5 configuration). These ports are reserved for future use.

In summary, the MAX5 offers the following major features:

- Adds 32 DIMM slots to either the x3850 X5 or the x3690 X5
- Adds 16 channels of memory bandwidth
- Improves snoop latencies

Figure 3-11 shows a diagram of the MAX5.

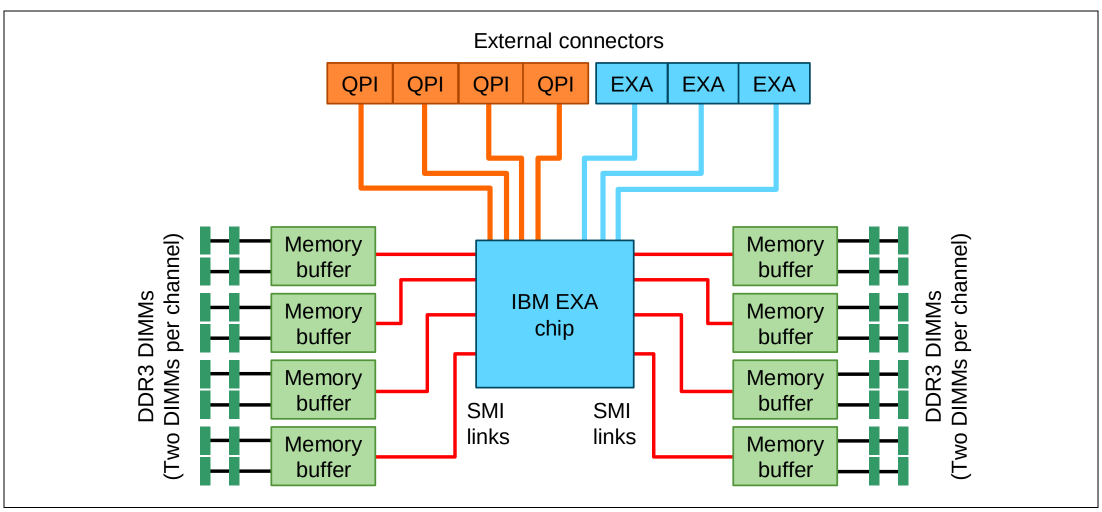


Figure 3-11 MAX5 block diagram

The MAX5 is connected to the x3850 X5 using four cables, connecting the QPI ports on the server to the four QPI ports on the MAX5. Figure 3-12 shows architecturally how a single-node x3850 X5 connects to a MAX5.

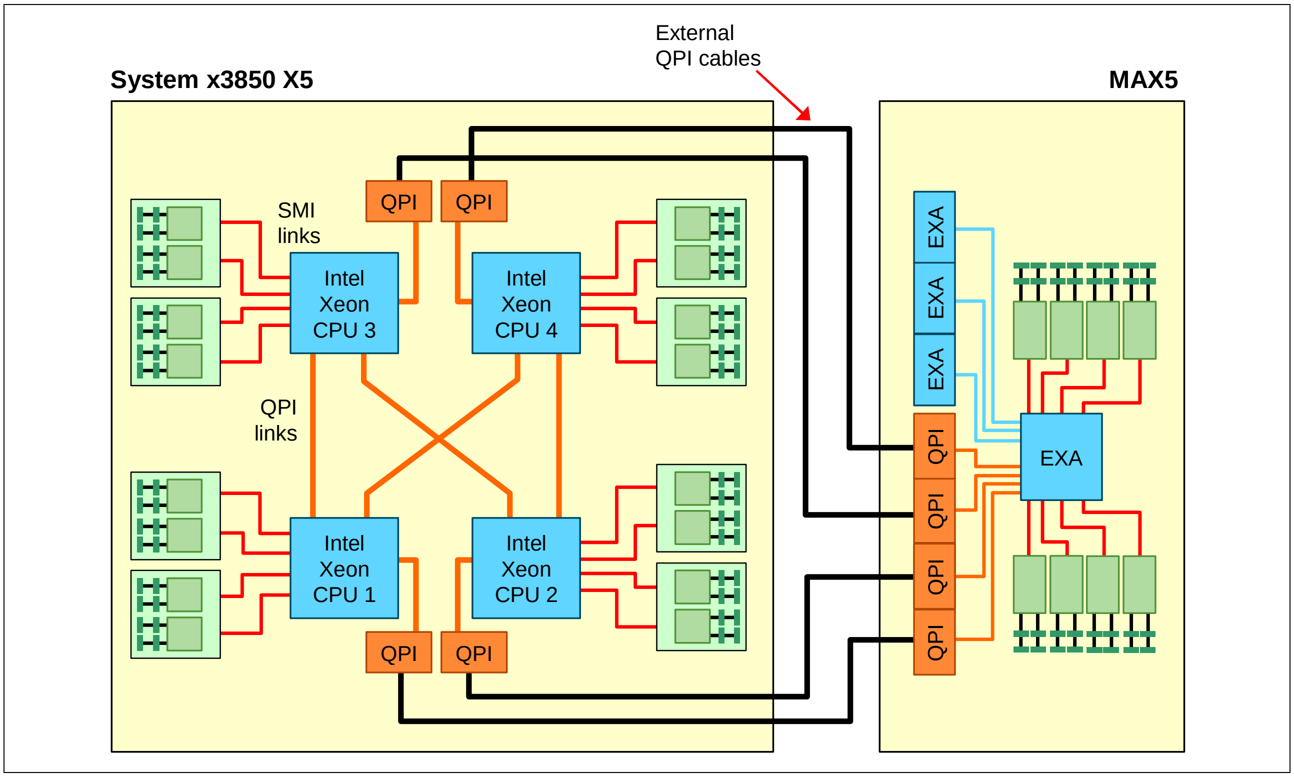


Figure 3-12 The x3850 X5: Connectivity of the system unit with the MAX5

Tip: As shown in Figure 3-12 on page 70, you maximize performance when you have four processors installed, because you then have four active QPI links to the MAX5. However, configurations of two and three processors are still supported. If only two processors are required, consider the use of the x3690 X5.

We describe the connectivity of the MAX5 to the x3850 X5 in 3.6, “Scalability” on page 70.

For memory configuration information, see 3.8.4, “Memory mirroring” on page 87. For information about power and fans, see 3.14, “Power supplies and fans of the x3850 X5 and MAX5” on page 112.

3.6 Scalability

In this section, we describe how to expand the x3850 X5 to increase the number of processors and the number of memory DIMMs.

The x3850 X5 currently supports the following scalable configurations:

- A single x3850 X5 server with four processor sockets. This configuration is sometimes referred to as a *single-node server*.
- A single x3850 X5 server with a single MAX5 memory expansion unit attached. This configuration is sometimes referred to as a *memory-expanded server*.

- Two x3850 X5 servers connected to form a single image 8-socket server. This configuration is sometimes referred to as a *2-node server*.

MAX5: The configuration of two nodes with MAX5 is not supported.

3.6.1 Memory scalability with MAX5

The MAX5 memory expansion unit permits the x3850 X5 to scale to an additional 32 DDR3 DIMM sockets.

Connecting the single-node x3850 X5 with the MAX5 memory expansion unit uses four QPI cables, part number 59Y6267, as listed in Table 3-7. Figure 3-13 shows the connectivity.

Tip: As shown in Figure 3-12 on page 70, you maximize performance when you have four processors installed because you have four active QPI links to the MAX5. However, configurations of two and three processors are still supported.

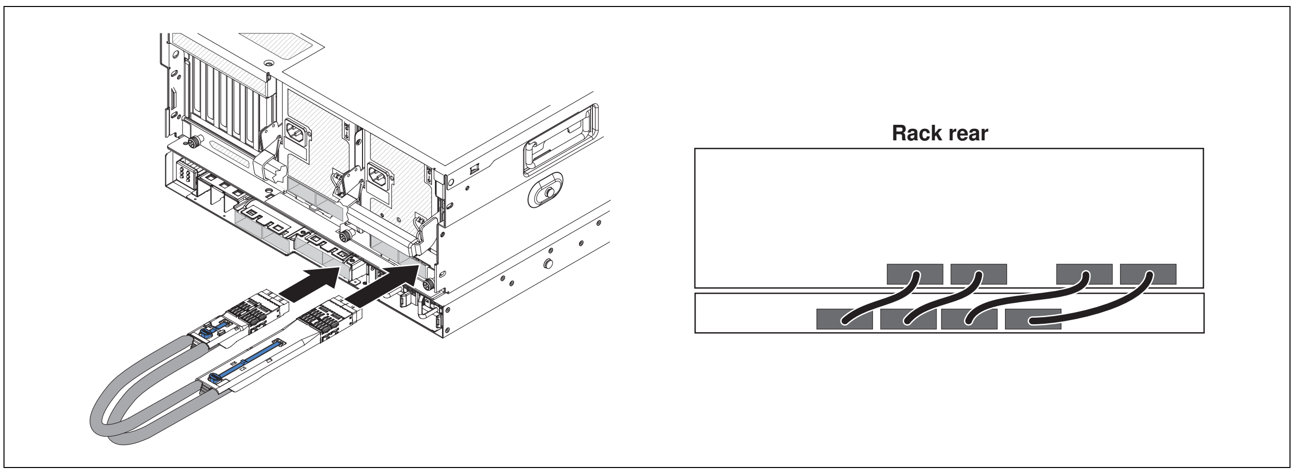


Figure 3-13 Connecting the MAX5 to a single-node x3850 X5

Connecting the MAX5 to a single-node x3850 X5 requires one IBM MAX5 to x3850 X5 Cable Kit, which consists of four QPI cables. See Table 3-7.

Table 3-7 Ordering information for the IBM MAX5 to x3850 X5 Cable Kit

Part number	Feature code	Description
59Y6267	4192	IBM MAX5 to x3850 X5 Cable Kit (quantity 4 cables)

3.6.2 Two-node scalability

The 2-node configuration also uses native Intel QPI scaling to create an 8-socket configuration. The two servers are physically connected to each other with a set of external QPI cables. The cables are connected to the server through the QPI bays, which are shown in Figure 3-7 on page 66. Figure 3-14 on page 72 shows the cable routing.

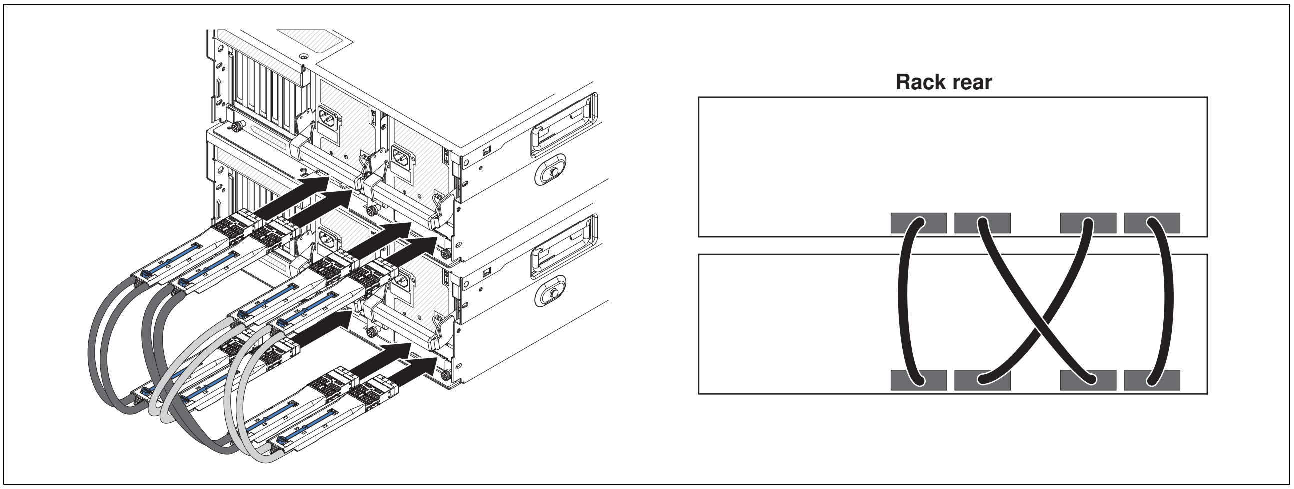


Figure 3-14 Cabling diagram for two node x3850 X5

Connecting the two x3850 X5 servers to form a 2-node system requires one IBM x3850 X5 and x3950 X5 QPI Scalability Kit, which consists of four QPI cables. See Table 3-8.

Table 3-8 Ordering information for the IBM x3850 X5 and x3950 X5 QPI Scalability Kit

Part number	Feature code	Description
46M0072	5103	IBM x3850 X5 and x3950 X5 QPI Scalability Kit (quantity 4 cables)

No QPI ports are visible on the rear of the server. The QPI scalability cables have long rigid connectors, allowing them to be inserted into the QPI bay until they connect to the QPI ports, which are located a few inches inside on the planar. Completing the QPI scaling of two x3850 X5 servers into a 2-node complex does not require any other option.

Intel E7520 and E7530: The Intel E7520 and E7530 processors cannot be used to scale to an 8-way 2-node complex. They support a maximum of four processors. At the time of this writing, the following models use those processors:

- 7145-ARx
- 7145-1Rx
- 7145-2Rx
- 7145-2Sx

Figure 3-15 on page 73 shows the QPI links that are used to connect two x3850 X5 servers to each other. Both nodes must have four processors each, and all processors must be identical.

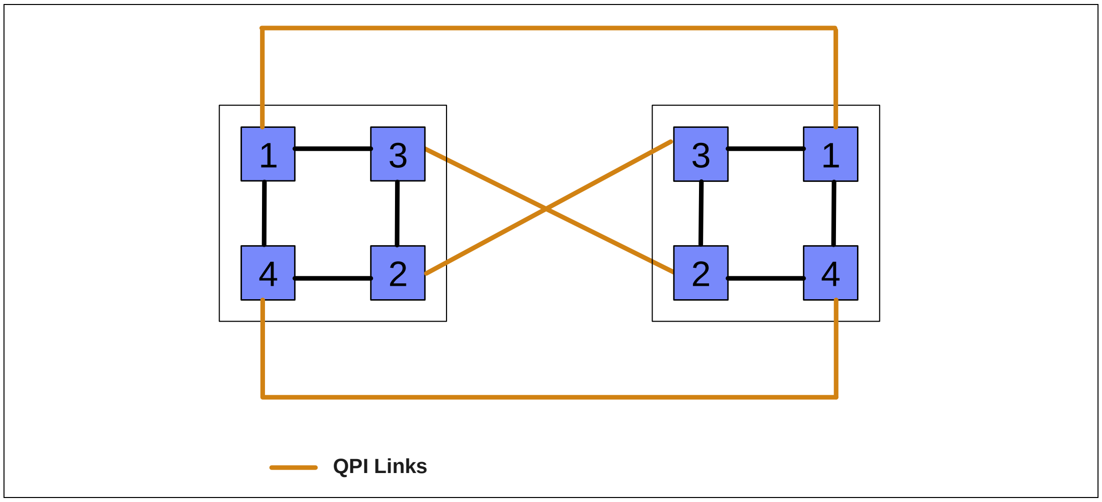


Figure 3-15 QPI links for a 2-node x3850 X5

QPI-based scaling is managed primarily through the Unified Extensible Firmware Interface (UEFI) firmware of the x3850 X5.

For the 2-node x3850 X5 scaled through the QPI ports, when those cables are connected, the two nodes act as one system until the cables are physically disconnected.

Firmware levels: It is important to ensure that both of the x3850 X5 servers have the identical UEFI, integrated management module (IMM), and Field-Programmable Gate Array (FPGA) levels before scaling. If they are not at the same levels, unexpected issues occur and the server might not boot. See 9.10, “Firmware update tools and methods” on page 509 for ways to check and update the firmware.

Partitioning: The x3850 X5 currently does not support partitioning.

3.7 Processor options

The x3850 X5 is supported with two, three, or four processors. Table 3-9 shows the option part numbers for the supported processors. In a 2-node system, you must have eight processors, which must all be identical. For a list of the processor options available in this solution, see 2.2, “Intel Xeon 6500 and 7500 family processors” on page 16.

Table 3-9 Available processor options for the x3850 X5

Part number	Feature code	Intel Xeon model	Speed	Cores	L3 cache	GT/s/ Memory speed[a]	Power (watts)	HT[b]	TB[c]
49Y4300	4513	X7560	2.26 GHz	8	24 MB	x6.4/1066 MHz	130 W	Yes	Yes
49Y4302	4517	X7550	2.00 GHz	8	18 MB	x6.4/1066 MHz	130 W	Yes	Yes
59Y6103	4527	X7542	2.66 GHz	6	18 MB	x5.86/978 MHz	130 W	No	Yes
49Y4304	4521	E7540	2.00 GHz	6	18 MB	x6.4/1066 MHz	105 W	Yes	Yes
49Y4305	4523	E7530[d]	1.86 GHz	6	12 MB	x5.86/978 MHz	105 W	Yes	Yes
49Y4306	4525	E7520[d]	1.86 GHz	4	18 MB	x4.8/800 MHz	95 W	Yes	No
49Y4301	4515	L7555	1.86 GHz	8	24 MB	x5.86/978 MHz	95 W	Yes	Yes
49Y4303	4519	L7545	1.86 GHz	6	18 MB	x5.86/978 MHz	95 W	Yes	Yes

a. GT/s is gigatransfers per second. For an explanation, see 2.3.1, “Memory speed” on page 22.
b. Intel Hyper-Threading Technology. For an explanation, see 2.2.2, “Hyper-Threading Technology” on page 17.
c. Intel Turbo Boost Technology. For an explanation, see 2.2.3, “Turbo Boost Technology” on page 18.
d. Scalable to a 4-socket maximum, and therefore, it cannot be used in a 2-node x3850 X5 complex that is scaled with native QPI cables.

With the exception of the E7520, all processors listed in Table 3-9 support Intel Turbo Boost Technology. When a processor operates below its thermal and electrical limits, Turbo Boost *dynamically* increases the clock frequency of the processor by 133 MHz on short and regular intervals until an upper limit is reached. See 2.2.3, “Turbo Boost Technology” on page 18 for more information.

With the exception of the X7542, all of the processors that are shown in Table 3-9 support Intel Hyper-Threading Technology, which is an Intel technology that can improve the parallelization of workloads. When Hyper-Threading is engaged in the BIOS, for each processor core that is physically present, the operating system addresses two. For more information, see 2.2.2, “Hyper-Threading Technology” on page 17.

All processor options include the heat-sink and CPU installation tool. This tool is extremely important due to the high possibility of bending pins on the processor socket when using the incorrect procedure.

The x3850 X5 includes at least two CPUs as standard. Two CPUs are required to access all seven of the PCIe slots (shown in Figure 3-7 on page 66):

- Either CPU 1 or CPU 2 is required for the operation of PCIe slots 5-7.
- Either CPU 3 or CPU 4 is required for the operation of PCIe Slots 1-4.

All CPUs are also required to access all memory cards on the x3850 X5 but they are not required to access memory on the MAX5, as explained in 3.8, “Memory” on page 76.

Use these population guidelines:

- Each CPU requires a minimum of two DIMMs to operate.
- All processors must be identical.
- Only configurations of two, three, or four processors are supported.
- The number of installed processors dictates what memory cards can be used:
  - Two installed processors enable four memory cards.
  - Three installed processors enable six memory cards.
  - Four installed processors enable all eight memory cards.
- A processor must be installed in socket 1 or 2 for the system to successfully boot.
- A processor is required in socket 3 or 4 to use PCIe slots 1 - 4. See Figure 3-7 on page 66.
- When installing three or four processors, use a QPI Wrap Card Kit (part number 49Y4379) to improve performance. The kit contains two wrap cards. See 3.4.2, “QPI Wrap Card” on page 66.
- When using a MAX5 memory expansion unit, as shown in Figure 3-12 on page 70, you maximize performance when you have four installed processors because there are four active QPI links to the MAX5. However, configurations of two and three processors are still supported.
- Consider the X7542 processor for CPU frequency-dependent workloads because it has the highest core frequency of the available processor models.
- If high processing capacity is not required for your application but high memory bandwidth is required, consider using four processors with fewer cores or a lower core frequency rather than two processors with more cores or a higher core frequency. Having four processors enables all memory channels and maximizes memory bandwidth. We describe this situation in 3.8, “Memory” on page 76.

3.8 Memory

Memory is installed in the x3850 X5 in memory cards. Up to eight memory cards can be installed in the server, and each card holds eight DIMMs. Therefore, the x3850 X5 supports up to 64 DIMMs.

This section includes the following topics:

- 3.8.1, “Memory cards and DIMMs” on page 76
- 3.8.2, “DIMM population sequence” on page 79
- 3.8.3, “Maximizing memory performance” on page 84
- 3.8.4, “Memory mirroring” on page 87
- 3.8.5, “Memory sparing” on page 89
- 3.8.6, “Effect on performance by using mirroring or sparing” on page 89

3.8.1 Memory cards and DIMMs

This section describes the available memory options for the x3850 X5 and the MAX5.

Memory cards for the x3850 X5

The x3850 X5, like its predecessor the x3850 M2, uses memory cards to which the memory DIMMs are attached, as shown in Figure 3-16.

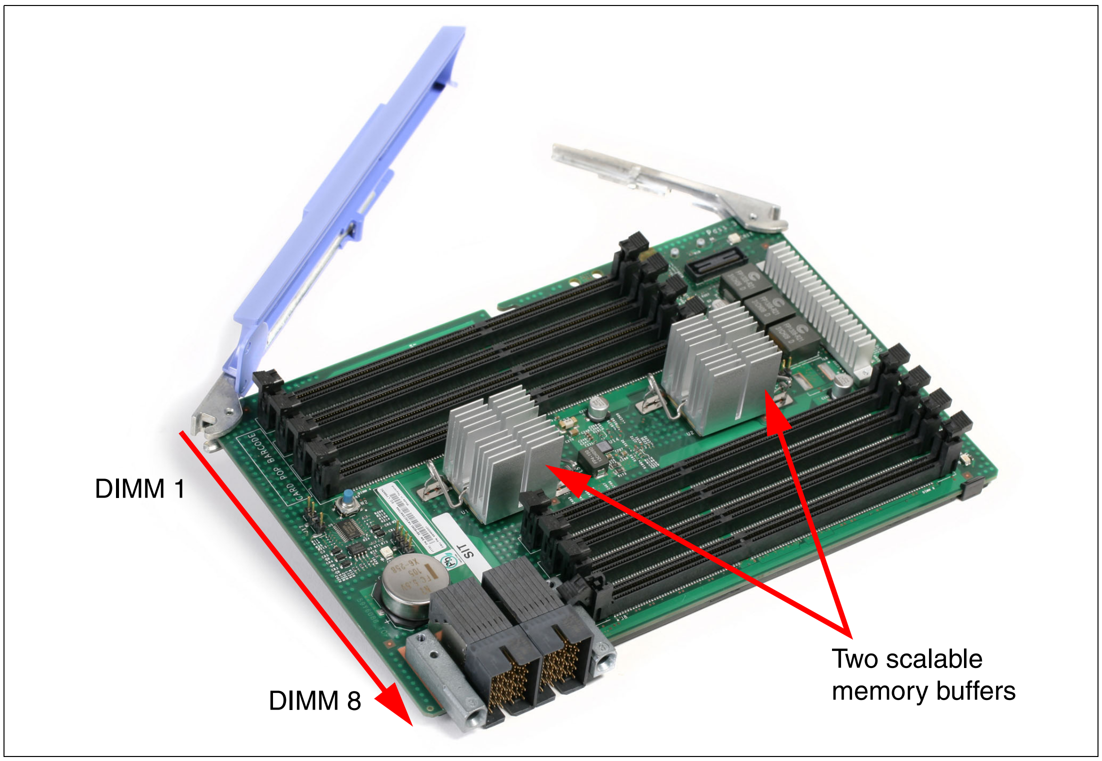


Figure 3-16 x3850 x5 memory card

Standard models contain two or more memory cards. You can configure additional cards, as listed in Table 3-10.

Table 3-10 IBM System x3850 X5 and x3950 X5 memory card

Part number	Feature code	Description
46M0071	5102	IBM x3850 X5 and x3950 X5 Memory Expansion Card

The memory cards are installed in the server, as shown in Figure 3-17. Each processor is electrically connected to two memory cards as shown (for example, processor 1 is connected to memory cards 1 and 2).

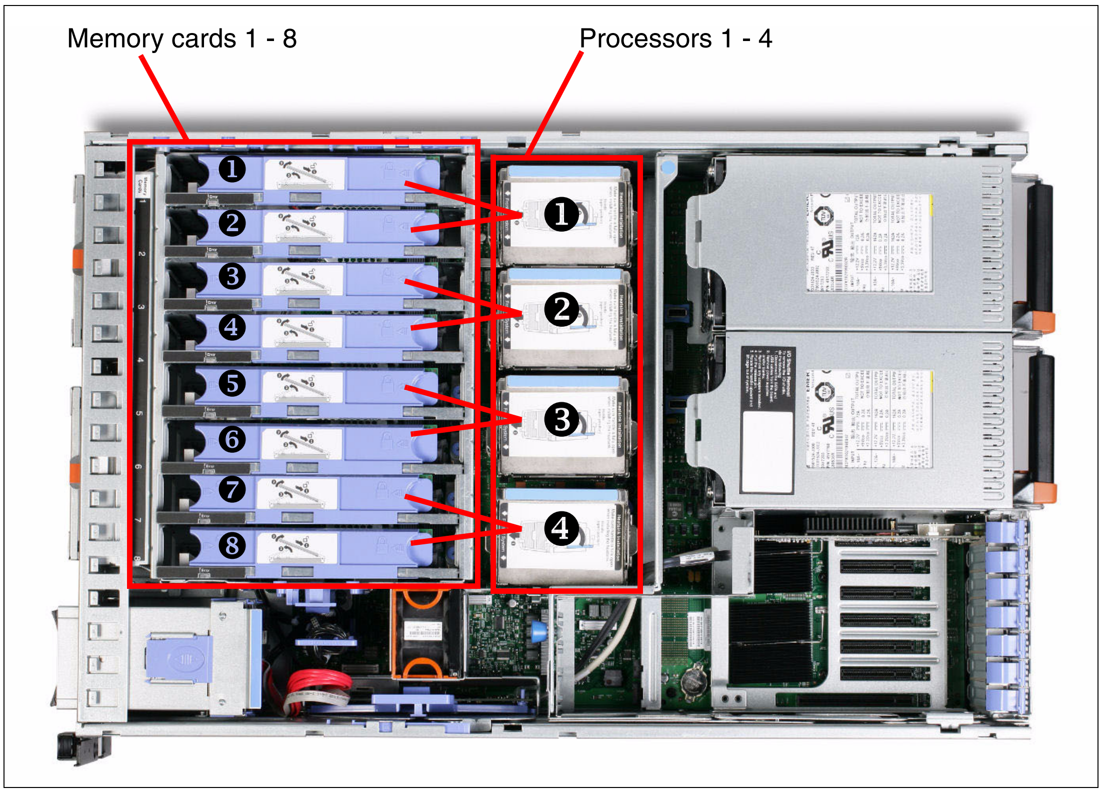


Figure 3-17 Memory card and processor enumeration

DIMMs for the x3850 X5

Table 3-11 shows the available DIMMs that are supported in the x3850 X5 server. The table also indicates the DIMM options that are also supported in the MAX5. When used in the MAX5, the DIMMs have separate feature codes, which are shown as - *fc*.

Table 3-11 x3850 X5 supported DIMMs

Part number	x3850 X5 feature code	Memory	Supported in MAX5	Memory speed[a]	Ranks
44T1592	1712	2 GB (1x 2GB) 1Rx8, 2 Gb PC3-10600R DDR3-1333	Yes (fc 2429)	1333 MHz[b]	Single x8
44T1599	1713	4 GB (1x 4GB), 2Rx8, 2 Gb PC3-10600R DDR3-1333	Yes (fc 2431)	1333 MHz	Dual x8
46C7448	1701	4 GB (1x 4GB), 4Rx8, 1 Gb PC3-8500 DDR3-1066	No	1066 MHz	Quad x8
46C7482	1706	8 GB (1x 8GB), 4Rx8, 2 Gb PC3-8500 DDR3-1066	Yes (fc 2432)	1066 MHz	Quad x8
46C7483	1707	16 GB (1x 16GB), 4Rx4, 2 Gb PC3-8500 DDR3-1066	Yes[c] (fc 2433)	1066 MHz	Quad x4

a. Memory speed is also controlled by the memory bus speed as specified by the processor model selected. The actual memory bus speed is the lower of both the processor memory bus speed and the DIMM memory bus speed.
b. Although 1333 MHz memory DIMMs are supported in the x3850 X5, the memory DIMMs run at a maximum speed of 1066 MHz.
c. The 16 GB memory option is supported in the MAX5 only when it is the only type of memory that is used in the MAX5. No other memory options can be used in the MAX5 if this option is installed in the MAX5. This DIMM also supports redundant bit steering (RBS) when used in the MAX5, as described in “Redundant bit steering” on page 29.

Guidelines:

- Memory options must be installed in matched pairs. Single options cannot be installed, so the options that are shown in Table 3-11 need to be ordered in quantities of two.
- You can achieve additional performance by enabling Hemisphere Mode, which is described in “Hemisphere Mode” on page 26. This mode requires that the memory options are installed in matched quads.
- The maximum memory speed that is supported by Xeon 7500 and 6500 (Nehalem-EX) processors is 1066 MHz (1,333 MHz speed is not supported). Although the 1333 MHz DIMMs are still supported in the x3850 X5, they can operate at a speed of at most 1066 MHz.

As with Intel Xeon 5500 processor (Nehalem-EP), the speed at which the memory that is connected to the Xeon 7500 and 6500 processors (Nehalem-EX) runs depends on the capabilities of the specific processor. With Nehalem-EX, the scalable memory interconnect (SMI) link runs from the memory controller that is integrated in the processor to the memory buffers on the memory cards.

The SMI link speed is derived from the processor QPI link speed:

- 6.4 GT/s QPI link speed capable of running memory speeds up to 1066 MHz
- 5.86 GT/s QPI link speed capable of running memory speeds up to 978 MHz
- 4.8 GT/s QPI link speed capable of running memory speeds up to 800 MHz

To see more information about how memory speed is calculated with QPI, see 2.3.1, “Memory speed” on page 22.

MAX5 memory

The MAX5 memory expansion unit has 32 DIMM sockets and is designed to augment the installed memory in the attached x3850 X5 server. Table 3-12 shows the available memory options that are supported in the MAX5 memory expansion unit. These options are a subset of the options that are supported in the x3850 X5 because the MAX5 requires that all DIMMs use identical DRAM technology: either 2 Gb x8 or 2 Gb x4 (but not both at the same time).

x3850 X5 memory options: The memory options listed here are also supported in the x3850 X5, but under other feature codes for configure-to-order (CTO) clients. Additional memory options are also supported in the x3850 X5 server but not in the MAX5; these options are listed in Table 3-11 on page 78.

Table 3-12 DIMMs supported in the MAX5

Part number	MAX5 feature code	Memory	Supported in MAX5	Memory speed[a]	Ranks
44T1592	2429	2 GB (1x 2GB) 1Rx8, 2Gbit PC3-10600R DDR3-1333	Yes	1333 MHz[b]	Single x8
44T1599	2431	4 GB (1x 4GB), 2Rx8, 2Gbit PC3-10600R DDR3-1333	Yes	1333 MHz	Dual x8
46C7482	2432	8 GB (1x 8GB), 4Rx8, 2Gbit PC3-8500 DDR3-1066	Yes	1066 MHz	Quad x8
46C7483	2433	16 GB (1x 16GB), 4Rx4, 2Gbit PC3-8500 DDR3-1066	Yes[c d]	1066 MHz	Quad x4

a. Memory speed is also controlled by the memory bus speed, as specified by the selected processor model. The actual memory bus speed is the lower of both the processor memory bus speed and the DIMM memory bus speed.
b. Although 1333 MHz memory DIMMs are supported in the x3690 X5, the memory DIMMs run at a maximum speed of 1066 MHz.
c. The 16 GB memory option is supported in the MAX5 only when it is the only type of memory used in the MAX5. No other memory options can be used in the MAX5 if this option is installed in the MAX5.
d. This DIMM supports redundant bit steering (RBS), as described in “Redundant bit steering” on page 29.

Use of the 16 GB memory option: The 16 GB memory option, 46C7483, is supported in the MAX5 only when it is the only type of memory that is used in the MAX5. No other memory options can be used in the MAX5 if this option is installed in the MAX5.

Redundant bit steering: Redundant bit steering (RBS) is not supported on the x3850 X5 itself, because the integrated memory controller of the Intel Xeon 7500 processors does not support the feature. See “Redundant bit steering” on page 29 for details.

The MAX5 memory expansion unit support RBS, but only with x4 memory and not x8 memory. As shown in Table 3-12, the 16 GB DIMM, part 46C7483, uses x4 DRAM technology. RBS is automatically enabled in the MAX5 memory port, if all DIMMs installed to that memory port are x4 DIMMs.

3.8.2 DIMM population sequence

This section describes the order in which to install the memory DIMMs in the x3850 X5 and MAX5.

Installing DIMMs in the x3850 X5 and MAX5 in the correct order is essential for system performance. See “Mixed DIMMs and the effect on performance” on page 86 for performance effects when this guideline is not followed.

> **Tip:** The tables in this section list only memory configurations that are considered the best practices in obtaining the optimal memory and processor performance.
>
> For a full list of supported memory configurations, see the *IBM System x3850 X5 Installation and User’s Guide* or the IBM *System x3850 X5 Problem Determination and Service Guide*. We list the download links to these documents in “Related publications” on page 541.

x3850 X5 single-node and 2-node configurations

Table 3-13 is the same if you use a single-node configuration or if you use a 2-node configuration. In a 2-node configuration, you install in the same order twice, once for each server.

Table 3-13 shows the NUMA-compliant memory installation sequence for two processors.

Table 3-13 NUMA-compliant DIMM installation (two processors): x3850 X5

Number of DIMMs	**Hemisphere Mode**[a]	**Processor 1**								**Processor 4**							
		Card 1				**Card 2**				**Card 7**				**Card 8**			
		DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5
4	N	x								x							
8	Y	x				x				x				x			
12	N	x	x			x				x	x			x			
16	Y	x	x			x	x			x	x			x	x		
20	N	x	x	x		x	x			x	x	x		x	x		
24	Y	x	x	x		x	x	x		x	x	x		x	x	x	
28	N	x	x	x	x	x	x	x		x	x	x	x	x	x	x	
32	Y	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

a. For more information about Hemisphere Mode and its importance, see 2.3.5, “Hemisphere Mode” on page 26.

Table 3-14 on page 81 shows the NUMA-compliant memory installation sequence for three processors.

Table 3-14 NUMA-compliant DIMM installation (three processors): x3850 X5

Number of DIMMs	Hemisphere Mode[a]	Processor 1								Processor 4								Processor 2 or 3							
		Card 1				Card 2				Card 7				Card 8				Card 3				Card 4			
		DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5
6	N	x								x								x							
12	Y	x				x				x				x				x				x			
18	N	x	x			x				x	x			x				x	x			x			
24	Y	x	x			x	x			x	x			x	x			x	x			x	x		
30	N	x	x	x		x	x			x	x	x		x	x			x	x	x		x	x		
36	Y	x	x	x		x	x	x		x	x	x		x	x	x		x	x	x		x	x	x	
42	N	x	x	x	x	x	x	x		x	x	x	x	x	x	x		x	x	x	x	x	x	x	
48	Y	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

a. For more information about Hemisphere Mode and its importance, see 2.3.5, "Hemisphere Mode" on page 26.

Three-processor system: For a 3-processor system, you can use either processor slot 2 or processor 3. Processor 3 uses cards 5 and 6 instead of cards 3 and 4, which are used for processor 2.

Table 3-15 shows the NUMA-compliant memory installation sequence for four processors.

Table 3-15 NUMA-compliant DIMM installation (four processors): x3850 X5

Number of DIMMs	Hemisphere Mode[a]	Processor 1								Processor 4								Processor 2								Processor 3							
		Card 1				Card 2				Card 7				Card 8				Card 3				Card 4				Card 5				Card 6			
		DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5
8	N	x								x								x								x							
16	Y	x				x				x				x				x				x				x				x			
24	N	x	x			x				x	x			x				x	x			x				x	x			x			
32	Y	x	x			x	x			x	x			x	x			x	x			x	x			x	x			x	x		
40	N	x	x	x		x	x			x	x	x		x	x			x	x	x		x	x			x	x	x		x	x		
48	Y	x	x	x		x	x	x		x	x	x		x	x	x		x	x	x		x	x	x		x	x	x		x	x	x	
56	N	x	x	x	x	x	x	x		x	x	x	x	x	x	x		x	x	x	x	x	x	x		x	x	x	x	x	x	x	

Number of DIMMs	Hemisphere Mode[a]	Processor 1								Processor 4								Processor 2								Processor 3							
		Card 1				Card 2				Card 7				Card 8				Card 3				Card 4				Card 5				Card 6			
		DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5
64	Y	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

a. For more information about Hemisphere Mode and its importance, see 2.3.5, “Hemisphere Mode” on page 26.

MAX5 configurations

The memory installed in the MAX5 operates at the same speed as the memory that is installed in the x3850 X5 server. As explained in 2.3.1, “Memory speed” on page 22, the memory speed is derived from the QPI link speed of the installed processors, which in turn dictates the maximum SMI link speed, which in turn dictates the memory speed.

Table 3-9 on page 74 summarizes the memory speeds of all the models of Intel Xeon 7500 series CPUs.

One important consideration when installing memory in MAX5 configurations is that the server must be fully populated before adding DIMMs to the MAX5. As we described in 2.3.2, “Memory DIMM placement” on page 23, you get the best performance by using all memory buffers and all DIMM sockets on the server first, and then add DIMMs to the MAX5.

Figure 3-18 on page 83 shows the numbering scheme for the DIMM slots on the MAX5, and the pairing of DIMMs in the MAX5. As DIMMs are added in pairs, they must be matched on a memory port (as shown by using the colors). For example, DIMM1 is matched to DIMM 8, DIMM 2 to DIMM 7, DIMM 20 to DIMM 21, and so on.

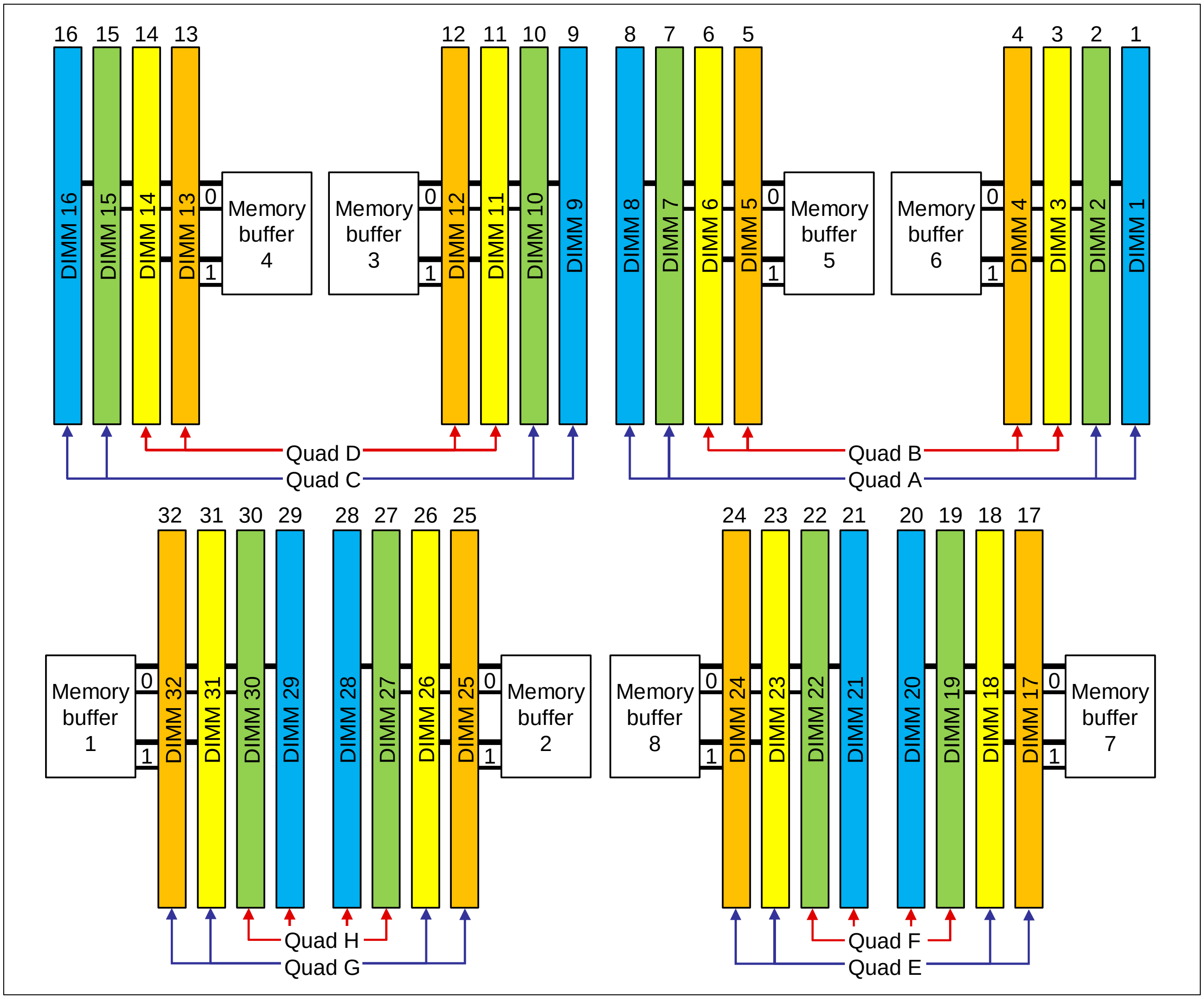


Figure 3-18 DIMM numbering on MAX5

Table 3-16 shows the population order of the MAX5 DIMM slots, ensuring that memory is balanced among the memory buffers. The colors in the table match the colors in Figure 3-18.

Table 3-16 DIMM installation sequence in the MAX5

DIMM pair	DIMM slots
1	28 and 29
2	9 and 16
3	1 and 8
4	20 and 21
5	26 and 31
6	11 and 14
7	3 and 6
8	18 and 23

DIMM pair	DIMM slots
9	27 and 30
10	10 and 15
11	2 and 7
12	19 and 22
13	25 and 32
14	12 and 13
15	4 and 5
16	17 and 24

MAX5 memory as seen by the operating system

MAX5 is capable of two modes of operation in terms of the way that memory is presented to the operating system:

- Memory in MAX5 can be split and assigned between the CPUs on the host system (partitioned mode). This mode is the default.
- Memory in MAX5 can be presented as a pool of space that is not assigned to any particular CPU (pooled mode).

By default, MAX5 is set to operate in partitioned mode because certain operating systems behave unpredictably when presented with a pool of memory space. Linux can work with memory that is presented either as a pool or pre-assigned between CPUs, however for performance reasons, if you are running Linux, change the setting to pooled mode.

You can change this default setting in UEFI.

VMware vSphere support: MAX5 requires VMware vSphere 4.1 or later.

3.8.3 Maximizing memory performance

In a single node x3850 X5 that is populated with four CPUs and eight memory cards, there are a total of 16 memory buffers, as shown in the system block diagram in Figure 3-7 on page 66. Memory buffers are listed as MB1 and MB2 on each of eight memory cards in that diagram. Each memory buffer has two memory channels, and each channel can have a maximum of two *DIMMs per channel* (DPC). A single-node x3850 X5 has the following maximums:

- Memory cards: 8
- Memory buffers: 16
- Memory channels: 32
- Number of DIMMs: 64

The x3850 X5 supports a variety of ways to install memory DIMMs in the eight memory cards. However, it is important to understand that because of the layout of the SMI links, memory buffers, and memory channels, you must install the DIMMs in the correct locations to maximize performance.

Figure 3-19 shows eight possible memory configurations for the two memory cards and 16 DIMMs connected to one processor socket. Each configuration has a relative performance score. Note the key information from this chart:

- The best performance is achieved by populating all memory DIMMs in two memory cards for each processor installed (configuration 1).
- Populating only one memory card per socket can result in approximately a 50% performance degradation (compare configuration 1 with 5).
- Memory performance is better if you install DIMMs on all memory channels than if you leave any memory channels empty (compare configuration 2 with 3).
- Two DIMMs per channel result in better performance that one DIMM per channel (compare configuration 1 with 2, and compare configuration 5 with 6).

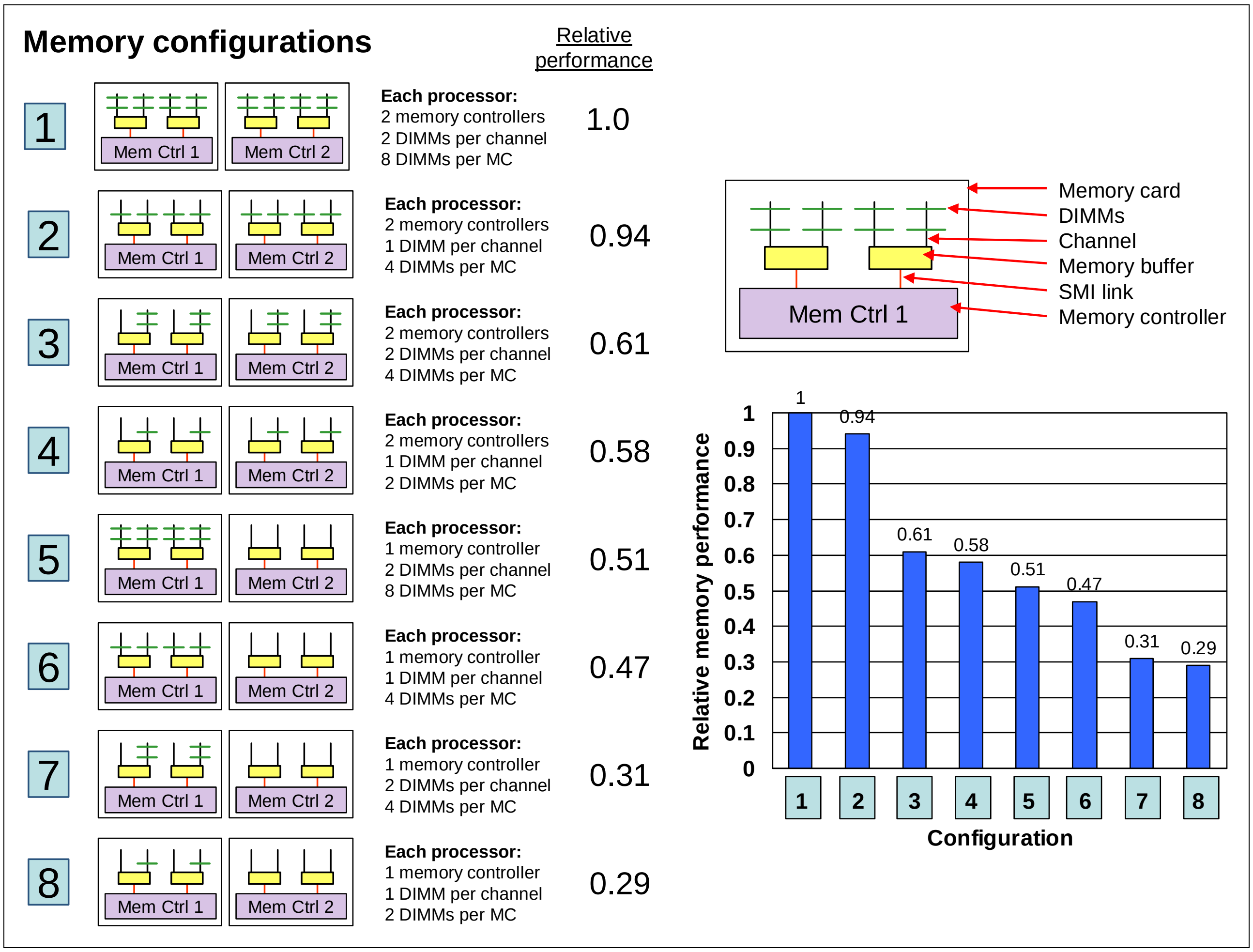


Figure 3-19 Relative memory performance based on DIMM placement (one processor and two memory cards shown)

Use the following general memory population rules:

- DIMMs must be installed in matching pairs.
- Each memory card requires at least two DIMMs.
- Each processor and memory card must have identical amounts of RAM.
- Install and populate two memory cards per processor or you can lose memory bandwidth.

- Populate one DIMM per channel on *every* channel on memory before populating a second DIMM in *any* channel.
- Populate DIMMs at the end of a memory channel first before populating the DIMM closer to the memory buffer. That is, install to sockets 1, 3, 6, and 8 first.
- If you have a mix of DIMM capacities (such as 4 GB and 8 GB DIMMs), insert the largest DIMMs first (spreading out the DIMMs across every memory channel), then move to next largest DIMMs, and finish with the smallest capacity DIMMs that you have.

Therefore, where memory performance is key to a successful deployment, the best configuration is to install 32 or 64 identical DIMMs across eight memory cards and four processors.

A system with fewer than four installed processors or fewer than eight installed memory cards has fewer memory channels, therefore less bandwidth and lower performance.

Mixed DIMMs and the effect on performance

Using DIMMs of various capacities (for example, 4 GB and 8 GB DIMMs) is supported. The capacities of the DIMMs might differ for several reasons:

- Not all applications require the full memory capacity that a homogenous memory population provides.
- Cost-saving requirements might dictate using a lower memory capacity for several of the platform's DIMMs.
- Certain configurations might attempt to use the DIMMs that came with the base platform, along with optional DIMMs of a separate type.

Figure 3-20 on page 87 illustrates the relative performance of three mixed memory configurations as compared to a baseline of a fully populated memory configuration. While these configurations use 4 GB (4R x8) and 2 GB (2R x8) DIMMs as specified, similar trends to this data are expected when using other mixed DIMM capacities. In all cases, memory is populated in minimum groups of four, as specified in the following configurations, to ensure that Hemisphere Mode is maintained.

Figure 3-20 on page 87 shows the following configurations:

- Configuration A: Full population of equivalent capacity DIMMs (2 GB). This represents an optimally balanced configuration.
- Configuration B: Each memory channel is balanced with the same memory capacity, but half of the DIMMs are of one capacity (4 GB), and half of the DIMMs are of another capacity (2 GB).
- Configuration C: Eight DIMMs of one capacity (4 GB) are populated across the eight memory channels, and four additional DIMMs (2 GB) are installed one per memory buffer, so that Hemisphere Mode is maintained.
- Configuration D: Four DIMMs of one capacity (4 GB) are populated across four memory channels, and four DIMMs of another capacity (2 GB) are populated on the other four memory channels, with configurations balanced across the memory buffers, so that Hemisphere Mode is maintained.

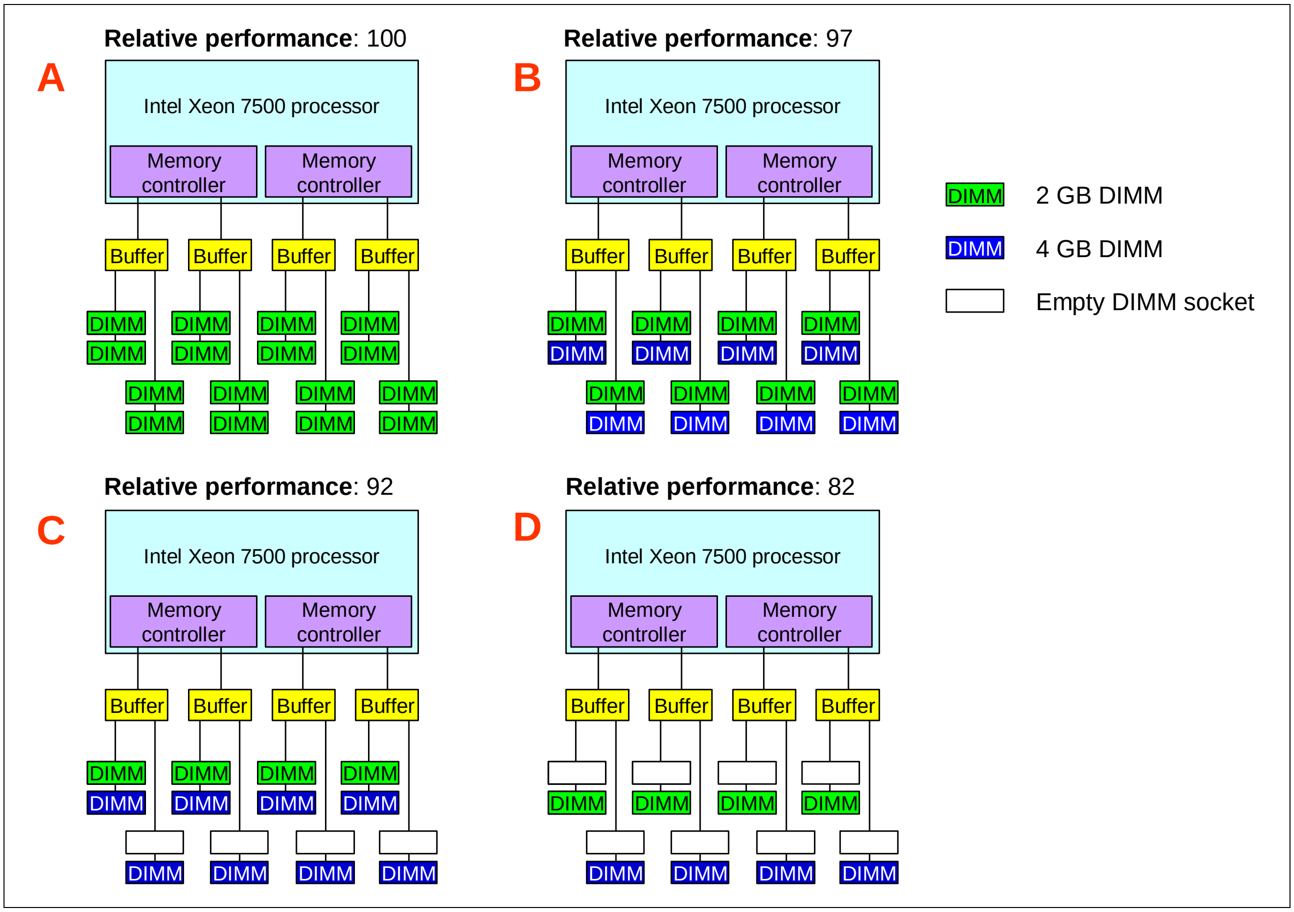


Figure 3-20 Relative memory performance using mixed DIMMs

As you can see, mixing DIMM sizes can cause performance loss up to 18%, even if all channels are occupied and Hemisphere Mode is maintained.

3.8.4 Memory mirroring

Memory mirroring is supported using x3850 X5 with or without the MAX5. To enable memory mirroring, you must install DIMMs in sets of four, one pair in each memory card. All DIMMs in each set must be the same size and type. Memory cards 1 and 2 mirror each other, cards 3 and 4 mirror each other, memory cards 5 and 6 mirror each other, and cards 7 and 8 mirror each other.

For x3850 X5, you install the memory evenly across all memory cards and then work your way to filling all eight memory cards for the best performance.

The source and destination cards that are used for memory mirroring are not selectable by the user. For a detailed understanding of memory mirroring, see “Memory mirroring” on page 28.

x3850 X5 memory mirroring population order

Table 3-17 on page 88 shows DIMM placements for each solution.

Table 3-17 x3850 X5 memory mirroring 4-processor 2-node

Number of DIMMs	Processor 1								Processor 4								Processor 2								Processor 3							
	Card 1				Card 2				Card 7				Card 8				Card 3				Card 4				Card 5				Card 6			
	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5	DIMM 1 and 8	DIMM 3 and 6	DIMM 2 and 7	DIMM 4 and 5
4	x				x																											
8	x				x				x				x																			
12	x				x				x				x				x				x											
16	x				x				x				x				x				x				x				x			
20	x	x			x	x			x				x				x				x				x				x			
24	x	x			x	x			x	x			x	x			x				x				x				x			
28	x	x			x	x			x	x			x	x			x	x			x	x			x				x			
32	x	x			x	x			x	x			x	x			x	x			x	x			x	x			x	x		
36	x	x	x		x	x	x		x	x			x	x			x	x			x	x			x	x			x	x		
40	x	x	x		x	x	x		x	x	x		x	x	x		x	x			x	x			x	x			x	x		
44	x	x	x		x	x	x		x	x	x		x	x	x		x	x	x		x	x	x		x	x			x	x		
48	x	x	x		x	x	x		x	x	x		x	x	x		x	x	x		x	x	x		x	x	x		x	x	x	
52	x	x	x	x	x	x	x	x	x	x	x		x	x	x		x	x	x		x	x	x		x	x	x		x	x	x	
56	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x		x	x	x		x	x	x		x	x	x	
60	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x								
64	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Table 3-18 shows the memory mirroring card pairs.

Table 3-18 Memory mirroring: Card pairs

Source card	Destination card
Memory card 2	Memory card 1
Memory card 4	Memory card 3
Memory card 6	Memory card 5
Memory card 8	Memory card 7

MAX5 memory mirroring population order

Table 3-19 on page 89 shows the installation guide for MAX5 memory mirroring.

Table 3-19 MAX5 memory mirroring setup

Number of DIMMs	MAX5																														
	DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16	DIMM 17	DIMM 18	DIMM 19	DIMM 20	DIMM 21	DIMM 22	DIMM 23	DIMM 24	DIMM 25	DIMM 26	DIMM 27	DIMM 28	DIMM 29	DIMM 30	DIMM 31	DIMM 32
4									x							x												x	x			
8	x							x	x							x				x	x							x	x			
12	x							x	x		x			x		x				x	x					x		x	x		x	
16	x		x			x		x	x		x			x		x		x		x	x		x			x		x	x		x	
20	x		x			x		x	x	x	x			x	x	x		x		x	x		x			x	x	x	x	x	x	
24	x	x	x			x	x	x	x	x	x			x	x	x		x	x	x	x	x	x			x	x	x	x	x	x	
28	x	x	x			x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x		x	x	x	x	x	x	x	x
32	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

3.8.5 Memory sparing

Sparing provides a degree of redundancy in the memory subsystem, but not to the extent of mirroring. For more information regarding memory sparing, see “Memory sparing” on page 29. Use these guidelines for installing memory for use with sparing. The two sparing options are DIMM sparing and rank sparing:

- DIMM sparing

  Two unused DIMMs are spared per memory card. These DIMMs must have the same rank and capacity as the largest DIMMs that we are sparing. The total size of the two unused DIMMs for sparing is subtracted from the usable capacity that is presented to the operating system. DIMM sparing is applied on all memory cards in the system.

- Rank sparing

  Two ranks per memory card are configured as spares. The ranks have to be as large as the rank relative to the highest capacity DIMM that we are sparing. The total size of the two unused ranks for sparing is subtracted from the usable capacity that is presented to the operating system. Rank sparing is applied on all memory cards in the system.

These options are configured by using the UEFI during boot.

3.8.6 Effect on performance by using mirroring or sparing

To understand the effect on performance by selecting various memory modes, we use a system that is configured with X7560 processors and populated with sixty-four 4 GB quad-rank DIMMs.

Figure 3-21 on page 90 shows the peak system-level memory throughput for various memory modes measured using an IBM-internal memory load generation tool. There is a 50% decrease in peak memory throughput when going from a normal (non-mirrored) memory configuration to a mirrored memory configuration.

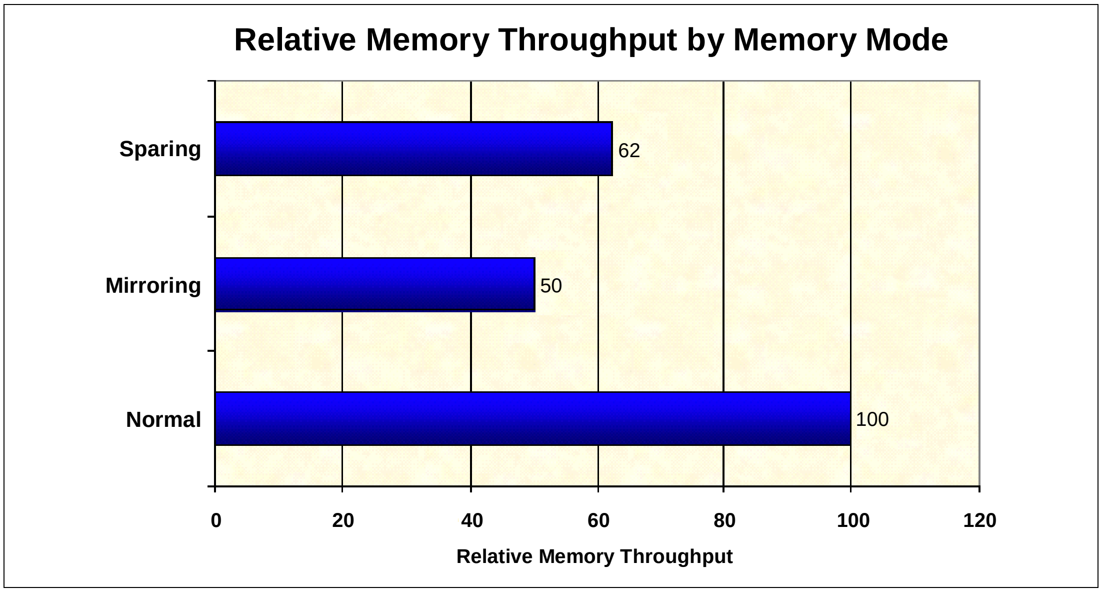


Figure 3-21 Relative memory throughput by memory mode

3.9 Storage

In this section, we look at the internal storage and RAID options for the x3850 X5, with suggestions where you can obtain the details about supported external storage arrays. This section includes the following topics:

- 3.9.1, “Internal disks” on page 90
- 3.9.2, “SAS and SSD 2.5-inch disk support” on page 91
- 3.9.3, “IBM eXFlash and 1.8-inch SSD support” on page 93
- 3.9.4, “SAS and SSD controllers” on page 96
- 3.9.6, “External storage connectivity” on page 101

3.9.1 Internal disks

The x3850 X5 supports one of the following sets of drives in the internal drive bays, accessible from the front of the system unit:

- Up to eight 2.5-inch SSDs
- Up to eight 2.5-inch SAS or SATA HDDs
- Up to sixteen 1.8-inch SSDs
- A mixture of up to four 2.5-inch drives and up to eight 1.8-inch SSDs

Figure 3-22 on page 91 shows the internal bays with eight 2.5-inch SAS drives.

Figure 3-22 Front view of the x3850 X5 showing eight 2.5-inch SAS drives

3.9.2 SAS and SSD 2.5-inch disk support

This section describes backplane, controller, and drive options for 2.5-inch disk drives and SSDs.

2.5-inch SAS disks and SSDs use the same backplane options. Most standard models of the x3850 X5 include one SAS backplane, supporting four 2.5-inch drives, as listed in 3.3, “Models” on page 64. You can add a second identical backplane to increase the supported number of SAS disks to eight (using part number 59Y6135). The standard backplane is always included in the lower of the two backplane bays.

“x3850 X5 backplane options” on page 91 lists the backplane option.

Table 3-20 x3850 X5 backplane options

Part number	Feature code	Description
59Y6135	3873	IBM Hot Swap SAS Hard Disk Drive Backplane (one standard, one optional); includes 250 mm SAS cable, supports four 2.5-inch drives

The SAS backplane uses a short SAS cable (included with the part number 59Y6135), and it is always controlled by the RAID adapter in the dedicated slot behind the disk cage, never from an adapter in the PCIe slots. The required power/signal “Y” cable is also included with the x3850 X5.

Up to two 2.5-inch backplanes (each holding up to four disks) can connect to a RAID controller installed in the dedicated RAID slot. Table 3-21 lists the supported RAID controllers. For more information about each RAID controller, see 3.9.4, “SAS and SSD controllers” on page 96.

Table 3-21 RAID controllers that are compatible with SAS backplane and SAS disk drives

Part number	Feature code	Description
44E8689	3577	ServeRAID BR10i (standard on most models; see 3.3, “Models” on page 64)
46M0831	0095	ServeRAID M1015 SAS/SATA Controller
46M0829	0093	ServeRAID M5015 SAS/SATA Controller[a]

Part number	Feature code	Description
46M0916	3877	ServeRAID M5014 SAS/SATA Controller
46M0969	3889	ServeRAID B5015 SSD
46M0930	5106	IBM ServeRAID M5000 Advance Feature Key: Adds RAID-6, RAID-60, and SED Data Encryption Key Management to the ServeRAID M5014, M5015, and M5025 controllers
81Y4426	A10C	IBM ServeRAID M5000 Performance Accelerator Key: Adds Cut Through I/O (CTIO) for SSD FastPath optimization on ServeRAID M5014, M5015, and M5025 controllers

a. The battery is not included with the ServeRAID M5015.

Table 3-22 lists the 2.5-inch SAS 10K and 15K RPM disk drives and the 2.5-inch SSD that are supported in the x3850 X5. These drives are supported with the SAS hard disk backplane, 59Y6135.

Table 3-22 Supported 2.5-inch SAS drives and 2.5-inch SSDs

Part number	Feature code	Description
42D0632	5537	IBM 146 GB 10K 6 Gbps SAS 2.5-inch SFF Slim-HS HDD
42D0637	5599	IBM 300 GB 10K 6 Gbps SAS 2.5-inch SFF Slim-HS HDD
42D0672	5522	IBM 73 GB 15K 6 Gbps SAS 2.5-inch SFF Slim-HS HDD
42D0677	5536	IBM 146 GB 15K 6 Gbps SAS 2.5-inch SFF Slim-HS HDD
43W7714	3745	IBM 50 GB SATA 2.5-inch SFF Slim-HS High IOPS SSD

Table 3-23 lists the 2.5-inch Nearline SATA 7.2K drives that are supported in the x3850 X5. These drives are supported with the SAS hard disk backplane, part number 59Y6135.

Table 3-23 Supported 2.5-inch Nearline SATA drives

Part number	Feature code	Description
42D0747	5405	IBM 160GB 7200 NL SATA 2.5" SFF Slim-HS HDD
42D0752	5407	IBM 500GB 7200 NL SATA 2.5" SFF Slim-HS HDD

The 2.5-inch drives require less space than 3.5-inch drives, consume half the power, produce less noise, can seek faster, and offer increased reliability.

Compatibility: As listed in Table 3-22, the 2.5-inch 50 GB SSD is also supported with the standard SAS backplane and the optional SAS backplane, part number 59Y6135. It is incompatible with the 1.8-inch SSD eXFlash backplane, part number 59Y6213.

A typical configuration can be two 2.5-inch SAS disks for the operating system and two High IOPS disks for data. Only the 2.5-inch High IOPS SSD disk can be used on the SAS backplane. The 1.8-inch disks for the eXFlash cannot be used on the SAS backplane.

3.9.3 IBM eXFlash and 1.8-inch SSD support

Database-optimized models of the x3950 X5 include one IBM eXFlash SSD backplane, supporting eight 1.8-inch solid-state drives, as listed in Table 3-3 on page 65. Other models also support the addition of an eXFlash backplane, controllers, and SSDs.

You can add a second eXFlash backplane to increase the supported number of SSDs to 16 (using part number 59Y6213, as listed in Table 3-24). See 3.9.3, “IBM eXFlash and 1.8-inch SSD support” on page 93 for more information.

Table 3-24 IBM eXFlash 8x 1.8-inch HS SAS SSD Backplane

Part number	Feature code	Description
59Y6213	4191	IBM eXFlash 8x 1.8-inch HS SAS SSD Backplane (two optional, replacing the standard SAS backplane); includes a set of cables

The IBM eXFlash 8x 1.8-inch HS SAS SSD Backplane, part number 59Y6213, supports eight 1.8-inch SSDs. The eight drive bays require the same physical space as four SAS hard disk bays. A single eXFlash backplane requires two SAS x4 input cables and a power/configuration cable, which are both shipped standard. Up to two SSD backplanes and 16 SSDs are supported in the x3850 X5 chassis.

For more information regarding eXFlash and SSD information, including a brief overview of the benefits of using eXFlash, see 2.8, “IBM eXFlash” on page 47.

Two-node configurations: Spanning an array on any disk type between two chassis is not possible because the RAID controllers operate separately.

Figure 3-23 shows an x3850 X5 with one of two eXFlash units installed.

Figure 3-23 IBM eXFlash with eight SSDs

Table 3-25 on page 94 lists the supported controllers.

Table 3-25 Controllers supported with the eXFlash SSD backplane option

Part number	Feature code	Description
46M0914	3876	IBM 6Gb SSD Host Bus Adapter (No RAID support)
46M0831	0095	ServeRAID M1015 SAS/SATA Controller
46M0829	0093	ServeRAID M5015 SAS/SATA Controller[a]
46M0916	3877	ServeRAID M5014 SAS/SATA Controller[a]
46M0969	3889	ServeRAID B5015 SSD[a]
46M0930	5106	IBM ServeRAID M5000 Advance Feature Key: Adds RAID-6, RAID-60, and self-encrypting drives (SED) Data Encryption Key Management to the ServeRAID M5014, M5015, and M5025 controllers
81Y4426	A10C	IBM ServeRAID M5000 Performance Accelerator Key: Adds Cut Through I/O (CTIO) for SSD FastPath optimization on ServeRAID M5014, M5015, and M5025 controllers

a. When using SSD drives, you must disable the write back cache to prevent latency and bottlenecks by using the controller settings or by adding the ServeRAID M5000 Series Performance Accelerator Key. See “ServeRAID M5000 Series Performance Accelerator Key” on page 95 for more information.

When ordering M5000 series controllers (M5014, M5015, or M5025) for use only with SSD drives, the cache battery must not be used for performance reasons. If using M5000 series controllers in a mixed SSD and SAS environment cache, order the battery along with the Performance Accelerator Key.

If the ServeRAID controller being used is already set up and you want to leave the battery attached, you can still disable the write back cache by using the MegaRAID web BIOS, as shown in Figure 3-24.

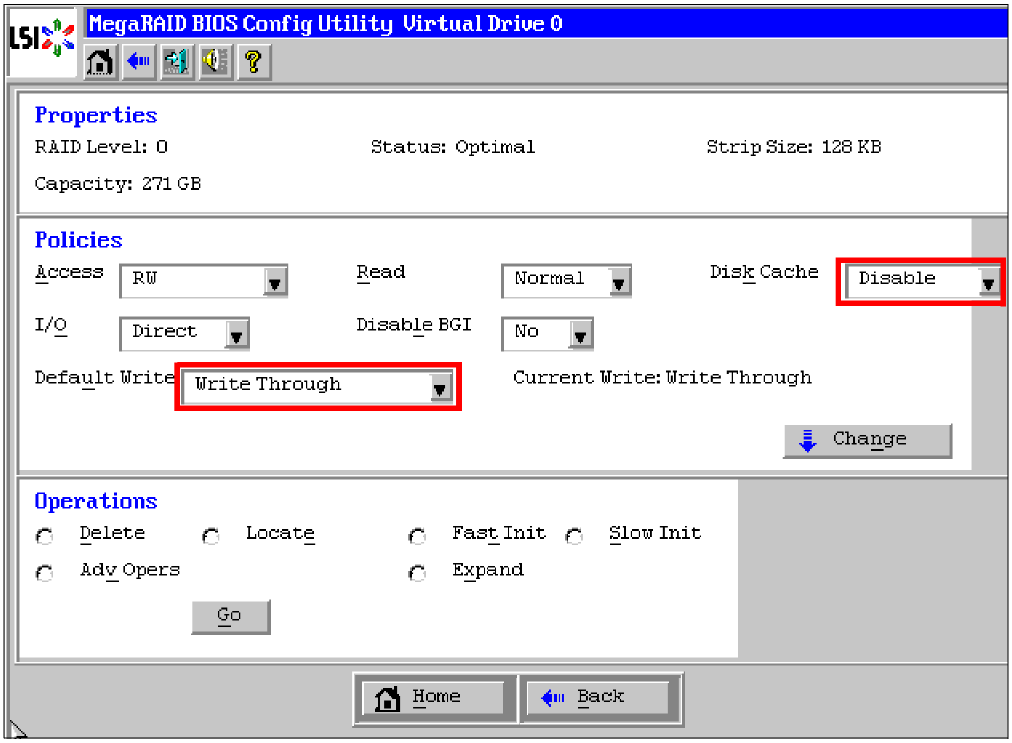


Figure 3-24 Disabling battery cache on controller in MegaRAID web BIOS

ServeRAID M5000 Series Performance Accelerator Key

ServeRAID M5000 Series Performance Accelerator Key for System x enables performance enhancements that are needed by emerging SSD technologies that are being used in a mixed SAS and SSD environment. You use a seamless field-upgradeable key. ServeRAID M5000 Series Performance Accelerator Key for System x provides these benefits:

- Performance optimization for SSDs: Improve SAS/SATA Controller performance to match an array of SSDs.
- Flash tiering enablement: A data-tiering enabler to support hybrid environments of SSDs and HDDs, realizing higher levels of performance.
- MegaRAID recovery: A data recovery feature that works both in pre-boot and OS environments.
- RAID 6, 60 enablement for added data protection.
- SED support enablement for encryption-equipped devices.
- Convenient upgrade with easy-to-use pluggable key.

We cover these controllers in detail in 3.9.4, “SAS and SSD controllers” on page 96.

1.8-inch hard drive options

Table 3-26 lists the supported 1.8-inch SSDs.

Table 3-26 IBM 1.8-inch SSD for use in the IBM eXFlash backplanes

Option number	Feature code	Description
43W7734	5314	IBM 50GB SATA 1.8-inch NHS SSD

The failure rate of SSDs is low because, in part, the drives have no moving parts. The 50 GB High IOPS SSD is a Single Level Cell (SLC) device with Enterprise Wear Leveling. As a consequence of both of these technologies, the additional layer of protection that is provided by a RAID controller might not always be necessary in every client environment and, in certain cases, RAID-0 might even be an acceptable option.

3.9.4 SAS and SSD controllers

Table 3-27 lists the disk controllers that are supported in the x3850 X5.

Table 3-27 Disk controllers that are compatible with the x3850 X5

Part number	Feature code	Name	Supports 2.5-inch SAS backplane	Supports eXFlash SSD backplane	Dedicated slot[a]	Battery	Cache	RAID support
44E8689	3577	ServeRAID BR10i[b]	Yes	No	Yes	No	None	0,1,1E
46M0831	0095	ServeRAID M1015	Yes	Yes	Yes	No	None	0,1,10,5,50[c]
46M0916	3877	ServeRAID M5014	Yes	Yes	Yes	Optional	256MB	0,1,10,5,50,6,60[d]
46M0829	0093	ServeRAID M5015	Yes	Yes	Yes	Yes[e]	512MB	0,1,10,5,50,6,60[d]
46M0830	0094	ServeRAID M5025	No	No	No	Yes	512MB	0,1,10,5,50,6,60[d]
None[f]	3876	IBM 6Gb SSD HBA	No	Yes	No	No	None	No
46M0969	3889	ServeRAID B5015 SSD	No	Yes	No	No	None	1,5

a. See 3.9.5, “Dedicated controller slot” on page 100.
b. The BR10i is standard on most models. See 3.3, “Models” on page 64.
c. M1015 support for RAID-5 and RAID-50 requires the M1000 Advanced Feature Key (46M0832, 9749).
d. M5014, M5015, and M5025 support for RAID-6 and RAID-60 requires the M5000 Advanced Feature Key (46M0930, fc 5106).
e. ServeRAID M5015 option part number 46M0829 includes the M5000 battery; however, the feature code 0093 does not contain the battery. Order feature code 5744 if you want to include the battery in the server configuration.
f. The IBM 6Gb SSD Host Bus Adapter is currently not available as a separately orderable option. Use the feature code to add the adapter to a customized order, using the configure-to-order (CTO) process. Part number 46M0914 is the L1 manufacturing part number. Part number 46M0983 is the pseudo option number, which is also used in manufacturing.

RAID levels 0 and 1 are standard on all models (except 7145-ARx) with the integrated BR10i ServeRAID controller. Model 7145-ARx has no RAID capability standard.

All servers, even those servers that are not standard with the BR10i (model 7145-ARx), include the blue mounting bracket (see Figure 3-25 on page 101), which allows for the easy installation of a supported RAID controller in the dedicated x8 PCIe slot behind the disk cage. Only RAID controllers that are supported with the 2.5-inch SAS backplane can be used in this slot. See Table 3-27 for a summary of these supported options.

ServeRAID BR10i Controller

The ServeRAID-BR10i has the following specifications:

- LSI 1068e-based adapter
- Two internal mini-SAS SFF-8087 connectors
- SAS 3 Gbps PCIe x8 host bus interface
- Fixed 64 KB stripe size

- Supports RAID-0, RAID-1, and RAID-1E
- No battery and no onboard cache

ServeRAID M5014 and M5015 controllers

The ServeRAID M5014 and M5015 adapter cards have the following specifications:

- Eight internal 6 Gbps SAS/SATA ports
- Two mini-SAS internal connectors (SFF-8087)
- Throughput of 6 Gbps per port
- An 800 MHz PowerPC® processor with LSI SAS2108 6 Gbps RAID on Chip (ROC) controller
- x8 PCI Express 2.0 host interface
- Onboard data cache (DDR2 running at 800 MHz):
  - ServeRAID M5015: 512 MB
  - ServeRAID M5014: 256 MB
- Intelligent battery backup unit with up to 48 hours of data retention:
  - ServeRAID M5015: Optional for feature code 0093, standard for part 46M0829
  - ServeRAID M5014: Optional

  > **Battery cache:** Battery cache is not needed when using all SSD drives. If using a controller in a mixed environment with SSD and SAS, you must order and use the battery and the performance enablement key.

- RAID levels 0, 1, 5, 10, and 50 support (RAID 6 and 60 support with the optional M5000 Advanced Feature Key)
- Connection of up to 32 SAS or SATA drives
- SAS and SATA drive support (however, the mixing of SAS and SATA in the same RAID array is not supported)
- Up to 64 logical volumes
- Logical unit number (LUN) sizes up to 64 TB
- Configurable stripe size up to 1 MB
- Compliance with Disk Data Format (DDF) configuration on disk (COD)
- Self-Monitoring, Analysis, and Reporting Technology (S.M.A.R.T.) support
- Support for the optional M5000 Series Performance Accelerator Key, which is recommended when using SSD drives in a mixed environment with SAS and SSD:
  - RAID levels 6 and 60
  - Performance optimization for SSDs
  - LSI SafeStore: Support for self-encrypting drive services, such as instant secure erase and local key management (which requires the use of self-encrypting drives)
- Support for the optional M5000 Advanced Feature Key, which enables the following features:
  - RAID levels 6 and 60
  - LSI SafeStore: Support for self-encrypting drive services, such as instant secure erase and local key management (which requires the use of self-encrypting drives)

Performance Key Accelerator: Performance Accelerator Key uses the same features as the Advanced Feature Key, but it also includes performance enhancements to enable SSD support in a mixed HDD environment.

For more information, see *ServeRAID M5015 and M5014 SAS/SATA Controllers for IBM System x*, TIPS0738, at the following website:

`http://www.redbooks.ibm.com/abstracts/tips0738.html?Open`

ServeRAID M5025 Controller

The key difference between the ServeRAID M5025 and M5015 RAID controllers is that the M5025 has two *external* SAS 2.0 x4 connectors and the M5015 has two *internal* SAS 2.0 x4 connectors. The ServeRAID M5025 Controller offers these benefits:

- Eight external 6 Gbps SAS 2.0 ports implemented through two four-lane (x4) connectors
- Two mini-SAS external connectors (SFF-8088)
- Six Gbps throughput per SAS port
- 800 MHz PowerPC processor with LSI SAS2108 6 Gbps RAID on Chip (ROC) controller
- PCI Express 2.0 x8 host interface
- 512 MB onboard data cache (DDR2 running at 800 MHz)
- Intelligent lithium polymer battery backup unit standard with up to 48 hours of data retention
- Support for RAID levels 0, 1, 5, 10, and 50 (RAID-6 and RAID-60 support with the optional M5000 Advanced Feature Key)
- Connections:
  - Up to 240 SAS or SATA drives
  - Up to nine daisy-chained enclosures per port
- SAS and SATA drives are supported, but mixing SAS and SATA drives in the same RAID array is not supported
- Support for up to 64 logical volumes
- Support for LUN sizes up to 64 TB
- Configurable stripe size up to 1024 KB
- Compliant with Disk Data Format (DDF) configuration on disk (COD)
- S.M.A.R.T. support
- Support for the optional M5000 Series Performance Accelerator Key, which is recommended when using SSD drives in a mixed environment with SAS and SSD:
  - RAID levels 6 and 60
  - Performance optimization for SSDs
  - LSI SafeStore: Support for self-encrypting drive services, such as instant secure erase and local key management (which requires the use of self-encrypting drives)
- Support for the optional M5000 Advanced Feature Key, which enables the following features:
  - RAID levels 6 and 60
  - LSI SafeStore: Support for self-encrypting drive services, such as instant secure erase and local key management (which requires the use of self-encrypting drives)

Performance: Performance Accelerator Key uses the same features as the Advanced Feature Key. However, it also includes performance enhancements to enable SSD support in a mixed HDD environment.

For more information, see *ServeRAID M5025 SAS/SATA Controller for IBM System x*, TIPS0739, at the following website:

http://www.redbooks.ibm.com/abstracts/tips0739.html?Open

ServeRAID M1015 Controller

The ServeRAID M1015 SAS/SATA Controller has the following specifications:

- Eight internal 6 Gbps SAS/SATA ports
- SAS and SATA drives support (but not in the same RAID volume)
- SSD support
- Two mini-SAS internal connectors (SFF-8087)
- Throughput of 6 Gbps per port
- LSI SAS2008 6 Gbps RAID on Chip (ROC) controller
- x8 PCI Express 2.0 host interface
- RAID levels 0, 1, and 10 support (RAID levels 5 and 50 with optional ServeRAID M1000 Series Advanced Feature Key)
- Connection of up to 32 SAS or SATA drives
- Up to 16 logical volumes
- LUN sizes up to 64 TB
- Configurable stripe size up to 64 KB
- Compliant with Disk Data Format (DDF) configuration on disk (COD)
- S.M.A.R.T. support

RAID-5, RAID-50, and self-encrypting drives (SED) technology are optional upgrades to the ServeRAID M1015 adapter with the addition of the ServeRAID M1000 Series Advanced Feature Key, part number 46M0832, feature code 9749.

For more information, see *ServeRAID M1015 SAS/SATA Controller for System x*, TIPS0740, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0740.html?Open

IBM 6Gb SSD Host Bus Adapter

The IBM 6Gb SSD Host Bus Adapter is an ideal host bus adapter (HBA) to connect to high-performance SSDs. With two x4 SFF-8087 connectors and a high performance PowerPC I/O processor, this HBA can support the bandwidth that SSDs can generate.

The IBM 6Gb SSD Host Bus Adapter has the following high-level specifications:

- PCI Express 2.0 host interface
- Six Gbps per port data transfer rate
- MD2 small form factor
- PCI Express 2.0 x8 host interface
- High performance I/O Processor: PowerPC 440 at 533MHz
- UEFI support

For more information, see *IBM 6Gb SSD Host Bus Adapter for IBM System x*, TIPS0744, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0744.html?Open

> **Important:** Two variants of the 6Gb Host Bus Adapter exist. The SSD variant has no external port and is part number 46M0914. Do not confuse it with the IBM 6Gb SAS HBA, part number 46M0907, which is not supported for use with eXFlash.

ServeRAID B5015 SSD Controller

The ServeRAID B5015 is a high-performance RAID controller that is optimized for SSDs. It has the following specifications:

- RAID 1 and 5 support
- Hot-spare support with automatic rebuild capability
- Background data scrubbing
- Stripe size of up to 1 MB
- Six Gbps per SAS port
- PCI Express 2.0 x8 host interface
- PCI MD2 low profile form factor
- Two x4 internal (SFF-8087) connectors
- SAS controller: PMC-Sierra PM8013 maxSAS 6 Gbps SAS RoC controller
- Up to eight disk drives per RAID adapter
- Performance that is optimized for SSDs
- Three multi-threading MIPS processing cores
- High performance contention-free architecture
- Up to four ServeRAID B5015 adapters supported in a system
- Support for up to four arrays/logical volumes

For more information, see *ServeRAID B5015 SSD Controller*, TIPS0763, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0763.html?Open

> **Important:** This controller is listed in power-on self test (POST) and in UEFI as a PMC-Sierra card. This controller uses the maxRAID Storage Manager for management, not MegaRAID.

3.9.5 Dedicated controller slot

As listed in Table 3-27 on page 96, certain supported controllers (including the ServeRAID BR10i that is standard in most models) can be installed in a single PCIe x8 dedicated slot on the side of the server, near the front. Figure 3-25 on page 101 shows the ServeRAID M5015 adapter installed on the side of the server, near the front with an installation bracket attached (blue plastic handle).

The blue plastic carrier is reusable and is included with the server (attached to the standard BR10i). The latch and edge clips allow the card to be removed and replaced with another supported card as required.

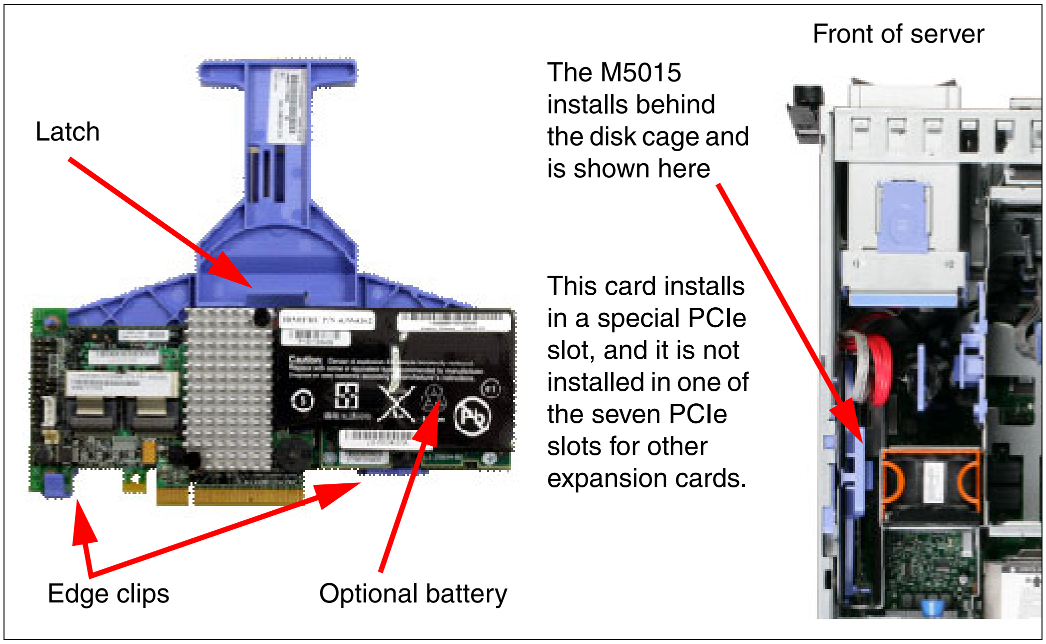


Figure 3-25 ServeRAID M5015 SAS/SATA Controller

RAID-6, RAID-60, and encryption are a further optional upgrade for the M5015 through the ServeRAID M5000 Series Advance Feature Key or the Performance Accelerator Key.

3.9.6 External storage connectivity

The ServeRAID M5025 offers two external SAS ports to connect to external storage. Table 3-28 lists the cards and support cables and feature keys.

Table 3-28 External ServeRAID card

Option	Feature code	Description
46M0830	0094	IBM 6Gb ServeRAID M5025 External RAID
39R6531	3707	IBM 3m SAS External Cable for ServeRAID M5025 to an EXP2512 (1747-HC1) or EXP2524 (1747-HC2)
39R6529	3708	IBM 1m SAS External Cable for interconnection between multiple EXP2512 (1747-HC1) or EXP2524 (1747-HC2)
46M0930	5106	IBM ServeRAID M5000 Advance Feature Key: Adds RAID-6, RAID-60, and SED Data Encryption Key Management to the ServeRAID M5025 controller

The M5025 has two *external* SAS 2.0 x4 connectors and supports the following features:

- Eight external 6 Gbps SAS 2.0 ports implemented through two four-lane (x4) connectors.
- Two mini-SAS external connectors (SFF-8088).
- Six Gbps throughput per SAS port.
- 800 MHz PowerPC processor with LSI SAS2108 6 Gbps RAID on Chip (ROC) controller.
- PCI Express 2.0 x8 host interface.
- 512 MB onboard data cache (DDR2 running at 800 MHz).

- Intelligent lithium polymer battery backup unit is standard with up to 48 hours of data retention.
- Support for RAID levels 0, 1, 5, 10, and 50 (RAID-6 and 60 support with either the optional M5000 Advanced Feature Key or the optional M5000 Performance Key).
- Connections:
  - Up to 240 SAS or SATA drives
  - Up to 9 daisy-chained enclosures per port
- SAS and SATA drives are supported, but mixing SAS and SATA in the same RAID array is not supported.
- Support for up to 64 logical volumes.
- Support for LUN sizes up to 64 TB.
- Configurable stripe size up to 1024 KB.
- Compliant with Disk Data Format (DDF) configuration on disk (COD).
- S.M.A.R.T. support.
- Support for the optional M5000 Advanced Feature Key, which enables the following features:
  - RAID levels 6 and 60
  - LSI SafeStore: Support for SED services, such as instant secure erase and local key management (which requires the use of self-encrypting drives).
- Support for SSD drives in a mixed environment with SAS and SSD with the optional M5000 Series Performance Accelerator Key, which enables the following features:
  - RAID levels 6 and 60
  - Performance optimizations for SSDs
  - LSI SafeStore: Support for SED services, such as instant secure erase and local key management (which requires the use of self-encrypting drives).

For more information, see the IBM Redbooks at-a-glance guide *ServeRAID M5025 SAS/SATA Controller for IBM System x*, TIPS0739, which is available at this website:

http://www.redbooks.ibm.com/abstracts/tips0739.html?Open

The x3850 X5 is qualified with a wide range of external storage options. To view the available solutions, see the Configuration and Options Guide, which is available at this website:

http://ibm.com/systems/xbc/cog/x3850x5_7145/x3850x5_7145io.html

The System Storage Interoperation Center (SSIC) is a search engine that provides details about supported configurations:

http://www.ibm.com/systems/support/storage/config/ssic

3.10 Optical drives

An optical drive is optional in the x3850 X5. Table 3-29 lists the supported part numbers.

Table 3-29 Optical drives

Part number	Feature code	Description
46M0901	4161	IBM UltraSlim Enhanced SATA DVD-ROM

Part number	Feature code	Description
46M0902	4163	IBM UltraSlim Enhanced SATA Multi-Burner

3.11 PCIe slots

The x3850 X5 has a total of seven PCI Express (PCIe) slots. Slot 7 holds the Emulex 10Gb Ethernet Adapter that is standard in most models (see 3.3, "Models" on page 64). We describe the Emulex 10Gb Ethernet Adapter in 3.12.1, "Standard Emulex 10Gb Ethernet Adapter" on page 104.

The RAID card that is used in the x3850 X5 to control 2.5-inch SAS disks has a dedicated slot behind the disk cage and does not consume one of the seven available PCIe slots. For further details about supported RAID cards, see 3.9.4, "SAS and SSD controllers" on page 96.

Table 3-30 lists the PCIe slots.

Table 3-30 PCI Express slots

Slot	Host interface	Length
1	PCI Express 2.0 x16	Full length
2	PCI Express 2.0 x4 (x8 mechanical)	Full length
3	PCI Express 2.0 x8	Full length
4	PCI Express 2.0 x8	Full length
5	PCI Express 2.0 x8	Half length
6	PCI Express 2.0 x8	Half length
7	PCI Express 2.0 x8	Half length (Emulex 10Gb Ethernet Adapter)
Dedicated	PCI Express 2.0 x8	Dedicated RAID controller side slot

All slots are PCI Express 2.0, full height, and *not* hot-swap. PCI Express 2.0 has several improvements over PCI Express 1.1 (as implemented in the x3850 M2). The chief benefit is the enhanced throughput. PCI Express 2.0 is rated for 500 MBps per lane (5 Gbps per lane); PCI Express 1.1 is rated for 250 MBps per lane (2.5 Gbps per lane).

Note the following information about the slots:

- Slot 1 can accommodate a double-wide x16 card, but access to slot 2 is then blocked.
- Slot 2 is described as *x4 (x8 mechanical)*. This host interface is sometimes shown as x4 (x8) and means that the slot is only capable of x4 speed, but is physically large enough to accommodate an x8 card. Any x8-rated card physically fits in the slot, but it runs at only x4 speed. Do not add RAID cards to this slot, because RAID cards in this slot cause bottlenecks and possible crashes.
- Slot 7 has been extended in length to 106 pins, making it a nonstandard connector. It still accepts PCI Express x8, x4, and x1 standard adapters. It is the only slot that is compatible with the extended edge connector on the Emulex 10Gb Ethernet Adapter, which is standard with most models.

- Slots 5 - 7, the onboard Broadcom-based Ethernet dual-port chip and the custom slot for the RAID controller are on the first PCIe bridge and require that either CPU 1 or 2 is installed and operational.
- Slots 1 - 4 are on the second PCIe bridge and require that either CPU 3 or 4 is installed and operational.

Table 3-31 shows the order in which to add cards to balance bandwidth between the two PCIe controllers. However, this installation order assumes that the cards are installed in matched pairs, or that they have similar throughput capabilities.

Table 3-31 Order for adding cards

Installation order	PCIe slot	Slot width	Slot bandwidth[a]
1	1	x16 PCIe slot	8 GBps (80 Gbps)
2	5	x8 PCIe slot	4 GBps (40 Gbps)
3	3	x8 PCIe slot	4 GBps (40 Gbps)
4	6	x8 PCIe slot	4 GBps (40 Gbps)
5	4	x8 PCIe slot	4 GBps (40 Gbps)
6	7	x8 PCIe slot	4 GBps (40 Gbps)
7	2	x4 PCIe slot	2 GBps (20 Gbps)

a. This column correctly shows bandwidth expressed as GB for gigabyte or Gb for gigabit. 10 bits of traffic correspond to 1 byte of data due to the 8:10 encoding scheme. A single PCIe 2.0 lane provides a unidirectional bandwidth of 500 MBps or 5 Gbps.

Two additional power connectors, one 2x4 and one 2x3, are provided on the planar for high-power adapters, such as graphics cards. If there is a requirement to use an x16 PCIe card that is not shown as supported in ServerProven, initiate the SPORE process.

To determine whether a vendor has qualified any x16 cards with the x3850 X5, see IBM ServerProven at the following website:

http://www.ibm.com/servers/eserver/serverproven/compat/us/serverproven

If the preferred vendor's logo is displayed, click it to assess options that the vendor has qualified on the x3850 X5. You can obtain the support caveats for third-party options in 3.12.2, "Optional adapters" on page 107.

In a 2-node configuration, all PCIe slots are available to the operating system that is running on the complex. They appear as devices on separate PCIe buses.

3.12 I/O cards

This section describes the I/O cards that are suitable for the x3850 X5.

3.12.1 Standard Emulex 10Gb Ethernet Adapter

As described in 3.3, "Models" on page 64, most models include the Emulex 10Gb Ethernet Adapter as standard. The card is installed in PCIe slot 7. Slot 7 is a nonstandard x8 slot, which is slightly longer than normal, and it is shown in Figure 3-26 on page 105.

Tip: The Emulex 10Gb Ethernet Adapter that is standard with most models is a custom version of the Emulex 10Gb Virtual Fabric Adapter for IBM System x, part number 49Y4250. However, the features and functions of the two adapters are identical.

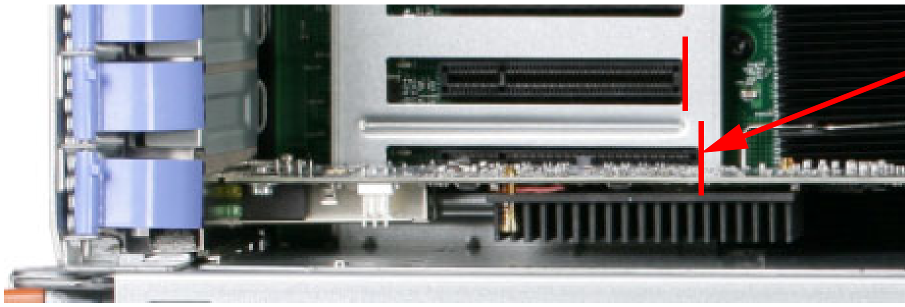


Figure 3-26 Top view of slot 6 and 7 showing that slot 7 is slightly longer than slot 6

The Emulex 10Gb Ethernet Adapter in the x3850 X5 has been customized with a special type of connector called an *extended edge connector*. The card itself is colored blue instead of green to indicate that it is nonstandard and cannot be installed in a standard x8 PCIe slot.

At the time of writing, only the x3850 X5 and the x3690 X5 have slots that are compatible with the custom-built Emulex 10Gb Ethernet Adapter that is shown in Figure 3-27.

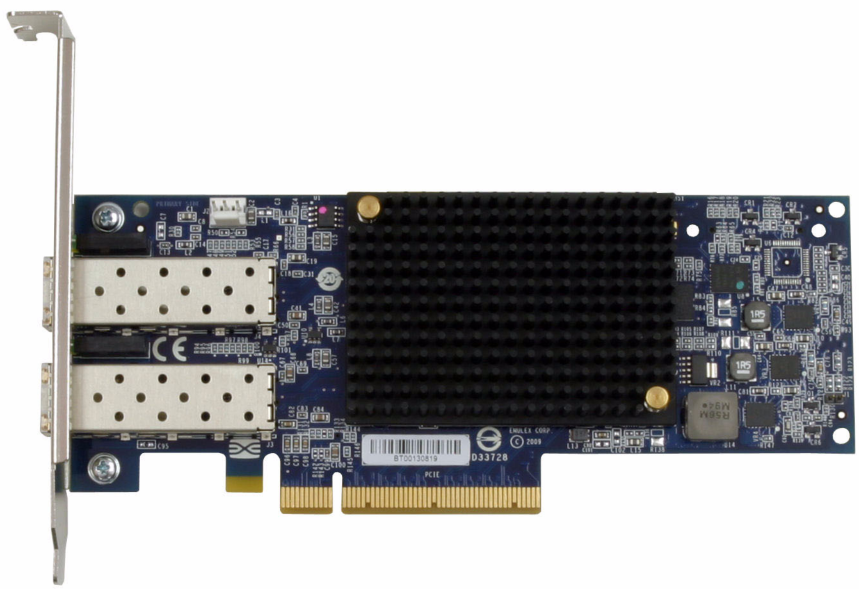


Figure 3-27 The Emulex 10Gb Ethernet Adapter has a blue circuit board and a longer connector

The Emulex 10Gb Ethernet Adapter is a customer-replaceable unit (CRU). To replace the adapter (for example, under warranty), order the CRU number, as shown in Table 3-32 on page 106. The table also shows the regular Emulex 10Gb Virtual Fabric Adapter (VFA) for IBM System x option, which differs only in the connector type (standard x8) and color of the circuit board (green).

Emulex VFA: The standard version of the Emulex VFA and the eX5 extended edge custom version can be used together as a redundant pair. This pair is a supported combination.

Table 3-32 Emulex adapter part numbers

Option description	Part number	Feature code	CRU number
Emulex 10Gb Ethernet Adapter for x3850 X5	None	1648	49Y4202
Emulex 10Gb Virtual Fabric Adapter for IBM System x	49Y4250	5749	Not applicable

General details about this card are in *Emulex 10Gb Virtual Fabric Adapter for IBM System x*, TIPS0762, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0762.html

Important: Although these cards are functionally identical, the availability of iSCSI and Fibre Channel over Ethernet (FCoE) upgrades for one card does not automatically mean availability for both cards. At the time of writing, the target availability of these features is the second quarter of 2011. Check the availability of iSCSI and FCoE feature upgrades with your local IBM representative.

The Emulex 10Gb Ethernet Adapter for x3850 X5 offers the following features:

- Dual-channel, 10 Gbps Ethernet controller
- Near line rate 10 Gbps performance
- Two SFP+ empty cages to support either of the following items:
  - SFP+ SR link is with SFP+ SR Module with LC connectors
  - SFP+ twinaxial copper link is with SFP+ direct-attached copper module/cable

  Note: Servers that include the Emulex 10Gb Ethernet Adapter do not include transceivers. You must order transceivers separately if needed, as listed in Table 3-33.

- TCP/IP stateless off-loads
- TCP chimney off-load
- Based on Emulex OneConnect technology
- FCoE support as a future feature entitlement upgrade
- Hardware parity, cyclic redundancy check (CRC), error checking and correcting (ECC), and other advanced error checking
- PCI Express 2.0 x8 host interface
- Low-profile form-factor design
- IPv4/IPv6 TCP, user datagram protocol (UDP) checksum off-load
- Virtual LAN (VLAN) insertion and extraction
- Support for jumbo frames up to 9000 bytes
- Preboot eXecutive Environment (PXE) 2.0 network boot support
- Interrupt coalescing
- Load balancing and failover support

- Deployment of this adapter and other Emulex OneConnect-based adapters with OneCommand Manager
- Interoperable with BNT 10Gb Top of Rack (ToR) switch for FCoE functions
- Interoperable with Cisco Nexus 5000 and Brocade 10Gb Ethernet switches for NIC/FCoE

SFP+ transceivers are not included with the server. You must order them separately. Table 3-33 lists the compatible transceivers.

Table 3-33 Transceiver ordering information

Option number	Feature code	Description
49Y4218	0064	QLogic 10Gb SFP+ SR Optical Transceiver
49Y4216	0069	Brocade 10Gb SFP+ SR Optical Transceiver
46C3447	5053	BNT SFP+ Transceiver

3.12.2 Optional adapters

Table 3-34 lists a selection of the expansion cards that are available for the x3850 X5.

Table 3-34 Available I/O adapters for the x3850 X5

Option	Feature code	Description
Networking		
59Y1887	5763	QLogic QLE7340 single-port 4X QDR IB x8 PCI-E 2.0 HCA
39Y6071	1485	NetXtreme II 1000 Express G Ethernet Adapter- PCIe
49Y4253	5749	Emulex 10GbE Virtual Fabric Adapter
49Y4243	5768	Intel Ethernet Quad Port Server Adapter I340-T4
49Y4233	5767	Intel Ethernet Dual Port Server Adapter I340-T2
49Y4223	5766	NetXtreme II 1000 Express Quad Port Ethernet Adapter
49Y4200	1648	Emulex 10Gb Dual-port Ethernet Adapter
42C1823	1637	Brocade 10Gb CNA
42C1803	5751	QLogic 10Gb CNA
42C1793	5451	NetXtreme II 10 GigE Express Fiber
42C1783	2995	NetXtreme II 1000 Express Dual Port Ethernet Adapter
42C1753	2975	PRO/1000 PF Server Adapter
39Y6139	2974	PRO/1000 PT Quad Port Server Adapter
39Y6129	2944	PRO/1000 PT Dual Port Server Adapter
Storage		
42D0486	3580	Emulex 8Gb FC Single-port HBA
42D0495	3581	Emulex 8Gb FC Dual-port HBA
42D0502	3578	QLogic 8Gb FC Single-port HBA

Option	Feature code	Description
42D0511	3579	QLogic 8Gb FC Dual-port HBA
46M6051	3589	Brocade 8Gb FC Single-port HBA
46M6052	3591	Brocade 8Gb FC Dual-port HBA
59Y1988	3885	Brocade 4Gb FC Single-port HBA
42C2182	3568	QLogic 4Gb FC Dual-Port PCIe HBA
43W7491	1698	Emulex 4GB FC Single-Port PCI-E HBA
43W7492	1699	Emulex 4GB FC Dual-Port PCI-E HBA
Graphics		
49Y6804	1826	NVIDIA Quadro FX 3800

This list is constantly updated and changed. To see the latest updates, see the following website:

http://ibm.com/systems/xbc/cog/x3850x5_7145/x3850x5_7145io.html

Tools, such as the Configurations and Options Guide (COG) or SSCT, contain information about supported part numbers. Many System x tools, including those tools that we have mentioned, are located on the following configuration tools website:

http://www.ibm.com/systems/x/hardware/configtools.html

See the ServerProven website for a complete list of the available options:

http://www.ibm.com/systems/info/x86servers/serverproven/compat/us/

In any circumstance where this list of options differs from the options that are shown in ServerProven, use ServerProven as the definitive resource. The main function of ServerProven is to show the options that have been successfully tested by IBM with a System x server.

Another useful page in the ServerProven site is the list of vendors. On the home page for ServerProven, click the **industry leaders** link, as shown in Figure 3-28.

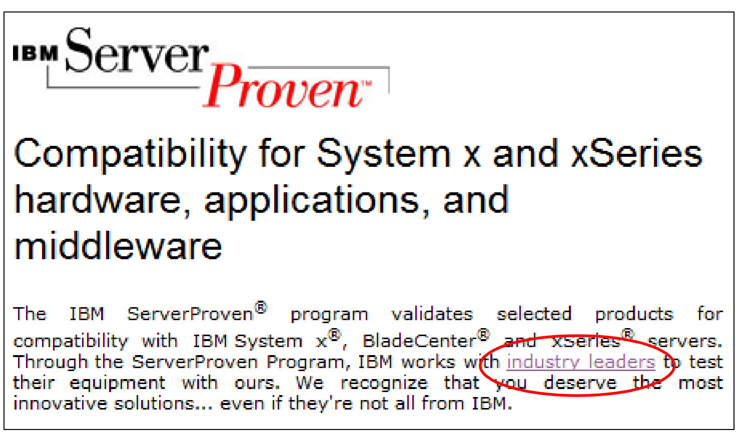


Figure 3-28 Link to vendor testing results

You can use the following ServerProven web address:

http://www.ibm.com/systems/info/x86servers/serverproven/compat/us/serverproven

This page lists the third-party vendors that have performed their own testing of their options with our servers. This support information means that those vendors agree to support the combinations that are shown in those particular pages.

Tip: To see the tested hardware, click the logo of the vendor. Clicking the About link under the logo takes you to a separate About page.

Although IBM supports the rest of the System x server, technical issues traced to the vendor card are, in most circumstances, directed to the vendor for resolution.

3.13 Standard onboard features

In this section, we look at several standard features in the x3850 X5.

3.13.1 Onboard Ethernet

The x3850 X5 has an embedded dual 10/100/1000 Ethernet controller, which is based on the Broadcom 5709C controller. The BCM5709C is a single-chip high performance multi-speed dual port Ethernet LAN controller. The controller contains two standard IEEE 802.3 Ethernet media access controls (MACs) that can operate in either full-duplex or half-duplex mode. Two direct memory access (DMA) engines maximize the bus throughput and minimize CPU overhead.

The onboard Ethernet offers these features:

- TCP off-load engine (TOE) acceleration
- Shared PCIe interface across two internal PCI functions with separate configuration space
- Integrated dual 10/100/1000 MAC and PHY devices able to share the bus through bridge-less arbitration
- Comprehensive nonvolatile memory interface
- Intelligent Peripheral Management Interface (IPMI)-enabled

3.13.2 Environmental data

x3850 X5 has the following environmental data:

- Heat output:
  - Minimum configuration: 734 Btu/hr (215 watts per hour)
  - Typical configuration: 2,730 Btu/hr (800 watts per hour)
  - Design maximum configuration:
    - 5,971 Btu/hr (1930 watts per hour) at 110 V ac
    - 6,739 Btu/hr (2150 watts per hour) at 220 V ac
- Electrical input: 100 - 127 V, 200 - 240 V, 50 - 60 Hz
- Approximate input kilovolt-amperes (kVA):
  - Minimum: 0.25 kVA
  - Typical: 0.85 kVA
  - Maximum: 1.95 kVA (110 V ac)
  - Maximum: 2.17 kVA (220 V ac)

3.13.3 Integrated Management Module (IMM)

The System x3850 X5 includes an IMM that provides industry-standard Intelligent Platform Management Interface (IPMI) 2.0-compliant systems management. You access the IMM through software that is compatible with IPMI 2.0 (xCAT, for example). You implement the IMM using industry-leading firmware from OSA and applications in conjunction with the Integrated Management Module.

The IMM delivers advanced control and monitoring features to manage your IBM System x3850 X5 server at virtually any time, from virtually anywhere. IMM enables easy console redirection with text and graphics, and keyboard and mouse support (operating system must support USB) over the system management LAN connections.

With video compression now built into the adapter hardware, it is designed to allow greater panel sizes and refresh rates that are becoming standard in the marketplace. This feature allows the user to display server activities from power-on to full operation remotely, with remote user interaction at virtually any time.

IMM monitors the following components:

- System voltages
- System temperatures
- Fan speed control
- Fan tachometer monitor
- Good Power signal monitor
- System ID and planar version detection
- System power and reset control
- Non-maskable interrupt (NMI) detection (system interrupts)
- SMI detection and generation (system interrupts)
- Serial port text console redirection
- System LED control (power, HDD, activity, alerts, and heartbeat)

IMM provides these features:

- An embedded web server, which gives you remote control from any standard web browser. No additional software is required on the remote administrator’s workstation.
- A command-line interface (CLI), which the administrator can use from a Telnet session.
- Secure Sockets Layer (SSL) and Lightweight Directory Access Protocol (LDAP).
- Built-in LAN and serial connectivity that support virtually any network infrastructure.
- Multiple alerting functions to warn systems administrators of potential problems through email, IPMI platform event traps (PETs), and Simple Network Management Protocol (SNMP).

3.13.4 UEFI

The x3850 X5 uses an integrated Unified Extensible Firmware Interface (UEFI) next-generation BIOS.

UEFI includes the following capabilities:

- Human-readable event logs; no more beep codes
- Complete setup solution by allowing adapter configuration function to be moved to UEFI
- Complete out-of-band coverage by Advanced Settings Utility to simplify remote setup

Using all of the features of UEFI requires an UEFI-aware operating system and adapters. UEFI is fully backward-compatible with BIOS.

For more information about UEFI, see the IBM white paper, *Introducing UEFI-Compliant Firmware on IBM System x and BladeCenter Servers*, which is available at the following website:

http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5083207

For UEFI menu setup, see 6.9, "UEFI settings" on page 259.

3.13.5 Integrated Trusted Platform Module (TPM)

The Trusted Platform Module (TPM) in the x3850 X5 is compliant with TPM 1.2. This integrated security chip performs cryptographic functions and stores private and public secure keys. It provides the hardware support for the Trusted Computing Group (TCG) specification.

Full disk encryption applications, such as the BitLocker Drive Encryption feature of Microsoft Windows Server 2008, can use this technology. The operating system uses it to protect the keys that encrypt the computer's operating system volume and provide integrity authentication for a trusted boot pathway (such as BIOS, boot sector, and others). A number of vendor full-disk encryption products also support the TPM chip.

The x3850 X5 uses the light path diagnostics panel's Remind button for the TPM Physical Presence function.

For details about this technology, see the Trusted Computing Group (TCG) TPM Main Specification at the following website:

http://www.trustedcomputinggroup.org/resources/tpm_main_specification

For more information about BitLocker and how TPM 1.2 fits into data security in a Windows environment, go to the following website:

http://technet.microsoft.com/en-us/windows/aa905062.aspx

3.13.6 Light path diagnostics

Light path diagnostics is a system of LEDs on various external and internal components of the server. When an error occurs, LEDs are lit throughout the server. By viewing the LEDs in a particular order, you can often identify the source of the error.

The server is designed so that LEDs remain lit when the server is connected to an ac power source but is not turned on, if the power supply is operating correctly. This feature helps you to isolate the problem when the operating system is shut down.

Figure 3-29 shows the light path diagnostics panel on the x3850 X5.

Figure 3-29 Light path diagnostics panel on the x3850 X5

Full details about the functionality and operation of the light path diagnostics in this system are in the *IBM System x3850 X5 and x3950 X5 Problem Determination and Service Guide,* which is available at the following website:

http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5083418

3.14 Power supplies and fans of the x3850 X5 and MAX5

This section describes the power and cooling features of the x3850 X5 server and the MAX5 memory expansion unit.

3.14.1 x3850 X5 power supplies and fans

The x3850 X5 includes the following power supplies:

- One or two dual-rated power supplies are standard (model-dependent; see 3.3, "Models" on page 64):
  - 1975 watts at 220 V ac input
  - 875 watts at 110 V ac input
- Hot-swappable and redundant at 220 V ac only, with two power supplies

The x3850 X5 includes the following fans to cool system components:

- Fan 1 = front left 120 mm (front access)
- Fan 2 = front right 120 mm (front access)
- Fan 3 = center right 60 mm (two fans) (top access)
- Fan 4 = back left 120 mm, part of power supply 2 (rear access)
- Fan 5 = back right 120 mm, part of power supply 1 (rear access)

The system is divided into the following cooling zones, which are shown in Figure 3-30 on page 113. Fans are redundant: two fans per cooling zone.

- Zone1 (left) = Fan 1, Fan 4, CPUs 1 and 2, memory cards 1 - 4, and power supply 2
- Zone2 (center) = Fan 2, Fan 5, CPUs 3 and 4, memory cards 5 - 8, and power supply 1
- Zone3 (right) = Fan 3, HDDs, SAS adapter, and PCIe slots 1 - 7

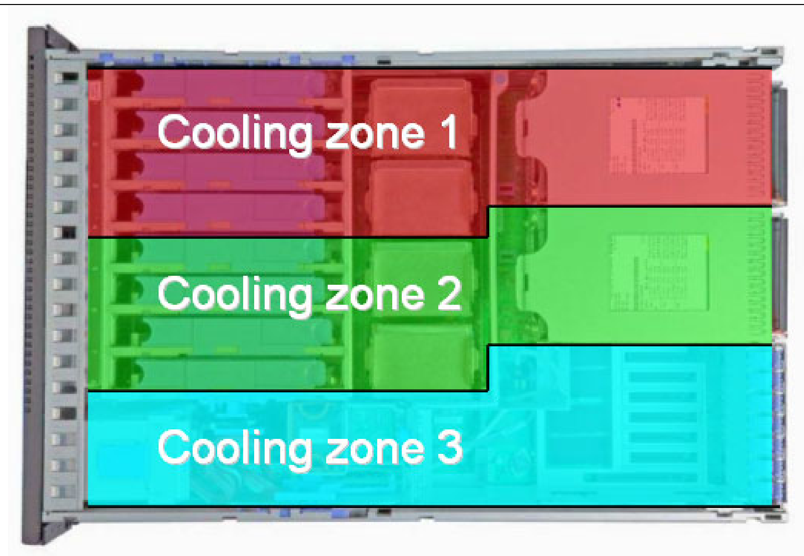


Figure 3-30 Cooling zones in the x3850 X5

Six strategically located hot-swap/redundant fans, combined with efficient airflow paths, provide highly effective system cooling for the eX5 systems. This technology is known as IBM Calibrated Vectored Cooling™ technology. The fans are arranged to cool three separate zones, with one pair of redundant fans per zone.

The fans automatically adjust speeds in response to changing thermal requirements, depending on the zone, redundancy, and internal temperatures. When the temperature inside the server increases, the fans speed up to maintain the proper ambient temperature. When the temperature returns to a normal operating level, the fans return to their default speed.

All x3850 X5 system fans are hot-swappable, except for Fan 3 in the bottom x3850 X5 of a 2-node complex, when QPI cables directly link the two servers.

3.14.2 MAX5 power supplies and fans

The MAX5 power subsystem consists of one or two hot-pluggable 675W power supplies. The power subsystem is designed for N+N (fully redundant) operation and hot-swap replacement.

Standard models of MAX5 have one power supply installed in power supply bay 1. For redundancy, install the second power supply that is listed in Table 3-35.

Table 3-35 Ordering information for the IBM MAX5 for System x

Part number	Feature code	Description
60Y0332	4782	IBM 675W HE Redundant Power Supply

A fan that is located inside each power supply cools the power modules.

MAX5 has five redundant hot-swap fans, which are all in one cooling zone. The IMM of the attached host controls the MAX5 fan speed, based on altitude and ambient temperature.

Fans also respond to certain conditions and come up to speed accordingly:

- If a fan fails, the remaining fans ramp up to full speed.
- As the internal temperature rises, all fans ramp to full speed.

3.15 Integrated virtualization

Selected models of the x3950 X5 include an installed USB 2.0 Flash Key that is preloaded with either VMware ESXi 4.0 or VMware ESXi 4.1, as shown in Figure 3-31. However, all models of x3850 X5 support several USB keys as options. For a complete list of USB virtualization options, see 2.9, “Integrated virtualization” on page 50.

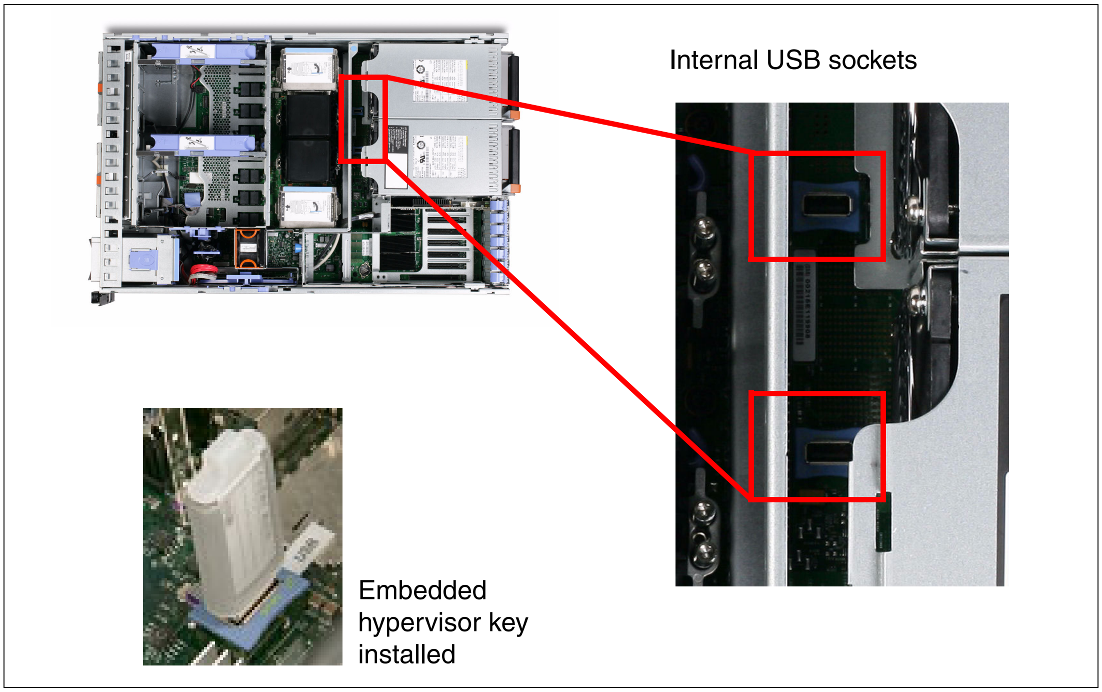


Figure 3-31 Location of internal USB ports for embedded hypervisor on the x3850 X5 and x3950 X5

3.16 Operating system support

The x3850 X5 supports the following operating systems:

- Microsoft Windows Server 2008 R2, Datacenter x64 Edition
- Microsoft Windows Server 2008 R2, Enterprise x64 Edition
- Microsoft Windows Server 2008 R2, Standard x64 Edition
- Microsoft Windows Server 2008 R2, Web x64 Edition
- Windows HPC Server 2008 R2
- SUSE LINUX Enterprise Server 10 for AMD64/EM64T
- SUSE LINUX Enterprise Server 11 for AMD64/EM64T
- SUSE LINUX Enterprise Server 10 with Xen for AMD64/EM64T
- SUSE LINUX Enterprise Server 11 with Xen for AMD64/EM64T
- Red Hat Enterprise Linux 5 Server x64 Edition
- Red Hat Enterprise Linux 5 Server with Xen x64 Edition
- VMware ESX 4.0 Update 1
- VMware ESX 4.1
- VMware ESXi 4.0 Update 1
- VMware ESXi 4.1

VMware ESXi 4.0 support: The use of MAX5 requires VMware ESXi 4.1. Version 4.0 is currently not supported with MAX5.

Because a short delay in qualification for several of these operating systems might exist, check the ServerProven Operating System support page for a current statement at the following website:

http://www.ibm.com/servers/eserver/serverproven/compat/us/nos/matrix.shtml

Table 3-36 summarizes the hardware maximums of the possible x3850 X5 configurations.

Table 3-36 Thread and memory maximums

	x3850 X5	**x3850 X5 with MAX5**	**Two-node x3850 X5**	**Two-node x3850 X5 with MAX5**
CPU threads	64	64	128	128
Memory capacity[a]	1 TB	1.5 TB	2 TB	3 TB

a. Using 16 GB DIMMs

3.17 Rack considerations

The x3850 X5 has the following physical specifications:

- Width: 440 mm (17.3 inches)
- Depth: 712 mm (28.0 inches)
- Height: 173 mm (6.8 inches) or 4 rack units (4U)
- Minimum configuration: 35.4 kg (78 lb)
- Maximum configuration: 49.9 kg (110 lb)

The x3850 X5 4U, rack-drawer models can be installed in a 19-inch rack cabinet that is designed for 26-inch deep devices, such as the NetBAY42 ER, NetBAY42 SR, NetBAY25 SR, or NetBAY11.

The 5U combination of MAX5 and x3850 X5 is mechanically joined and functions as a single unit. Adding the MAX5 to the x3850 X5 requires a change of the Electronic Industries Alliance (EIA) flange kit. The EIA flange kit, which ships standard with the 4U x3850 X5, must be removed and replaced with the 5U flange kit that ships standard with the MAX5.

If using a non-IBM rack, the cabinet must meet the EIA-310-D standards with a depth of at least 71.1 cm (28 in). Adequate space must be maintained from the slide assembly to the front door of the rack cabinet to allow sufficient space for the door to close and provide adequate air flow:

- 5 cm (2 inches) for the front bezel (approximate)
- 2.5 cm (1 inch) for air flow (approximate)

4

IBM System x3690 X5

The x3690 X5 servers are powerful 2-socket rack-mount servers with 4-core, 6-core, and 8-core Intel Xeon EX processors.

You can combine certain models of the x3690 X5 servers with the IBM MAX5 memory expansion for up to 1 TB of memory in a powerful 2-socket system. MAX5 is additionally available as an option for all of the other x3690 X5 models.

The x3690 X5 server belongs to the family of a new generation of Enterprise X-Architecture servers. The server delivers innovation with enhanced reliability and availability features to enable optimal performance for databases, enterprise applications, and virtualized environments.

This chapter contains the following topics:

4.1 Product features

The x3690 X5 is a 2U, 2-socket, scalable system that offers up to four times the memory capacity of current 2-socket servers. It has the following features:

- Up to two sockets for Intel Xeon 6500 or Xeon 7500 processors. Depending on the processor model, processors have four, six, or eight cores.
- Memory that is implemented using high-speed PC3-10600 and PC3-8500 DDR3 memory technology at up to 1066 MHz bus speed.
- Up to 32 dual inline memory modules (DIMMs) in the base system (16 on the system planar and 16 on an optional memory mezzanine), plus an additional 32 DIMMs with an optional 1U MAX5 memory expansion unit, for a total of 64 DIMM sockets.
- Uses Intel QuickPath Interconnect (QPI) technology for processor-to-processor connectivity and Intel and Scalable Memory Interconnect (SMI) processor-to-memory connectivity:
  - Intel QPI link topology at up to 6.4 Gbps with four QPI links per CPU
  - Intel SMI link topology at up to 6.4 Gbps with four SMI links per CPU
- Advanced networking capabilities with a Broadcom 5709 dual Gb Ethernet controller that is standard in all models.
- Emulex 10Gb dual-port Ethernet adapter that is standard on certain models, and optional on all other models.
- Power management savings.
- Memory ProteXion with Chipkill + 1b, memory mirroring, memory sparing, Intel scalable memory interconnect (SMI) Lane Failover, SMI Packet Retry, and SMI Clock Failover.
- Serial Attached SCSI (SAS)-based internal storage with RAID-0, RAID-1, or RAID-10 to maximize throughput and ease of installation; other RAID levels with optional RAID adapters.
- Up to 16 hot-swap 2.5-inch SAS hard disk drives (HDDs) and up to 8 TB of maximum internal storage or 16 hot-swap 2.5 inch solid-state drive (SSD) HDDs up to 800 Gb. The system includes (as standard) one HDD backplane that can hold four drives, and a second and third backplane are optional for an additional 12 drives. Adding more than two backplanes requires an additional SAS controller card.
- New eXFlash high-I/O operations per second (IOPS) solid-state storage technology for larger, faster databases. See 2.8, “IBM eXFlash” on page 47 for more information.
- A maximum of five PCIe 2.0 slots, depending on the option order for Peripheral Component Interconnect (PCI) riser card 1:
  - Four x8 PCIe slots with one x4 PCIe slot, using riser card option number 60Y0329.
  - One x16, two x8, and one x4 PCIe slots using riser card option number 60Y0331 for a 3/4 length adapter or option number 60Y0337 for a full-length adapter.
  - Two x8 PCIe slots and one x4 PCIe slot, with no PCI riser card 1 installed.
- Integrated Management Module (IMM) for enhanced systems management capabilities.
- 2U rack-optimized, tool-free chassis.
- Rear access hot-swap redundant power supplies for easy access.
- Top access hot-swap fan modules.

Figure 4-1 on page 119 shows the x3690 X5.

Figure 4-1 IBM System x3690 X5

Figure 4-1 shows the x3690 X5 server with 16 hot-swap 2.5-inch SAS disk drives installed.

The x3690 server has these physical specifications:

- Height: 86 mm (3.5 inches, 2U)
- Depth: 698 mm (27.4 inches)
- Width: 429 mm (16.8 inches)
- Maximum weight: 31.3 kg (69 lb) when fully configured

Each disk drive has an orange-colored bar. This color denotes that these disks are hot-swappable. The color coding that is used throughout the system is orange for hot-swappable and blue for non-hot-swappable. The only hot-swappable parts in this server are the HDDs, fans, and power supplies. All other parts require that the server is powered off before removing that component.

4.1.1 System components

Figure 4-2 shows the components on the front of the system.

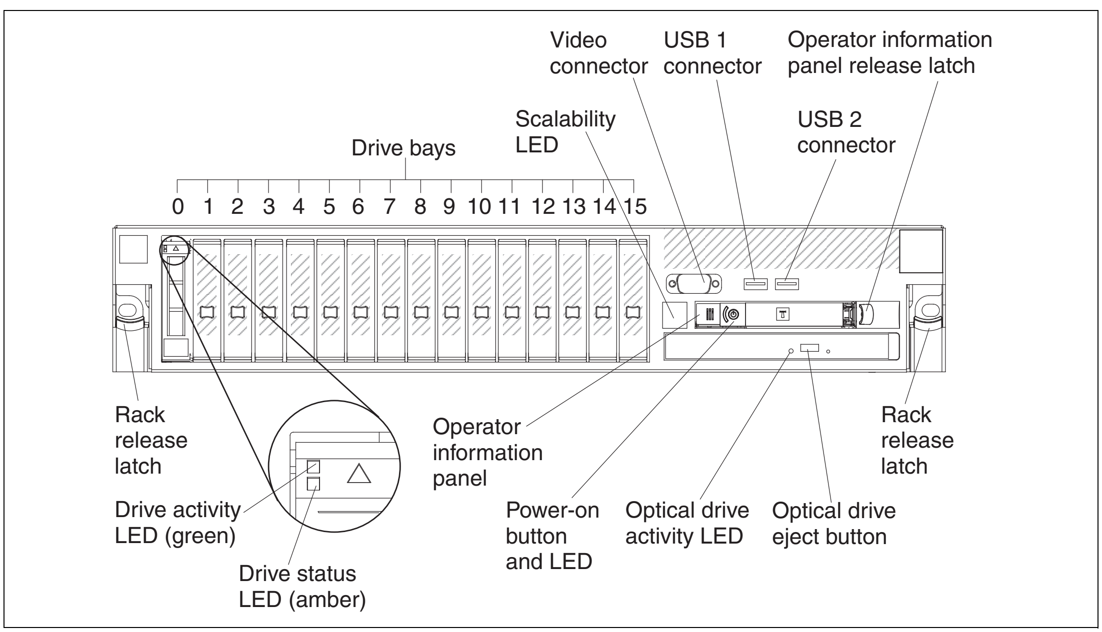


Figure 4-2 Front view of x3690 X5

Figure 4-3 on page 120 shows the rear of the system.

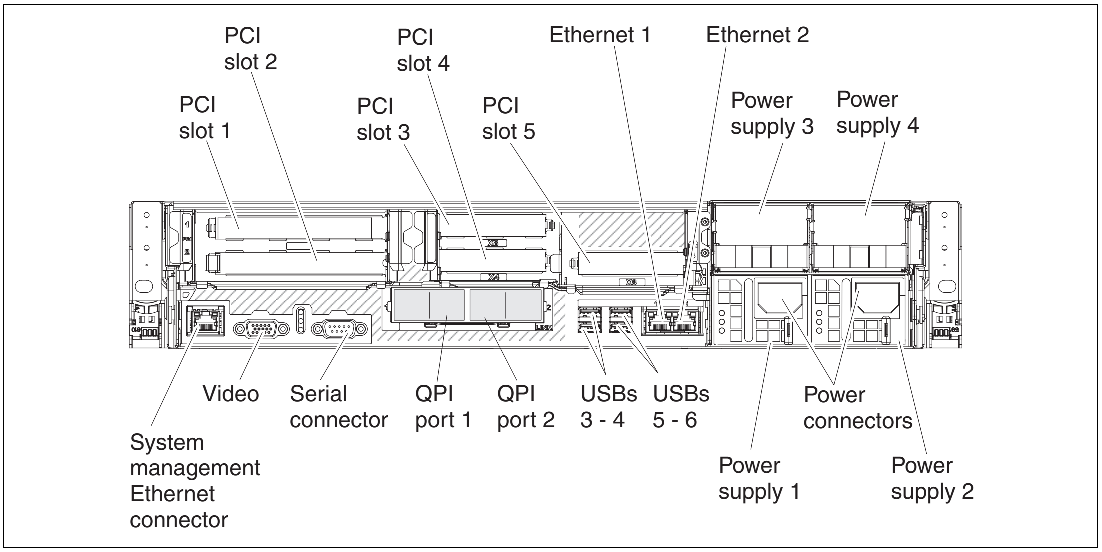


Figure 4-3 Rear view of x3690 X5

Figure 4-4 shows the system with the top cover removed.

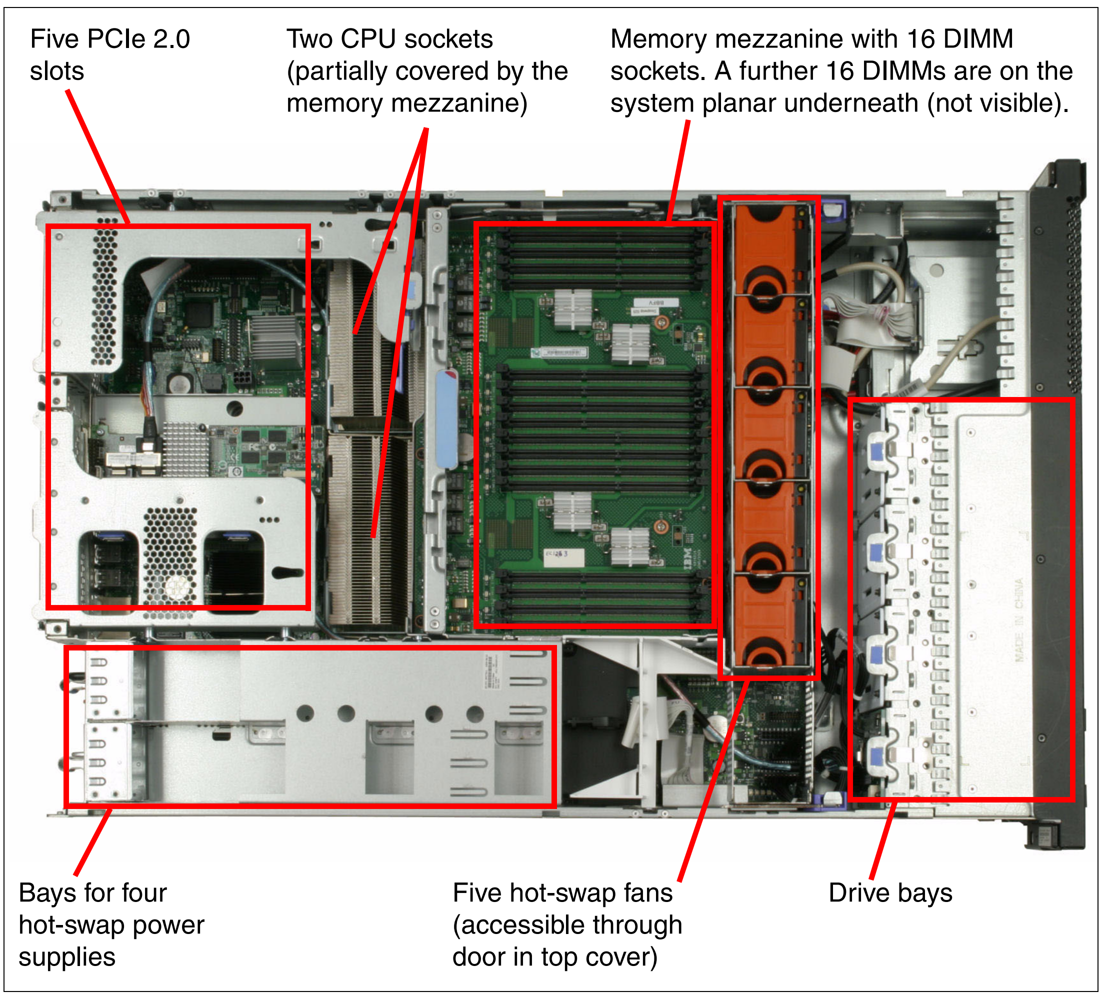


Figure 4-4 The x3690 X5 internals

important: The x3950 X5 top cover cannot be removed and the server remain powered on. If the top cover is removed, the server powers off immediately.

4.1.2 IBM MAX5 memory expansion unit

The IBM MAX5 for System x (MAX5) memory expansion unit has 32 DDR3 DIMM sockets, one or two 675-watt power supplies, and five 40 mm hot-swap speed-controlled fans. It provides added memory and multinode scaling support for host servers.

The MAX5 memory expansion unit is based on eX5, the next generation of Enterprise X-Architecture (EXA). The MAX5 expansion unit is designed for performance, expandability, and scalability. The fans and power supplies use hot-swap technology for easier replacement without requiring that you turn off the expansion module.

Figure 4-5 shows the x3690 X5 with the attached MAX5.

Figure 4-5 x3690 X5 with the attached MAX5 memory expansion unit

The MAX5 has the following specifications:

- IBM EXA5 chip set.
- Intel memory controller with eight memory ports (four DIMMs on each port).
- Intel QPI architecture technology to connect the MAX5 to the x3690 X5. There are two QPI links and each QPI link operates at up to 6.4 GT/s depending on the processors installed.
- Memory DIMMs:
  - Minimum: 2 DIMMs, 4 GB.
  - Maximum: 32 DIMM connectors (up to 512 GB of memory using 16 GB DIMMs).
  - Type of DIMMs: PC3-10600, 1067 MHz, error checking and correction (ECC), and DDR3 registered SDRAM DIMMs.
  - Support for 2 GB, 4 GB, 8 GB, and 16 GB DIMMs.

- Five hot-swap 40 mm fans.
- Power supply:
  - Hot-swap power supplies with built-in fans for redundancy support.
  - 675-watt (110 - 220 V ac auto-sensing).
  - One power supply standard; a second redundant power supply is optional.
- Light path diagnostics LEDs:
  - Board LED
  - Configuration LED
  - Fan LEDs
  - Link LED (for QPI and EXA5 links)
  - Locate LED
  - Memory LEDs
  - Power-on LED
  - Power supply LEDs
- Physical specifications:
  - Width: 483 mm (19.0 inches)
  - Depth: 724 mm (28.5 inches)
  - Height: 44 mm (1.73 inches) (1U rack unit)
  - Basic configuration: 12.8 kg (28.2 lb)
  - Maximum configuration: 15.4 kg (33.9 lb)

Tip: The MAX5 that is used with the x3690 X5 is the same as the MAX5 offered with the x3850 X5.

With the addition of the MAX5 memory expansion unit, the x3690 X5 gains an additional 32 DIMM sockets for a total of 64 DIMM sockets. Using 16 GB DIMMs, you can install a total of 1 TB of RAM.

All DIMM sockets in the MAX5 are accessible, regardless of whether one or two processors are installed in the x3690 X5.

Figure 4-6 shows the ports at the rear of the MAX5 memory expansion unit. When connecting the MAX5 to an x3690 X5, the QPI ports are used. The EXA ports are unused.

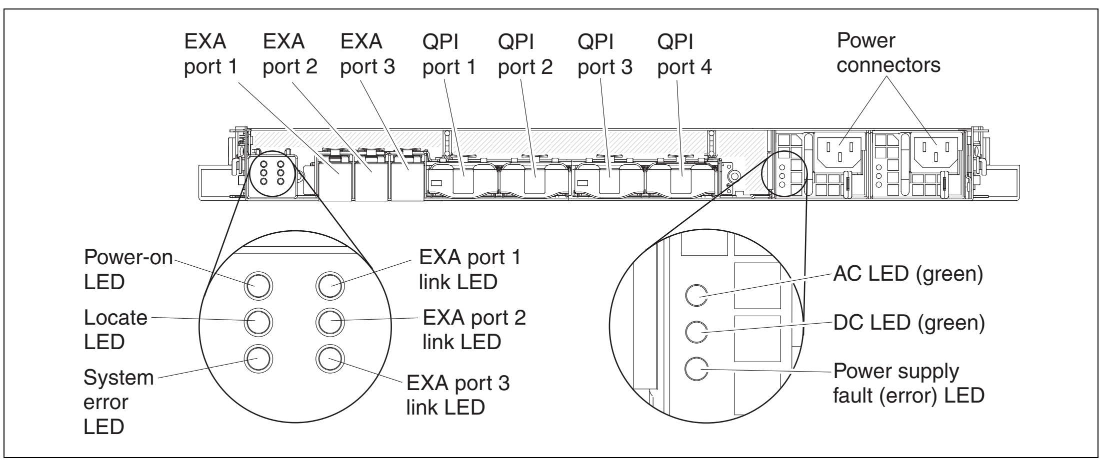


Figure 4-6 MAX5 connectors and LEDs

Figure 4-7 on page 123 shows the internals of the MAX5, including the IBM EXA chip that acts as the interface to the QPI links from the x3690 X5.

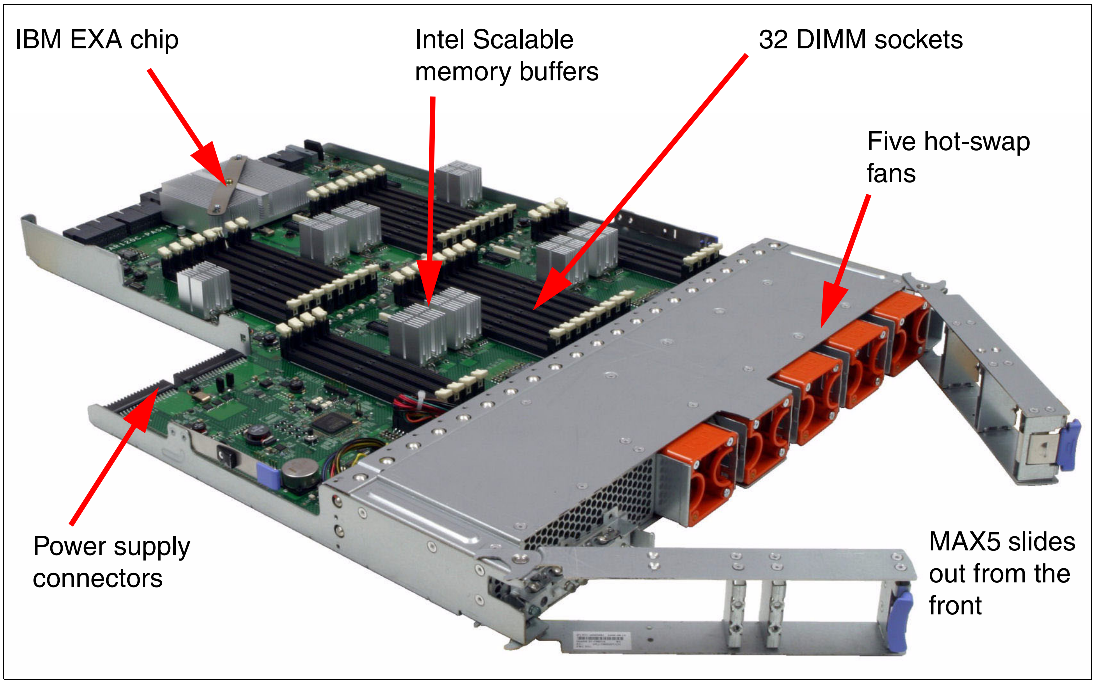


Figure 4-7 MAX5 memory expansion unit internals

For an in-depth look at the MAX5 offering, see 4.5, “MAX5” on page 127.

4.2 Target workloads

The x3690 X5 is an excellent choice for business applications that demand performance and memory. It provides maximum performance and memory for virtualization and database applications in a 2U package. It is a powerful and scalable system that allows certain workloads to migrate onto a 2-socket design, and it delivers enterprise computing in a dense package. Target workloads include the following items:

- Virtualization, consolidation, or virtual desktop

  The x3690 X5 with only two sockets can support as many virtual machines as older 4-socket servers because of having five times more memory than current 2-socket, x86-based servers. The result can lead to client savings on hardware and also on software licensing.

- Database

  The larger memory capacity of the x3690 X5 also offers leadership database performance. The x3690 X5 features the IBM eXFlash internal storage using SSDs to maximize the number of IOPS.

4.3 Models

In addition to the details in the tables in this chapter, each standard model has the following specifications:

- The servers have 16 DIMM sockets on the system planar. The additional 16-DIMM socket memory mezzanine (memory tray) is optional on most models and must be ordered separately. See 4.8, “Memory” on page 131 for details.
- The MAX5 is optional on certain models and standard on others.
- The optical drive is not standard and must be ordered separately if an optical drive is required. See 4.9.8, “Optical drives” on page 163 for details.
- As noted in the tables, most models have drive bays standard (*std*). However, disk drives are not standard and must be ordered separately. In the tables, *max* indicates maximum.

Base x3690 X5 models

Table 4-1 provides the standard models of the x3690 X5. The MAX5 memory expansion unit is standard on specific models, as indicated.

Table 4-1 x3690 X5 models

Model	Intel Xeon processors (two maximum)	Memory speed	MAX5	Standard memory[a]	Memory tray	ServeRAID M1015 standard	10Gb Ethernet standard[b]	Power supplies std/max	Drive bays std/max
7148-ARx	1x E7520 4C, 1.86 GHz, 95W	800 MHz	Opt	Server: 2x 4GB	Opt	Opt	Opt	1/4	None
7148-1Rx	1x E7520 4C, 1.86 GHz, 95W	800 MHz	Opt	Server: 2x 4GB	Opt	Std	Opt	1/4	4x 2.5”/16
7148-2Rx	1x E6540 6C, 2.00 GHz, 105W	1066 MHz	Opt	Server: 2x 4GB	Opt	Std	Opt	1/4	4x 2.5”/16
7148-3Rx	1x X6550 8C, 2.00 GHz, 130W	1066 MHz	Opt	Server: 2x 4GB	Opt	Std	Opt	1/4	4x 2.5”/16
7148-3Gx	1x X6550 8C, 2.00 GHz, 130W	1066 MHz	Opt	Server: 2x 4GB	Opt	Std	Std	1/4	4x 2.5”/16
7148-4Rx	1x X7560 8C, 2.26 GHz, 130W	1066 MHz	Opt	Server: 2x 4GB	Opt	Std	Opt	1/4	4x 2.5”/16
7148-3Sx	1x X7550 8C, 2.00GHz, 130W	1066 MHz	Std	Server: 2x 4GB MAX5: 2x 4GB	Opt	Std	Opt	Server: 2/4 MAX5: 1/2	4x 2.5”/16
7148-4Sx	1x X7560 8C, 2.26GHz, 130W	1066 MHz	Std	Server: 2x 4GB MAX5: 2x 4GB	Opt	Std	Opt	Server: 2/4 MAX5: 1/2	4x 2.5”/16

a. Up to 64 DIMM sockets: Each server has 16 DIMM sockets standard or 32 sockets with the addition of the internal memory tray (mezzanine). With the addition of the MAX5 memory expansion unit, 64 DIMM sockets total are available.

b. Emulex 10Gb Ethernet Adapter

Workload-optimized x3690 X5 models

Table 4-2 lists the workload-optimized models.

Model 3Dx is designed for database applications and uses SSDs for the best I/O performance. Backplane connections for sixteen 1.8-inch SSDs are standard and there is space for an additional 16 SSD. You must order the actual SSDs separately. No SAS controllers are standard, which allows you to select from the available cards, as described in 4.9, “Storage” on page 145. The MAX5 is optional on this model.

Model 2Dx is designed for virtualization applications and includes VMware ESXi 4.1 on an integrated USB memory key. Backplane connections for four 2.5-inch SAS drives are standard and there is space for an additional twelve 2.5-inch disk drives. You must order the actual drives separately. See 4.9, “Storage” on page 145 for details.

Table 4-2 x3690 X5 workload-optimized models

Model	Intel Xeon processors (two maximum)	Memory speed	MAX5	Standard memory[a]	Memory tray	ServeRAID M1015 std	10Gb Ethernet standard[b]	Power supplies std/max	Drive bays std/max
Database workload-optimized models									
7148-3Dx	2x X6550 8C, 2.00 GHz, 130W	1066 MHz	Opt	Server: 4x 4 GB	Std	Opt	Opt	Server: 4/4	16x 1.8”/24
Virtualization workload-optimized models									
7148-2Dx	2x E6540 6C, 2.00 GHz, 105W	1066 MHz	Std	Server: 32x 4GB MAX5: 32x 4GB	Std	Opt	Std	Server: 4/4 MAX5: 2/2	4x 2.5”/16

a. Up to 64 DIMM sockets: Each server has 16 DIMM sockets standard or 32 sockets with the addition of the internal memory tray (mezzanine). With the addition of the MAX5 memory expansion unit, 64 DIMM sockets total are available.

b. Emulex 10Gb Ethernet Adapter.

4.4 System architecture

Figure 4-8 shows the block diagram of the x3690 X5.

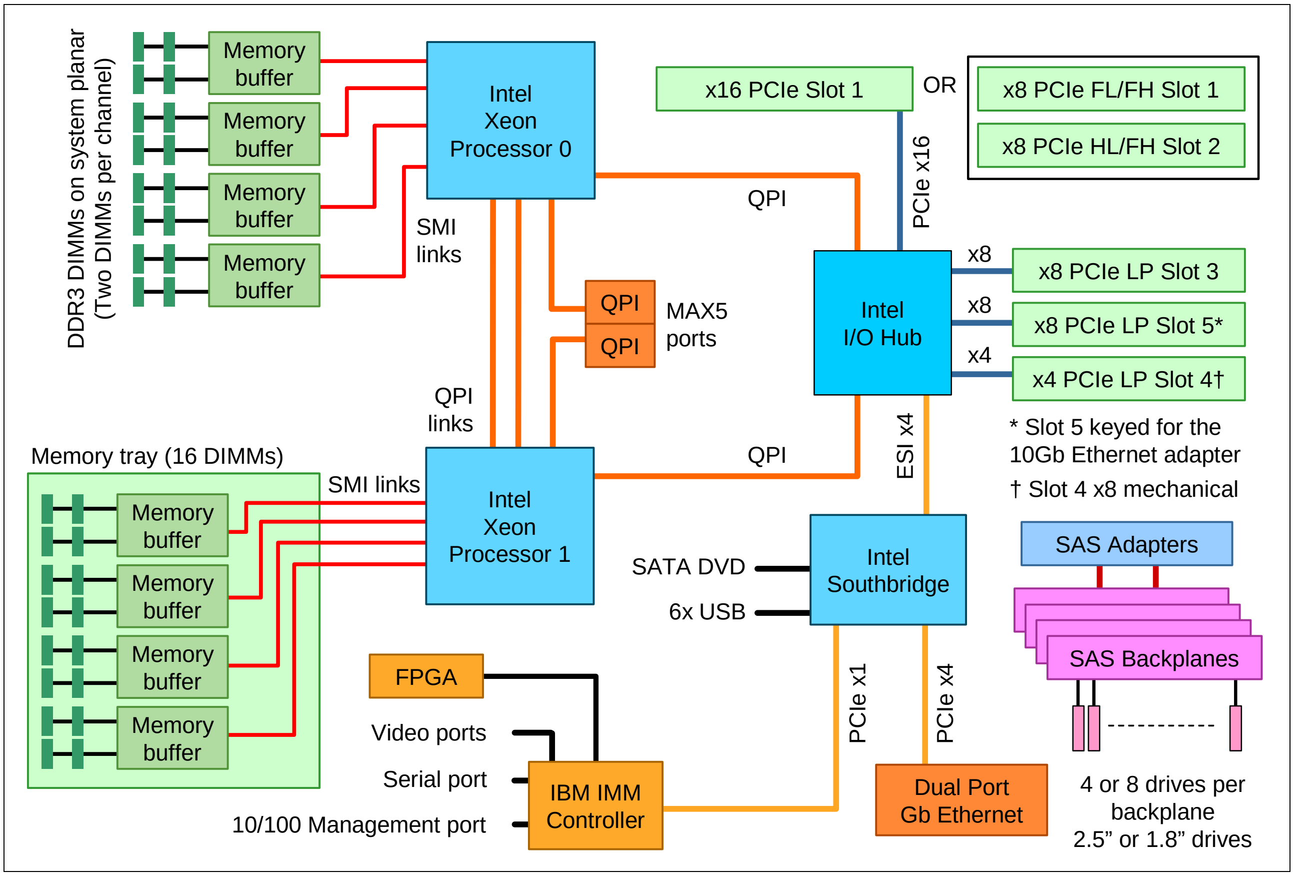


Figure 4-8 x3690 X5 block diagram

Figure 4-9 shows the block diagram of the MAX5. The MAX5 is connected to the x3690 X5 using two cables connecting the two QPI ports on the server to two of the QPI ports on the MAX5. The EXA ports and other two QPI are unused in this configuration.

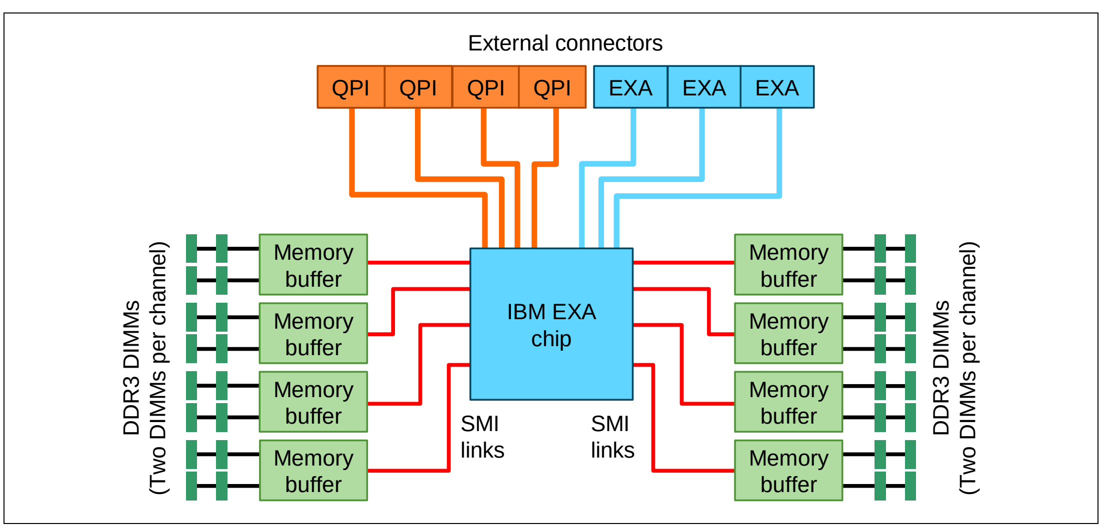


Figure 4-9 MAX5 block diagram

4.5 MAX5

As introduced in 4.1.2, “IBM MAX5 memory expansion unit” on page 121, the MAX5 memory expansion drawer is available for the x3690 X5. Certain standard models include the MAX5, as described in 4.3, “Models” on page 124, and the MAX5 can also be ordered separately, as listed in Table 4-3.

Table 4-3 Ordering information for the IBM MAX5 for System x

Part number	Feature code	Description
59Y6265	4199	IBM MAX5 for System x
60Y0332	4782	IBM 675W HE Redundant Power Supply
59Y6269	7481	IBM MAX5 to x3690 X5 Cable Kit (two cables)

The eX5 chip set in the MAX5 is an IBM unique design that attaches to the QPI links as a node controller, giving it direct access to all CPU bus transactions. It increases the number of DIMMs supported in a system by a total of 32, and also adds another 16 channels of memory bandwidth, boosting overall throughput. The MAX5 adds additional memory performance.

The eX5 chip connects directly through QPI links to both CPUs in the x3690 X5, and it maintains a directory of each CPU’s last-level cache. This directory allows the eX5 chip to respond to memory requests prior to the end of a broadcast snoop cycle, thereby improving performance. For more information about eX5 technology, see 2.1, “eX5 chip set” on page 16.

Figure 4-10 shows a diagram of the MAX5.

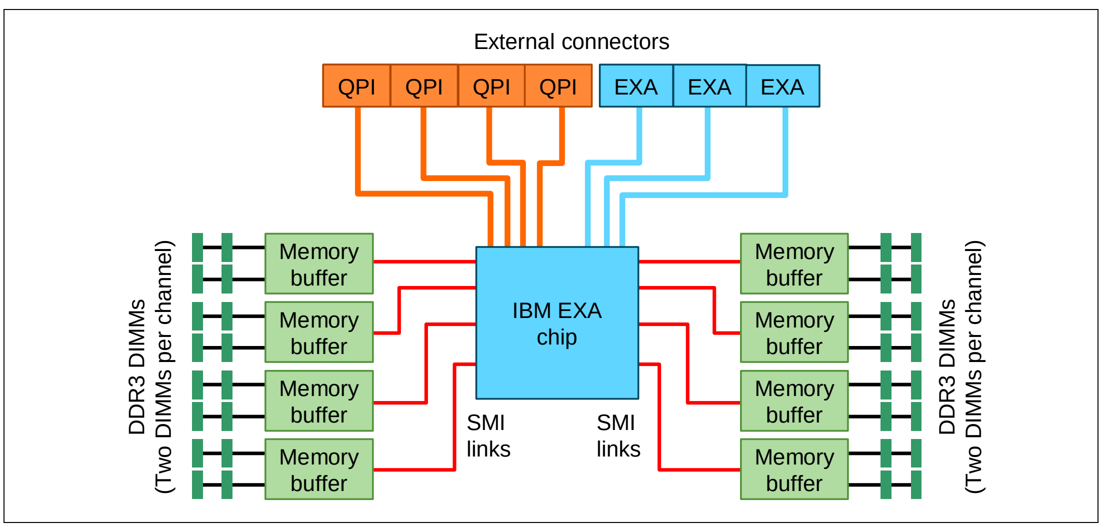


Figure 4-10 MAX5 block diagram

The MAX5 is connected to the x3690 X5 using two cables, connecting the QPI ports on the server to two of the four QPI ports on the MAX5. The other two QPI ports of the MAX5 are unused. The EXA ports are for future scaling capabilities.

Figure 4-11 shows architecturally how a single-node x3690 X5 is connected to a MAX5.

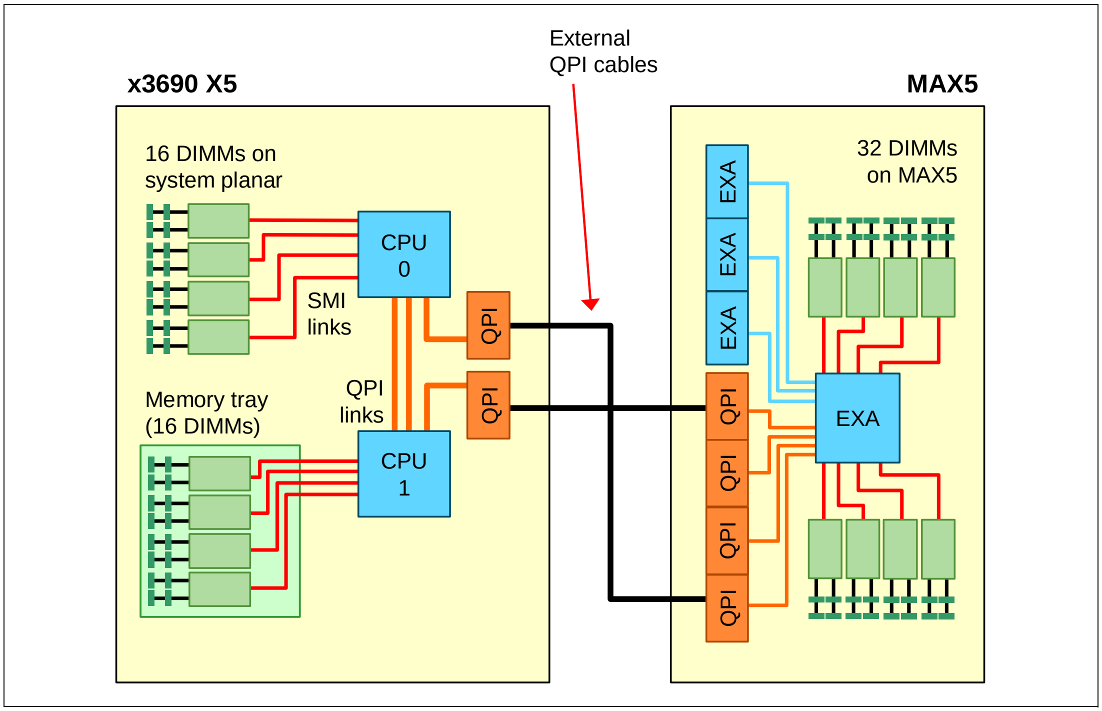


Figure 4-11 Connectivity of the x3690 X5 with a MAX5 memory expansion unit

As shown in Figure 4-11, the x3690 X5 attaches to the MAX5 using QPI links; you can see that the eX5 chip set in the MAX5 simultaneously connects to both CPUs in the server.

One benefit of this connectivity is that the MAX5 is able to store a copy of the contents of the last-level cache of all the CPUs in the server. Therefore, when a CPU requests content stored in the cache of another CPU, the MAX5 not only has that same data stored in its own cache, it is able to return the acknowledgement of the snoop *and* the data to the requesting CPU in the same transaction. For more information about QPI links and snooping, see 2.2.4, “QuickPath Interconnect (QPI)” on page 18.

Tip: The Xeon E6510 processor does not support the use of the MAX5.

Connectivity of the MAX5 to the x3690 X5 is described in 4.6, “Scalability” on page 128.

For memory configuration information, see 4.8.3, “MAX5 memory” on page 136. For a description of the power and fans, see 4.12, “Power supplies” on page 173.

4.6 Scalability

In this section, we describe how the x3690 X5 can be expanded to increase the number of memory DIMMs.

The x3690 X5 supports the following configurations:

- A single x3690 X5 server with two processor sockets. This configuration is sometimes referred to as a *single-node server*.
- A single x3690 X5 server with a single MAX5 memory expansion unit attached. This configuration is sometimes referred to as a *memory-expanded server*.

Two-node configurations: The 2-node configurations (with and without MAX5) are not supported.

The MAX5 memory expansion unit permits the x3690 X5 to scale to an additional 32 DDR3 DIMM sockets.

Connecting the single-node x3690 X5 with the MAX5 memory expansion unit uses two QPI cables, part number 59Y6269, as listed in Table 4-4. Figure 4-12 shows the connectivity.

Number of processors: The MAX5 is supported with either one or two processors installed in the x3690 X5. However, the recommendation is to have two processors installed and memory installed in every DIMM socket in the server to maximize performance.

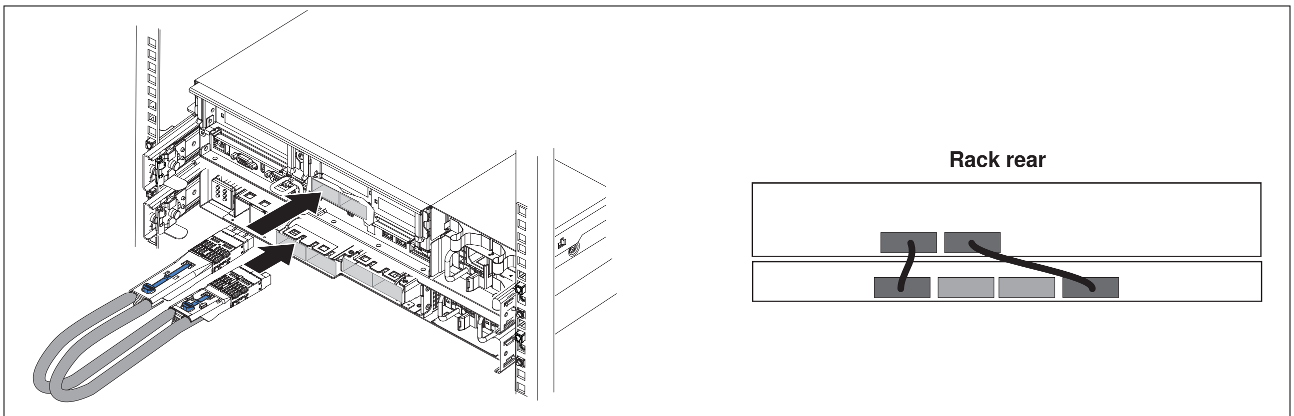


Figure 4-12 Connecting the MAX5 to a single-node x3690 X5

Connecting the MAX5 to a single-node x3690 X5 requires one IBM MAX5 to x3690 X5 Cable Kit, which consists of two QPI cables. See Table 4-4.

Table 4-4 Ordering information for the IBM MAX5 to x3690 X5 Cable Kit

Part number	Feature code	Description
59Y6269	7481	IBM MAX5 to x3690 X5 Cable Kit (two cables)

Note: There is no setup on the Integrated Management Module (IMM) for scalability. Ensure that you have updated all firmware before attaching the MAX5, especially the Field Programmable Gate Array (FPGA). For more information about updating firmware, see 9.10, “Firmware update tools and methods” on page 509.

4.7 Processor options

Several Intel Xeon 6500 and Xeon 7500 processor options are available for the x3690 X5, as listed in Table 4-5.

Table 4-5 x3690 X5 processor options

Part number	Feature code	Intel model # of cores (C)	Core speed	L3 cache	QPI Link GT/s	Memory speed	Power TDP[a]	HT[b]	TB[c]	MAX5
Advanced processors (X)										
60Y0311	4469	Xeon X7560 8C	2.26 GHz	24 MB	6.4 GT/s	1066 MHz	130W	Yes	Yes	Yes
60Y0313	4471	Xeon X7550 8C	2.00 GHz	18 MB	6.4 GT/s	1066 MHz	130W	Yes	Yes	Yes
60Y0321	4479	Xeon X7542 6C	2.66 GHz	18 MB	5.86 GT/s	977 MHz	130W	No	Yes	Yes
60Y0319	4477	Xeon X6550 8C	2.00 GHz	18 MB	6.4 GT/s	1066 MHz	130W	Yes	Yes	Yes
Standard processors (E)										
60Y0315	4473	Xeon E7540 6C	2.00 GHz	18 MB	6.4 GT/s	1066 MHz	105W	Yes	Yes	Yes
60Y0316	4474	Xeon E7530 6C	1.86 GHz	12 MB	5.86 GT/s	977 MHz	105W	Yes	Yes	Yes
60Y0317	4475	Xeon E7520 4C	1.86 GHz	18 MB	4.8 GT/s	800 MHz	95W	Yes	No	Yes
60Y0320	4478	Xeon E6540 6C	2.00 GHz	18 MB	6.4 GT/s	1066 MHz	105W	Yes	Yes	Yes
60Y0318	4476	Xeon E6510 4C	1.73 GHz	12 MB	4.8 GT/s	800 MHz	105W	Yes	No	No
Low-power processors (L)										
60Y0312	4470	Xeon L7555 8C	1.86 GHz	24 MB	5.86 GT/s	977 MHz	95W	Yes	Yes	Yes
60Y0314	4472	Xeon L7545 6C	1.86 GHz	18 MB	5.86 GT/s	977 MHz	95W	Yes	Yes	Yes

a. Thermal design power
b. Intel Hyper-Threading Technology
c. Intel Turbo Boost Technology

Clarification:

- The Xeon E6510 does not support the use of the MAX5 memory expansion unit.
- The x3690 X5 announcement letter incorrectly reports that the L3 cache of the X7542 is 12 MB.
- Also, the announcement letter incorrectly states that the E6540 does not have Turbo Boost mode.

See 2.2, “Intel Xeon 6500 and 7500 family processors” on page 16 for an in-depth description of the Intel Xeon 6500/7500 processor family and features.

Most processors support Intel Turbo Boost Technology with a couple of exceptions, as listed in Table 4-5. When a CPU operates beneath its thermal and electrical limits, Turbo Boost dynamically increases the processor’s clock frequency by 133 MHz on short and regular intervals until an upper limit is reached. See 2.2.3, “Turbo Boost Technology” on page 18 for more information.

With the exception of the X7542, all CPUs that are listed support Intel Hyper-Threading Technology. Hyper-Threading Technology (HT) is an Intel technology that is used to improve the parallelization of workloads. When Hyper-Threading is engaged in the BIOS, for each processor core that is physically present, the operating system addresses two. For more information, see 2.2.2, “Hyper-Threading Technology” on page 17.

All CPU options include a heat-sink.

The x3690 X5 models include one CPU standard. All five PCIe slots are usable, even with only one processor installed, as shown in Figure 4-8 on page 126.

The second CPU is required to access the memory in the memory mezzanine (if the memory mezzanine is installed). The second CPU can be installed without the memory mezzanine, but its only access to memory is through the primary CPU. For optimal performance, if two CPUs are installed, install a memory mezzanine also.

Follow these population guidelines:

- Each CPU requires a minimum of two DIMMs to operate. If the memory mezzanine is installed, it needs a minimum of two DIMMs installed.
- Both processors must be identical.
- Consider the X7542 processor for CPU frequency-dependent workloads, because it has the highest core frequency of the available processor models.
- The MAX5 is supported with either one or two processors installed in the x3690 X5, although the recommendation is to have two processors installed and memory installed in every DIMM socket in the server to maximize performance.
- If high processing capacity is not required for your application but high memory bandwidth is required, consider using two processors with fewer cores or a lower core frequency rather than two processors with more cores or a higher core frequency. Having four processors can enable all memory channels and can maximize memory bandwidth. We describe this technique in 4.8, “Memory” on page 131.

4.8 Memory

The x3690 X5 offers up to 32 DIMM sockets that are internal to the server chassis, plus an additional 32 DIMM sockets in the MAX5 memory expansion unit.

This section covers the following topics:

Implement the memory DIMMs that are internal to the x3690 X5 chassis:

- 16 DIMM sockets on the system planar
- 16 DIMM sockets in an optional memory mezzanine

> **Tip:** The memory mezzanine is referred to in the announcement letter as the *memory expansion card*. It is referred to as the *memory tray*, in the *Installation and User’s Guide - IBM System x3690 X5*.

The memory mezzanine is an optional component and orderable as listed in Table 4-6.

Table 4-6 x3690 X5 memory mezzanine option part number

Option	Feature code	Description
60Y0323	9278	IBM x3690 X5 16-DIMM Internal Memory Expansion

Figure 4-13 shows the memory mezzanine and DIMMs.

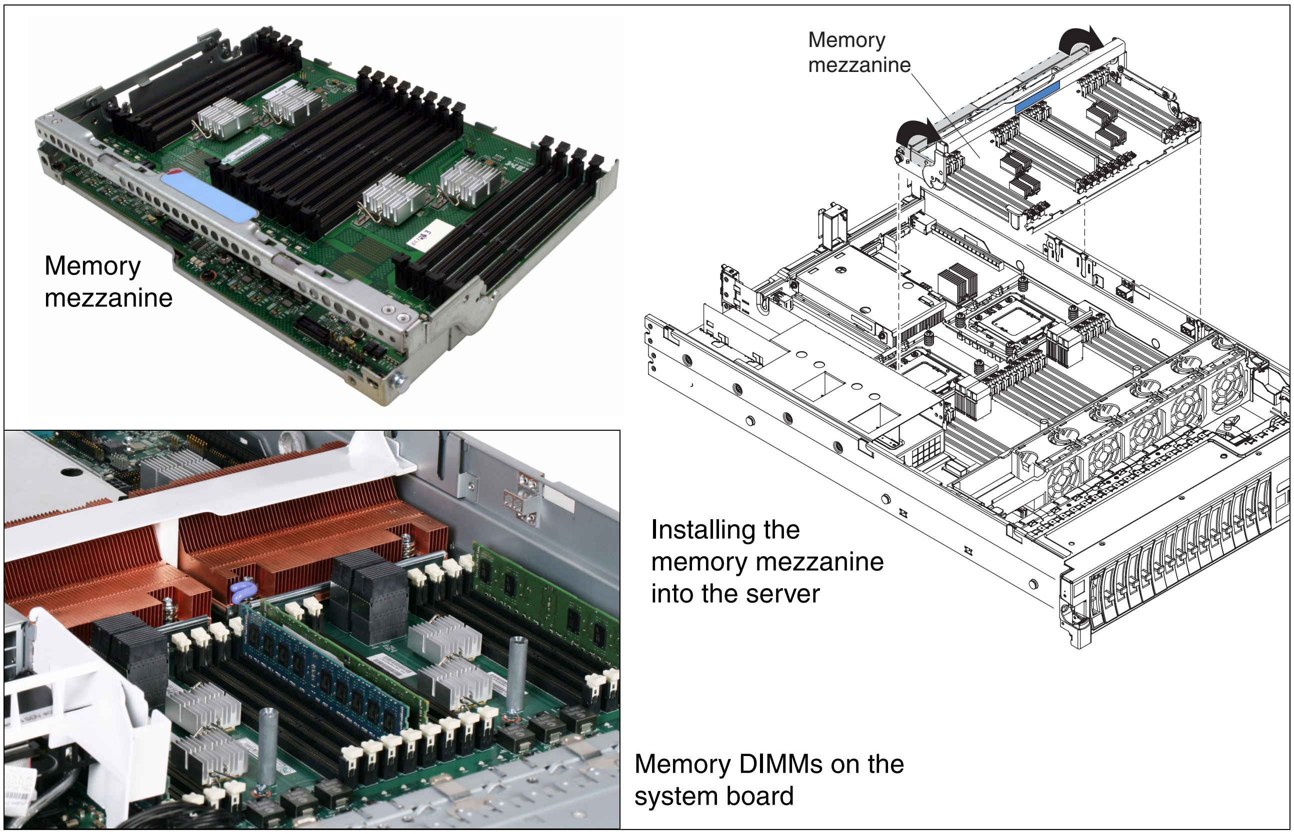


Figure 4-13 Location of the memory DIMMs

With the Intel Xeon 6500 and 7500 processors, the memory controller is integrated into the processor, as shown in the architecture block diagram in Figure 4-8 on page 126:

- Processor 0 connects directly to the memory buffers and memory DIMM sockets on the system planar.
- Processor 1 connects directly to the memory buffers and memory in the memory mezzanine.

If you plan to install the memory mezzanine, you are required to also install the second processor.

The x3690 X5 uses the Intel scalable memory buffer to provide DDR3 SDRAM memory functions. The memory buffers connect to the memory controller in each processor through Intel Scalable Memory Interconnect links. Each memory buffer has two memory channels, and the DIMM sockets are connected to the memory buffer with two DIMMs per memory channel (2 DPC).

The memory uses DDR3 technology and operates at memory speeds of 800, 978, and 1066 MHz. The memory speed is dictated by the memory speed of the processor (see Table 4-5 on page 130). For more information about how this is calculated, see 2.3.1, “Memory speed” on page 22.

4.8.1 Memory DIMM options

Table 4-7 shows the available memory options that are supported in the x3690 X5 server.

Table 4-7 x3690 X5/MAX5 Supported DIMMs

Part number	x3690 X5 feature code	Memory	Supported in MAX5	Memory speed[a]	Ranks
44T1592	1712	2 GB (1x 2GB) 1Rx8, 2Gbit PC3-10600R DDR3-1333	Yes (fc 2429)	1333 MHz[b]	Single x8
44T1599	1713	4 GB (1x 4GB), 2Rx8, 2Gbit PC3-10600R DDR3-1333	Yes (fc 2431)	1333 MHz	Dual x8
46C7448	1701	4 GB (1x 4GB), 4Rx8, 1Gbit PC3-8500 DDR3-1066	No	1066 MHz	Quad x8
46C7482	1706	8 GB (1x 8GB), 4Rx8, 2Gbit PC3-8500 DDR3-1066	Yes (fc 2432)	1066 MHz	Quad x8
46C7483	1707	16 GB (1x 16GB), 4Rx4, 2Gbit PC3-8500 DDR3-1066	Yes[c] (fc 2433)	1066 MHz	Quad x4

a. Memory speed is also controlled by the memory bus speed, as specified by the selected processor model. The actual memory bus speed is the lower of both the processor memory bus speed and the DIMM memory bus speed.
b. Although 1333 MHz memory DIMMs are supported in the x3690 X5, the memory DIMMs run at a maximum speed of 1066 MHz.
c. The 16 GB memory option is only supported in the MAX5 when it is the only type of memory used in the MAX5. You cannot use any other memory options in the MAX5 if this option is installed in the MAX5. This DIMM also supports redundant bit steering (RBS) when used in the MAX5, as described in “Redundant bit steering” on page 29.

Details:

- Only certain memory options supported in the x3690 X5 are also supported in the MAX5, as indicated in Table 4-7. When ordering DIMMs for the MAX5 with feature codes (configure-to-order (CTO) clients only), use the special MAX5 feature codes 24*xx* as listed. See “MAX5 memory options” on page 137 for more information.
- Memory options must be installed in matched pairs. Single options cannot be installed; therefore, the options listed must be ordered in a quantity of two.
- The maximum memory speed that is supported by Xeon 7500 and 6500 (Nehalem-EX) processors is 1066 MHz. DIMMs will not run at 1333 MHz.
- Mixing DIMMs is supported by all DIMMs, except 16 GB DIMMs, and operates at the speed of the slowest installed DIMM. We do not recommend mixing DIMMs for performance reasons.

4.8.2 x3690 X5 memory population order

Memory DIMM installation is key to maximizing system performance. In this section, we specify how to install DIMMs.

Figure 4-14 on page 134 shows the slot numbering for DIMM installation.

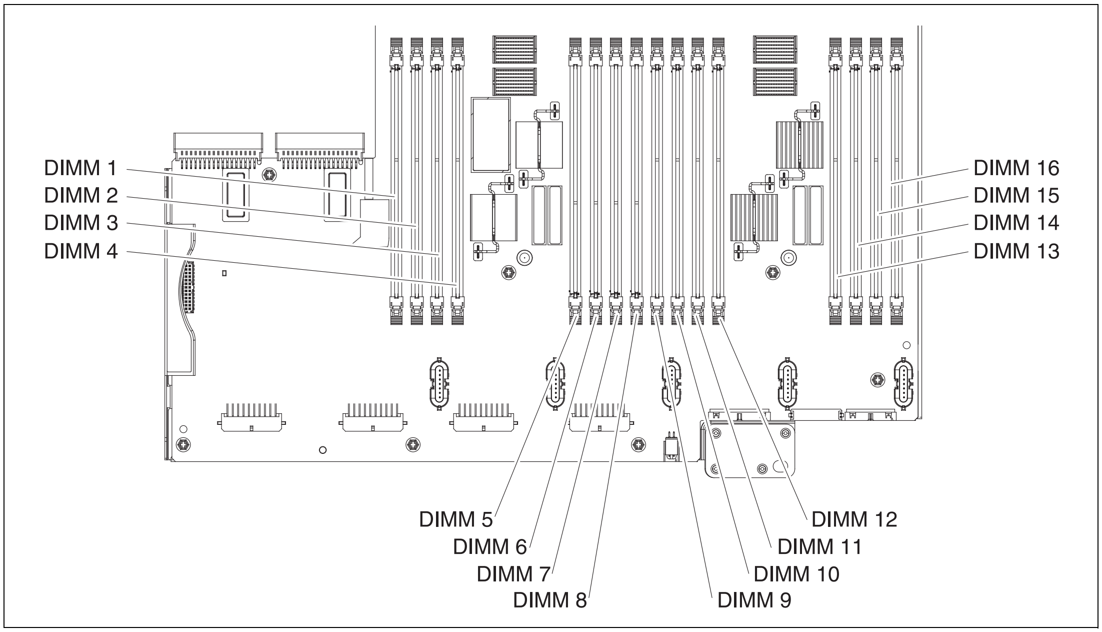


Figure 4-14 x3690 X5 planar showing memory DIMM locations

One or two processors without the memory mezzanine

In this configuration, all of the system's memory directly attaches to processor 1. If a second processor is installed, it accesses main memory through processor 1, resulting in a performance degradation to processor 2.

Tip: For performance reasons, install and populate the memory mezzanine if you install the second processor.

When the memory mezzanine is not installed, install the DIMMs in the order that is listed in Table 4-8 on page 135. Only certain DIMM combinations result in Hemisphere Mode being enabled. Hemisphere Mode improves memory performance, as described in 2.3.5, "Hemisphere Mode" on page 26.

Table 4-8 One or two processor DIMM installation when the memory mezzanine is not installed

Number of processors	Number of DIMMs	Hemisphere Mode[a]	Memory buffer				Memory buffer				Memory buffer				Memory buffer			
			DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16
1 or 2	2	N	x							x								
1 or 2	4	Y	x							x	x							x
1 or 2	6	N	x		x			x		x	x							x
1 or 2	8	Y	x		x			x		x	x		x			x		x
1 or 2	10	N	x	x	x			x	x	x	x		x			x		x
1 or 2	12	Y	x	x	x			x	x	x	x	x	x			x	x	x
1 or 2	14	N	x	x	x	x	x	x	x	x	x	x	x			x	x	x
1 or 2	16	Y	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

a. For more information about Hemisphere Mode and its importance, see 2.3.5, "Hemisphere Mode" on page 26.

Important: If using two processors with no mezzanine board, memory is *not* in a nonuniform memory access (NUMA)-compliant state, which causes great performance degradation. See 2.3.4, "Nonuniform memory architecture (NUMA)" on page 26 for details.

Important VMware ESX considerations: When installing and running VMware ESX on this server, the operating system might fail to install or boot with the following error message when the server memory configuration is not NUMA-compliant:

"`NUMA node 1 has no memory`"

There are only three possible configurations to support VMware:

- One processor is installed and no mezzanine board is installed.
- Two processors are installed and matching memory is installed on both the system board and the mezzanine board.
- Two processors are installed, no internal memory is installed, and the memory installed in an attached MAX5 memory expansion is configured as non-pooled memory.

Two processors with memory mezzanine installed

With two processors installed in the system, memory is evenly distributed between both processors, which maximizes system performance.

Install the memory in the order that is listed in Table 4-9 on page 136. You are required to install a minimum of four DIMMs.

Figure 4-15 on page 136 shows the DIMM numbering on the memory mezzanine.

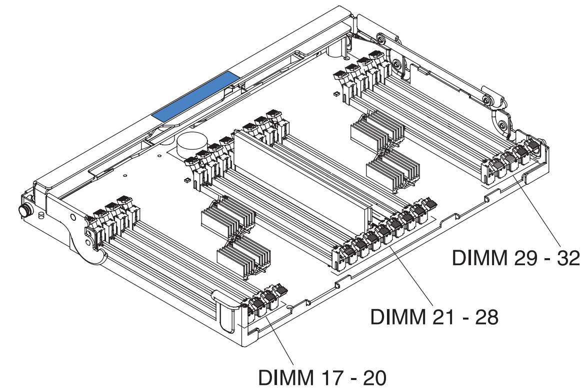


Figure 4-15 Memory mezzanine tray

Table 4-9 NUMA-compliant DIMM installation: Two processors and the memory mezzanine installed

Number of DIMMs	Hemisphere Mode[a]	Processor 1 (planar DIMMs)																Processor 2 (mezzanine DIMMs)															
		Buffer				Buffer				Buffer				Buffer				Buffer				Buffer				Buffer				Buffer			
		DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16	DIMM 17	DIMM 18	DIMM 19	DIMM 20	DIMM 21	DIMM 22	DIMM 23	DIMM 24	DIMM 25	DIMM 26	DIMM 27	DIMM 28	DIMM 29	DIMM 30	DIMM 31	DIMM 32
4	N	x							x									x							x								
8	Y	x							x	x							x	x							x	x							x
12	N	x		x			x		x	x							x	x		x			x		x	x							x
16	Y	x		x			x		x	x		x			x		x	x		x			x		x	x		x			x		x
20	N	x	x	x			x	x	x	x		x			x		x	x	x	x			x	x	x	x		x			x		x
24	Y	x	x	x			x	x	x	x	x	x			x	x	x	x	x	x			x	x	x	x	x	x			x	x	x
28	N	x	x	x	x	x	x	x	x	x	x	x			x	x	x	x	x	x	x	x	x	x	x	x	x	x			x	x	x
32	Y	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

a. For more information about Hemisphere Mode and its importance, see 2.3.5, "Hemisphere Mode" on page 26.

> **Tip:** Table 4-9 lists only memory configurations that are considered best practice in obtaining optimal memory and processor performance. For a full list of supported memory configurations, see the *IBM System x3690 X5 Installation and User Guide* or the *IBM System x3690 X5 Problem Determination and Service Guide*. You can obtain both of these documents at the following website:
>
> `http://www.ibm.com/support`

4.8.3 MAX5 memory

The MAX5 memory expansion unit has 32 DIMM sockets. It is designed to augment the memory that is installed in the attached x3690 X5 server.

MAX5 memory options

Table 4-10 shows the available memory options that are supported in the MAX5 memory expansion unit. These options are a subset of the options that are supported in the x3690 X5 because the MAX5 requires that all DIMMs use identical DRAM technology, which is either 2 Gbit x8 or 2 Gbit x4 (but not both at the same time).

> **Memory options:** The memory options listed here are also supported in the x3690 X5 (but under separate feature codes for CTO clients). Table 4-7 on page 133 lists additional memory options, which are also supported in the x3690 X5 server but not in the MAX5.

Table 4-10 DIMMs supported in the MAX5

Part number	MAX5 feature code	Memory	Supported in MAX5	Memory speed[a]	Ranks
44T1592	2429	2 GB (1x 2GB) 1Rx8, 2Gbit PC3-10600R DDR3-1333	Yes	1333 MHz[b]	Single x8
44T1599	2431	4 GB (1x 4GB), 2Rx8, 2Gbit PC3-10600R DDR3-1333	Yes	1333 MHz	Dual x8
46C7482	2432	8 GB (1x 8GB), 4Rx8, 2Gbit PC3-8500 DDR3-1066	Yes	1066 MHz	Quad x8
46C7483	2433	16 GB (1x 16GB), 4Rx4, 2Gbit PC3-8500 DDR3-1066	Yes[c d]	1066 MHz	Quad x4

a. Memory speed is also controlled by the memory bus speed, as specified by the processor model selected. The actual memory bus speed is the lower of both the processor memory bus speed and the DIMM memory bus speed.
b. Although 1333 MHz memory DIMMs are supported in the x3690 X5, the memory DIMMs run at a maximum speed of 1066 MHz.
c. The 16 GB memory option is only supported in the MAX5 when it is the only type of memory used in the MAX5. No other memory options can be used in the MAX5 if this option is installed in the MAX5.
d. This DIMM supports redundant bit steering (RBS), as described in "Redundant bit steering" on page 29.

> **Use of the 16 GB memory option:** The 16 GB memory option, part number 46C7483, is supported in the MAX5 only when it is the only type of memory used in the MAX5. No other memory options can be used in the MAX5 if this option is installed in the MAX5.

> **Redundant bit steering:** Redundant bit steering (RBS) is not supported on the x3690 X5 itself, because the integrated memory controller of the Intel Xeon 7500 processor does not support the feature. See "Redundant bit steering" on page 29 for details.
>
> The MAX5 memory expansion unit supports RBS, but only with x4 memory and not x8 memory. As shown in Table 4-10, the 16 GB DIMM, part number 46C7483, uses x4 DRAM technology. RBS is automatically enabled in the MAX5 memory port, if all DIMMs that are installed to that memory port are x4 DIMMs.

MAX5 memory population order

The memory installed in the MAX5 operates at the same speed as the memory that is installed in the x3690 X5 server. As explained in 2.3.1, "Memory speed" on page 22, the memory speed is derived from the QPI link speed of the installed processors, which in turn dictates the maximum SMI link speed, which in turn dictates the memory speed.

4.7, "Processor options" on page 130, summarizes the memory speeds for all of the models of Intel Xeon 7500 series CPUs.

One important consideration when installing memory in MAX5 configurations is that the server must be fully populated before adding DIMMs to the MAX5. As we describe in 2.3.2, “Memory DIMM placement” on page 23, you get the best performance by using all memory buffers and all DIMM sockets on the server first, and then adding DIMMs to the MAX5.

Figure 4-16 shows the numbering scheme for the DIMM slots on the MAX5, and the pairing of DIMMs in the MAX5. Because DIMMs are added in pairs, they must be matched on a memory port (as shown using the colors). For example, DIMM1 is matched to DIMM 8, DIMM 2 to DIMM 7, DIMM 20 to DIMM 21, and so on.

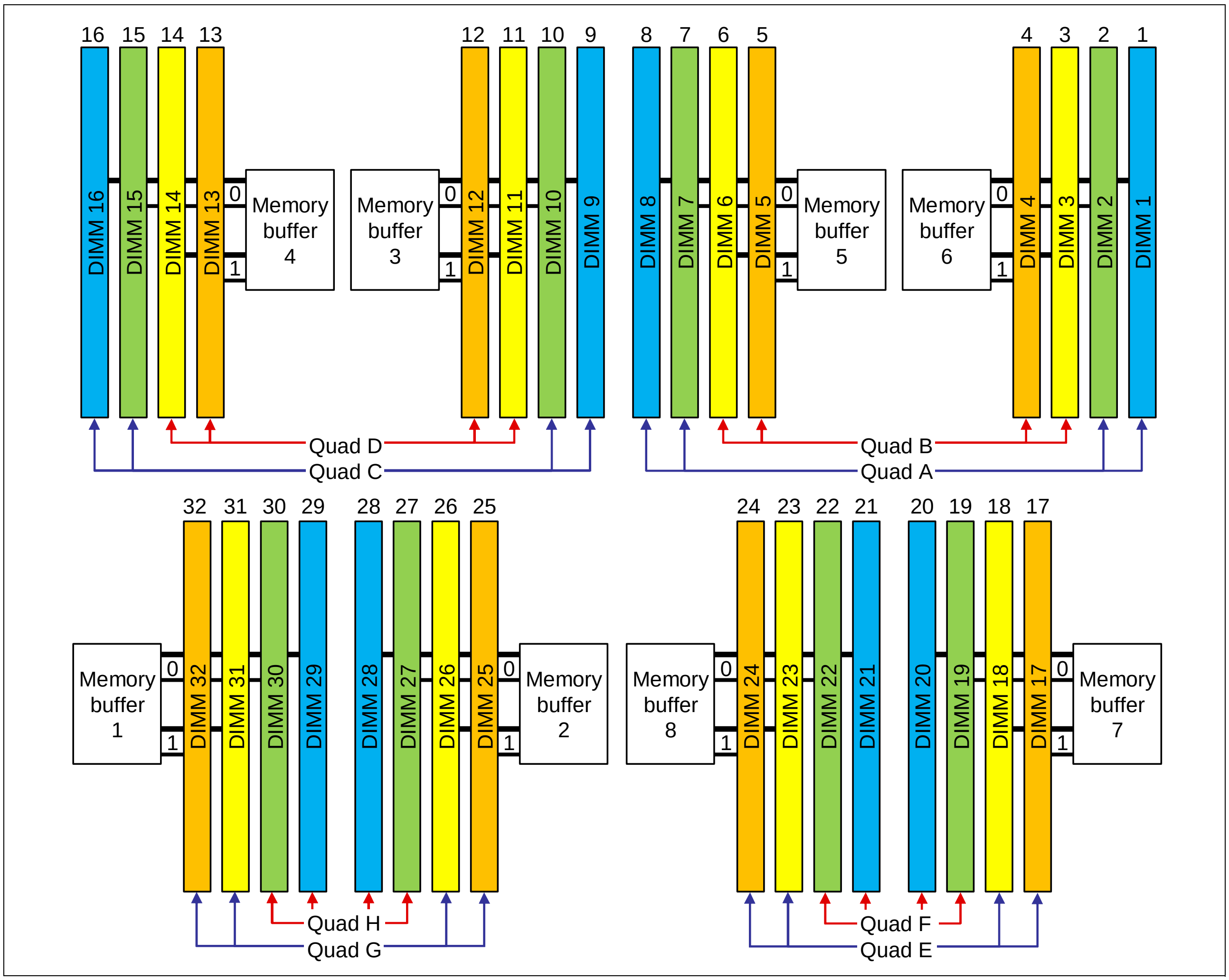


Figure 4-16 DIMM numbering on MAX5

Table 4-11 shows the population order of the MAX5 DIMM slots, which ensures that memory is balanced among the memory buffers. The colors in Table 4-11 match the colors in Figure 4-16.

Table 4-11 DIMM installation sequence in the MAX5

DIMM pair	DIMM slot
1	28 and 29
2	9 and 16
3	1 and 8

DIMM pair	DIMM slot
4	20 and 21
5	26 and 31
6	11 and 14
7	3 and 6
8	18 and 23
9	27 and 30
10	10 and 15
11	2 and 7
12	19 and 22
13	25 and 32
14	12 and 13
15	4 and 5
16	17 and 24

MAX5 memory as seen by the operating system

MAX5 is capable of two modes of operation in terms of the way that memory is presented to the operating system:

- Memory in MAX5 can be split and assigned between the CPUs on the host system (*Non-Pooled mode*). This mode is the default.
- Memory in MAX5 can be presented as a pool of space that is not assigned to any particular CPU (*Pooled mode*).

By default, MAX5 is set to operate in partitioned mode because certain operating systems behave unpredictably when presented with a pool of memory space. Linux can work with memory that is presented either as a pool or pre-assigned between CPUs; however for performance reasons, if you are running Linux, change the setting to pooled mode. VMware requires that the MAX5 memory is in *non-pooled mode*.

You can change this default setting in Unified Extensible Firmware Interface (UEFI). See 7.8, “UEFI settings” on page 337 for details.

VMware vSphere support: MAX5 requires VMware vSphere 4.1 or later.

4.8.4 Memory balance

The Xeon 7500 Series processor uses a nonuniform memory access (NUMA) architecture, as described in 2.3.4, “Nonuniform memory architecture (NUMA)” on page 26. Because NUMA is used, it is important to ensure that all memory controllers in the system are utilized by configuring all processors with memory. Populating all processors in an identical fashion is optimal to provide a balanced system and also required by VMware.

Looking at Figure 4-17 on page 140 as an example, Processor 0 has DIMMs populated, but no DIMMs are populated that are connected to Processor 1. In this case, Processor 0 has access to low latency local memory and high memory bandwidth. However, Processor 1 has access only to remote or "far" memory. So, threads executing on Processor 1 have a longer latency to access memory as compared to threads on Processor 0. This situation is due to the latency penalty incurred to traverse the QPI links to access the data on the other processor's memory controller. The bandwidth to remote memory is also limited by the capability of the QPI links. The latency to access remote memory is more than 50% higher than local memory access.

For these reasons, we advise that you populate all processors with memory, remembering the requirements that are necessary to ensure optimal interleaving and Hemisphere Mode.

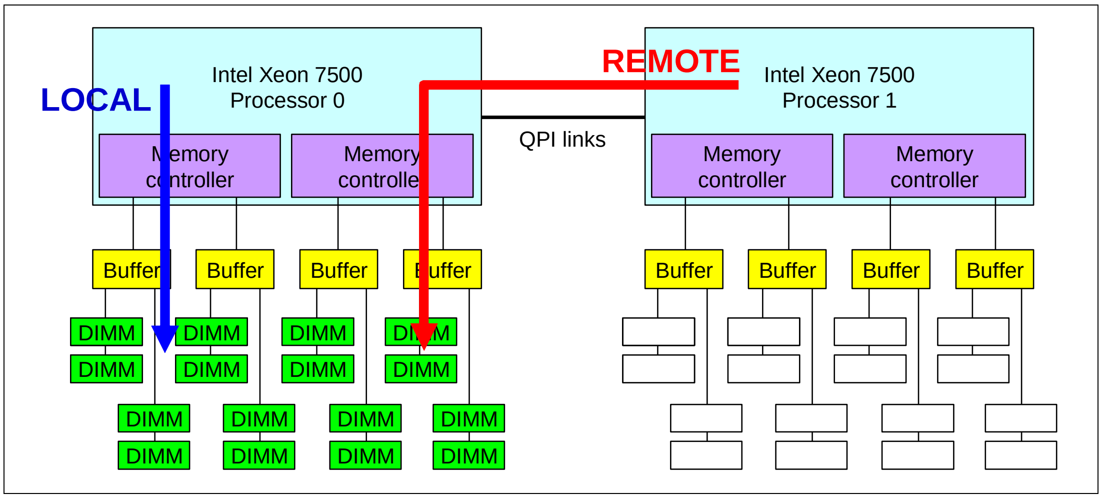


Figure 4-17 Memory latency when not spreading DIMMs across both processors

4.8.5 Mixing DIMMs and the performance effect

Using DIMMs of various capacities is supported for several reasons:

- Not all applications require the full memory capacity that a homogenous memory population provides.
- Cost-saving requirements might dictate using a lower memory capacity for part of the platform's DIMMs.

Figure 4-18 on page 141 illustrates the relative performance of three mixed memory configurations as compared to a baseline of a fully populated memory configuration. While these configurations use 4 GB (4R x8) and 2 GB (2R x8) DIMMs as specified, similar trends to this data are expected when using other mixed DIMM capacities. In all cases, memory is populated in minimum groups of four, as specified in the following configurations, to ensure that Hemisphere Mode is maintained.

Figure 4-18 on page 141 shows the following configurations:

- Configuration A: Full population of equivalent capacity DIMMs (2 GB). This configuration represents an optimally balanced configuration.
- Configuration B: Each memory channel is balanced with the same memory capacity, but half of the DIMMs are of one capacity (4 GB) and half are of another capacity (2 GB).
- Configuration C: Eight DIMMs of one capacity (4 GB) are populated across the eight memory channels, and four additional DIMMs of another capacity (2 GB) are installed one per memory buffer, so that Hemisphere Mode is maintained.
- Configuration D: Four DIMMs of one capacity (4 GB) are populated across four memory channels, and four DIMMs of another capacity (2 GB) are populated on the other four

memory channels, with configurations balanced across the memory buffers, so that Hemisphere Mode is maintained.

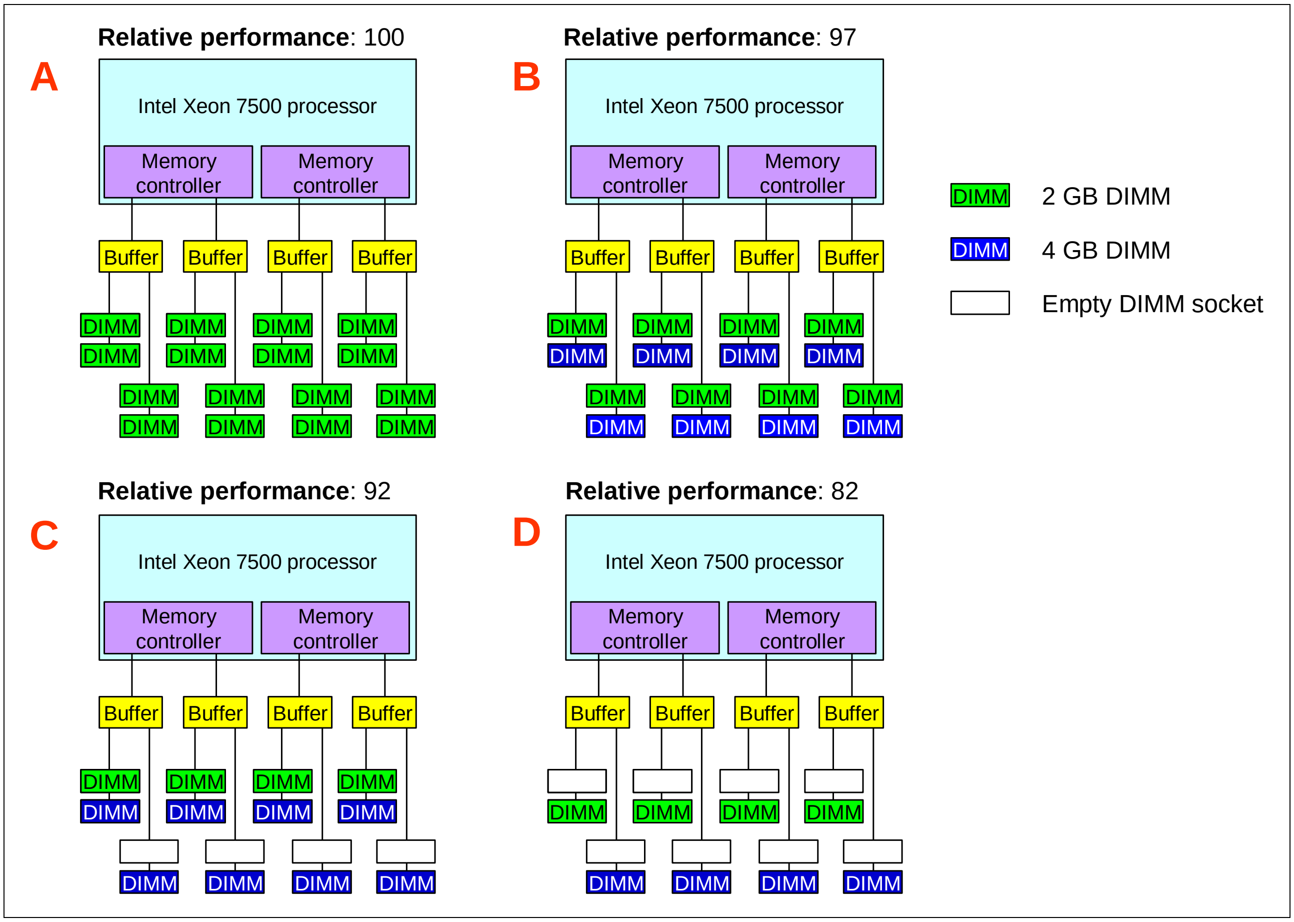


Figure 4-18 Relative memory performance using mixed DIMMs

As you can see, mixing DIMM sizes can cause performance loss up to 18%, even if all channels are occupied and Hemisphere Mode is maintained.

4.8.6 Memory mirroring

Memory mirroring is supported using x3690 X5 and MAX5. The DIMMs must be installed in sets of four. The DIMMs in each set must be the same size and type, which is applicable also when the memory mezzanine is installed in the server and if a MAX5 memory expansion unit is attached to the server. You must install DIMMs in sets of four DIMMs for memory-mirroring mode in each server and memory tray and in the MAX5.

The maximum available memory is reduced to half of the installed memory when memory mirroring is enabled. Partial mirroring (mirroring of part but not all of the installed memory) is not supported. For a detailed understanding of memory mirroring, see “Memory mirroring” on page 28.

DIMM installation for 3690 X5

Table 4-12 lists the DIMM installation sequence for memory-mirroring mode when one or two processors are installed in the server and no memory mezzanine tray is installed in the server.

Table 4-13 shows the DIMM population order for memory-mirroring mode without the mezzanine installed.

Table 4-12 Mirror DIMM installation: Two processors and no memory mezzanine installed

Number of DIMMs	Processor 1 (planar DIMMs)																Processor 2 (No mezzanine DIMMs)															
	Buffer				Buffer				Buffer				Buffer				Buffer				Buffer				Buffer				Buffer			
	DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16	DIMM 17	DIMM 18	DIMM 19	DIMM 20	DIMM 21	DIMM 22	DIMM 23	DIMM 24	DIMM 25	DIMM 26	DIMM 27	DIMM 28	DIMM 29	DIMM 30	DIMM 31	DIMM 32
4	x							x	x							x																
8	x		x			x		x	x		x			x		x																
12	x	x	x			x	x	x	x	x	x			x	x	x																
16	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x																

Table 4-13 DIMM population order: Memory-mirroring mode without the mezzanine installed

Sets of DIMMs	Number of installed processors	DIMM connector population sequence with no memory tray
Set 1	1 or 2	1, 8, 9, 16
Set 2	1 or 2	3, 6, 11, 14
Set 3	1 or 2	2, 7, 10, 15
Set 4	1 or 2	4, 5, 12, 13

Table 4-14 lists the DIMM installation sequence for memory-mirroring mode when two processors and a memory tray are installed in the server.

Table 4-14 Mirror DIMM installation: Two processors and memory mezzanine installed

Number of DIMMs	Processor 1 (planar DIMMs)																Processor 2 (with mezzanine DIMMs)															
	Buffer				Buffer				Buffer				Buffer				Buffer				Buffer				Buffer				Buffer			
	DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16	DIMM 17	DIMM 18	DIMM 19	DIMM 20	DIMM 21	DIMM 22	DIMM 23	DIMM 24	DIMM 25	DIMM 26	DIMM 27	DIMM 28	DIMM 29	DIMM 30	DIMM 31	DIMM 32
8	x							x	x							x	x							x	x							x
16	x		x			x		x	x		x			x		x	x		x			x		x	x		x			x		x
24	x	x	x			x	x	x	x	x	x			x	x	x	x	x	x			x	x	x	x	x	x			x	x	x
32	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Table 4-15 DIMM population order: Memory-mirroring mode with the mezzanine installed

Sets of DIMMs	Number of installed processors	DIMM connector population sequence on the system board	DIMM connector population sequence on the memory tray
Set 1	2	1, 8, 9, 16	17, 24, 25, 32
Set 2	2	3, 6, 11, 14	19, 22, 27, 30
Set 3	2	2, 7, 10, 15	18, 23, 26, 31
Set 4	2	4, 5, 12, 13	20, 21, 28, 29

DIMM installation: MAX5

Table 4-16 shows the installation guide for MAX5 memory mirroring.

Table 4-16 MAX5 memory mirroring setup

Number of DIMMs	MAX5																															
	DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16	DIMM 17	DIMM 18	DIMM 19	DIMM 20	DIMM 21	DIMM 22	DIMM 23	DIMM 24	DIMM 25	DIMM 26	DIMM 27	DIMM 28	DIMM 29	DIMM 30	DIMM 31	DIMM 32
4									x							x												x	x			
8	x							x	x							x				x	x							x	x			
12	x							x	x		x			x		x				x	x					x		x	x		x	
16	x		x			x		x	x		x			x		x		x		x	x		x			x		x	x		x	
20	x		x			x		x	x	x	x			x	x	x		x		x	x		x			x	x	x	x	x	x	
24	x	x	x			x	x	x	x	x	x			x	x	x		x	x	x	x	x	x			x	x	x	x	x	x	
28	x	x	x			x	x	x	x	x	x	x	x	x	x	x		x	x	x	x	x	x		x	x	x	x	x	x	x	x
32	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

4.8.7 Memory sparing

Sparing provides a degree of redundancy in the memory subsystem, but not to the extent that mirroring does. For more information regarding memory sparing, see “Memory sparing” on page 29. This section contains guidelines for installing memory for use with sparing. The two sparing options are DIMM sparing and rank sparing:

- DIMM sparing

  Two unused DIMMs are spared per memory card. These DIMMs must have the same rank and capacity as the largest DIMMs that we are sparing. The size of the two unused DIMMs for sparing is subtracted from the usable capacity presented to the operating system. DIMM sparing is applied on all memory cards in the system.

- Rank sparing

  Two ranks per memory card are configured as spares. The ranks have to be as large as the rank relative to the highest capacity DIMM that we are sparing. The size of the two unused ranks for sparing is subtracted from the usable capacity presented to the operating system. Rank sparing is applied on all memory cards in the system.

These options are configured by using the UEFI during the boot sequence.

4.8.8 Effect on performance of using mirroring or sparing

To understand the effect on performance of selecting various memory modes, we use a system configured with X7560 processors and populated with sixty-four 4 GB quad-rank DIMMs.

Figure 4-19 shows the peak system-level memory throughput for various memory modes measured using an IBM-internal memory load generation tool. As shown, there is a 50% decrease in peak memory throughput when going from a normal (non-mirrored) configuration to a mirrored memory configuration.

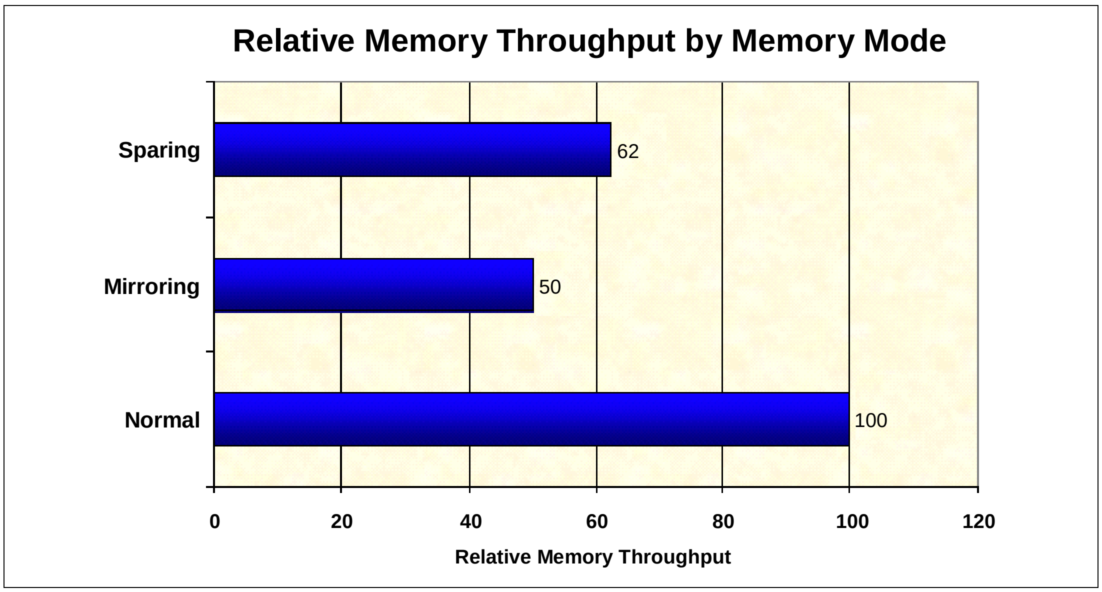


Figure 4-19 Relative memory throughput by memory mode

4.9 Storage

The x3690 X5 has internal capacity of up to sixteen 2.5-inch drives, as shown in Figure 4-20. The server supports 2.5-inch disk drives or solid-state drives (SSDs), or 1.8-inch SSDs.

Figure 4-20 Front of the x3690 X5 with sixteen 2.5-inch drive bays

This section covers the following topics:

- 4.9.1, "2.5-inch SAS drive support" on page 145
- 4.9.2, "IBM eXFlash and SSD disk support" on page 149
- 4.9.3, "SAS and SSD controller summary" on page 152
- 4.9.4, "Battery backup placement" on page 155
- 4.9.5, "ServeRAID Expansion Adapter" on page 157
- 4.9.6, "Drive combinations" on page 158
- 4.9.7, "External SAS storage" on page 162
- 4.9.8, "Optical drives" on page 163

See the IBM ServerProven website for the latest supported options:

http://www.ibm.com/systems/info/x86servers/serverproven/compat/us/

4.9.1 2.5-inch SAS drive support

The server supports up to sixteen 2.5-inch disk drives. These drives are connected to the server using hot-swap backplanes, either four-drive backplanes or eight-drive backplanes or a combination of the two.

Backplanes

Most standard models of the x3690 X5 include one SAS backplane supporting four 2.5-inch SAS disks, as listed in 4.3, "Models" on page 124. Additional backplanes can be added to increase the supported number of SAS disks to 16 (using part number 60Y0338 for a 8x backplane and part number 60Y0369 for a 4x backplane). The database model 7148-3Dx has two IBM eXFlash 8x 1.8-inch HS SAS SSD backplanes as standard. See 4.3, "Models" on page 124 for details. The standard backplanes are installed in the leftmost sections.

Table 4-17 on page 146 lists the backplane options. These backplanes support both SAS and SSD 2.5-inch drives. The specific combinations of the backplanes that are supported are listed in 4.9.6, "Drive combinations" on page 158.

Table 4-17   x3690 X5 hard drive backplanes

Part number	Feature code	Backplane	Drives supported	SAS cables included[a]
60Y0339	9287	IBM 4x 2.5” HS SAS HDD Backplane	Four 2.5” SAS drives	1 short, 1 long
60Y0381	1790	IBM 8x 2.5” HS SAS HDD Backplane	Eight 2.5” SAS drives	2 short, 2 long

a. See the next paragraph for a description of short and long cables. The option part numbers include the cables. If you order a configuration by using feature codes, use Table 4-18.

As listed in Table 4-17, the backplane option part numbers include the necessary cables to connect the backplane to the SAS controller. The *short* SAS cable is needed when installing a hard drive backplane for 2.5-inch bays 1 - 8 (the left half of the drive bays in the server, when looking from the front). The *long* SAS cable is used for hard drive backplanes for 2.5-inch bays 9 - 16 (the right half of the drive bays, when looking from the front).

When configuring an order by using feature codes, for example, with configure-to-order (CTO), the feature codes for the backplanes do not include the cables. You must order the cables separately, as listed in Table 4-18.

Table 4-18   x3690 X5 SAS cable options (not needed if ordering backplane part numbers)

Part number	Feature code	Description	When used
69Y2322	6428	x3690 X5 short SAS cable	For backplanes of bays 1 - 8
69Y2323	6429	x3690 X5 long SAS cable	For backplanes of bays 9 - 16

Using the ServeRAID Expansion Adapter: When using this adapter, the adapter must be installed in PCIe slot 1. Short SAS cables are used to connect the two ports of the ServeRAID controller to the two controller I/O ports on the expander. All four backplane SAS cable connections are connected to the ServeRAID Expander using the long SAS cables that are shown in Table 4-18 on page 146.

Figure 4-21 and Figure 4-23 on page 150 show the backplanes and their cable connections.

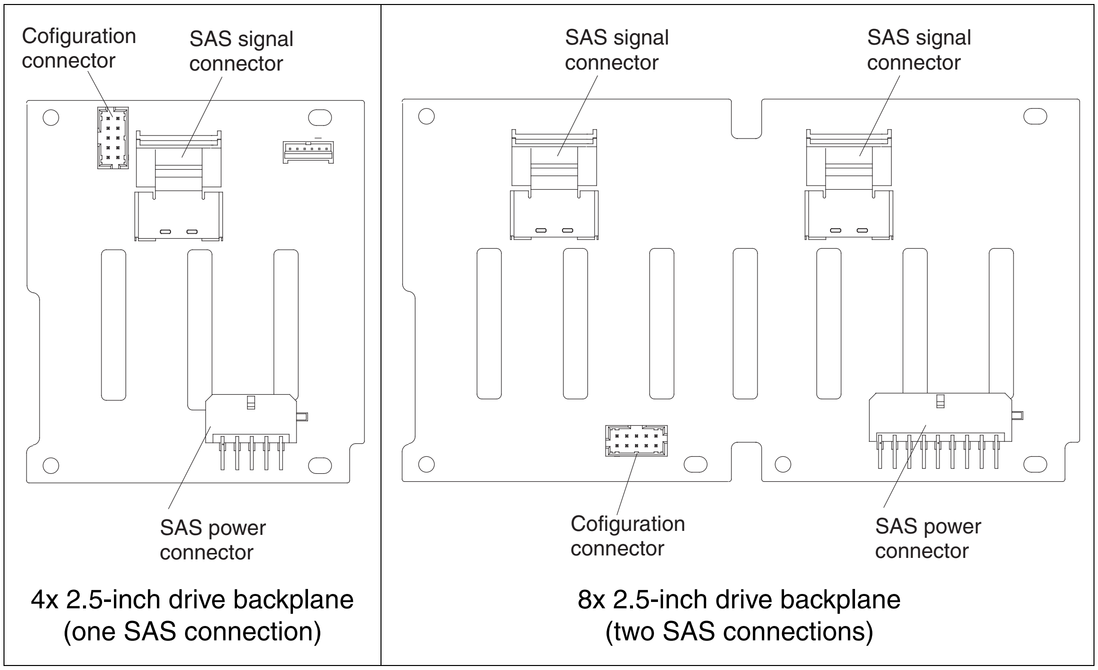


Figure 4-21 The 2.5-inch SAS backplanes (rear view)

Using 2.5-inch disk drives

Table 4-19 lists the 2.5-inch SAS 10K and 15K RPM disk drives and the 2.5-inch SSDs that are supported in the x3690 X5. These drives are supported with the SAS hard disk backplane, part numbers 60Y0338 and 60Y0369.

Table 4-19 The 2.5-inch disk drive options for the x3690 X5

Part number	Feature code	Description	Backplane used
42D0672	5522	IBM 73GB 15K 6Gbps SAS 2.5-inch SFF Slim-HS HDD	4x/8x SAS HDD
42D0632	5537	IBM 146GB 10K 6Gbps SAS 2.5-inch SFF Slim-HS HDD	4x/8x SAS HDD
42D0677	5536	IBM 146GB 15K 6Gbps SAS 2.5-inch SFF Slim-HS HDD	4x/8x SAS HDD
42D0637	5599	IBM 300GB 10K 6Gbps SAS 2.5-inch SFF Slim-HS HDD	4x/8x SAS HDD
43W7714	3745	IBM 50GB SATA 2.5-inch SFF Slim-HS High IOPS SSD	4x/8x SAS HDD
44W2266	5413	IBM 300GB 10K 6Gbps SAS 2.5" SFF Slim-HS SED[a]	4x/8x SAS HDD
44W2296	5412	IBM 146GB 15K 6Gbps SAS 2.5" SFF Slim-HS SED[a]	4x/8x SAS HDD

a. Using the self-encrypting drive (SED) feature of these drives requires a ServeRAID M5014 or M5015 RAID controller, plus either of the ServeRAID M5000 keys, as listed in Table 4-22 on page 149, to add SED support.

Table 4-20 lists the 2.5-inch Nearline SATA 7.2K drives that are supported in the x3690 X5. These drives are supported with the SAS hard disk backplane, 60Y0338 and 60Y0369.

Table 4-20 Supported 2.5-inch Nearline SATA drives

Part number	Feature code	Description
42D0709	5409	IBM 500GB 7200 NL SATA 2.5'' SFF Slim-HS HDD

Self-encrypting drives (SEDs) are also an available option as listed in Table 4-19 on page 147. SEDs provide cost-effective advanced data security with Advanced Encryption Standard (AES) 128 disk encryption. To make use of the encryption capabilities, you must also use either a ServeRAID M5014 or M5015 RAID controller, plus either the ServeRAID M5000 Advance Feature Key or the Performance Accelerator Key. See "Controller options with 2.5-inch drives" on page 149 for details.

For more information about SEDs, see the IBM Redbooks at-a-glance guide, *Self-Encrypting Drives for IBM System x*, TIPS0761, which is available at this website:

http://www.ibm.com/redbooks/abstracts/tips0761.html

Single 500 GB SATA drive

The x3690 X5 optionally supports the IBM x3690 X5 Single SATA HDD Bay, which contains a single 500 GB SATA drive with mounting hardware. You can use the single SATA drive as a boot drive when the system is populated with eXFlash SSDs.

The single SATA HDD bay (Table 4-21) is installed in the rightmost HDD bay, closest to the information panel and encompassing drive bays 12 - 15, as shown in Figure 4-22. Note, however, that no additional drives can be used in bays 12 - 14, because the bays are covered by a filler panel.

DVD-ROM drive: Because the single SATA drive uses the same connector on the system board as the DVD-ROM drive, the DVD-ROM drive cannot be installed when using the SATA drive.

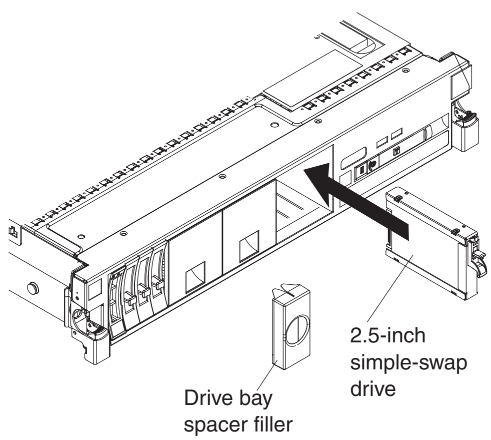


Figure 4-22 Location of the single SATA drive in the x3690 X5

Table 4-21 x3690 X5 Single SATA HDD Bay kit

Option	Feature code	Description
60Y0333	9284	IBM x3690 X5 Single SATA HDD Bay kit

The IBM x3690 X5 Single SATA HDD Bay kit includes the following components:

- 500 GB 7200 RPM 2.5-inch simple-swap SATA drive
- Simple-swap drive backplane and cable
- 4x4 drive bay filler panel
- Drive bay spacer filler

Follow these installation steps for the single SATA drive:

1. Install the single SATA HDD bay assembly into the last bay for drive bays 12 - 15.
2. Disconnect the optical drive's cable from the planar connector.
3. Plug the cable from the single SATA HDD bay into the connector on the planar.
4. Install the SATA drive into drive bay 15.

The 2.5-inch drives require less space than the 3.5-inch drives, consume half the power, produce less noise, can seek faster, and offer increased reliability.

Compatibility: As listed in Table 4-19 on page 147, the 2.5-inch 50 GB SSD is also supported with the standard SAS backplane and the option SAS backplane, part numbers 60Y0338 and 60Y0369. It is not compatible with the 1.8-inch SSD eXFlash backplane, 60Y0359.

A typical configuration can be two 2.5-inch SAS disks for the operating system and two High IOPS disks for data. Only the 2.5-inch High IOPS SSD disk can be used on the SAS backplane. The 1.8-inch disks for the eXFlash cannot be used on the SAS backplane.

Controller options with 2.5-inch drives

Table 4-22 lists the SAS controllers that are supported in the x3690 X5. Most models of the x3690 X5 have a ServeRAID M1015 installed as standard. See 4.3, "Models" on page 124.

Table 4-22 RAID controllers compatible with SAS backplane and SAS disk drives

Part number	Feature code	Description
44E8689	3577	ServeRAID BR10i
46M0831	0095	ServeRAID M1015 SAS/SATA Controller (standard on most models; see 4.3, "Models" on page 124)
46M0832	9749	IBM ServeRAID M1000 Advance Feature Key: Adds RAID-5 and RAID-50 to the ServeRAID M1015 controller
46M0829	0093	ServeRAID M5015 SAS/SATA Controller[a]
46M0916	3877	ServeRAID M5014 SAS/SATA Controller
46M0969	3889	ServeRAID B5015 SSD
46M0930	5106	IBM ServeRAID M5000 Advance Feature Key: Adds RAID-6, RAID-60, and SED Data Encryption Key Management to the ServeRAID M5014, M5015, and M5025 controllers
81Y4426	A10C	IBM ServeRAID M5000 Performance Accelerator Key: Adds Cut Through I/O (CTIO) for SSD FastPath optimization on ServeRAID M5014, M5015, and M5025 controllers.

a. The battery is not included with the ServeRAID M5015 if ordered using the feature code, and it is not needed if using all SSD.

4.9.2 IBM eXFlash and SSD disk support

IBM eXFlash is the name of the feature of the x3690 X5 that offers high-performance 1.8-inch SSDs via optimized eXFlash SSD backplanes and SSD controllers.

IBM eXFlash is available as an option on all models; however, workload-optimized models of the x3690 X5 include one IBM eXFlash SSD backplane that supports eight 1.8-inch SSDs, as listed in “Workload-optimized x3690 X5 models” on page 125. You can add two more eXFlash backplanes to increase the supported number of SSDs to 24.

The IBM eXFlash 8x 1.8-inch HS SAS SSD Backplane, part number 60Y0360, supports eight 1.8-inch High IOPS SSDs, as shown in Table 4-23. The eight drive bays require the same physical space as four SAS hard disk bays. A single eXFlash backplane requires two SAS x4 input cables and one custom power/configuration cable (shipped standard). Up to four SSD backplanes and 24 SSDs are supported in the x3690 X5 chassis.

For more information regarding eXFlash and SSD information, including a brief overview of the benefits from using eXFlash, see 2.8, “IBM eXFlash” on page 47.

Table 4-23 x3690 X5 hard drive backplanes

Part number	Feature code	Backplane	Drives supported	SAS cables included
60Y0360	9281	IBM eXFlash 8x 1.8” HS SAS SSD Backplane	Eight 1.8” solid-state drives	2 short, 2 long

Figure 4-23 shows the 8x 1.8-inch SSD backplane with its two SAS connectors.

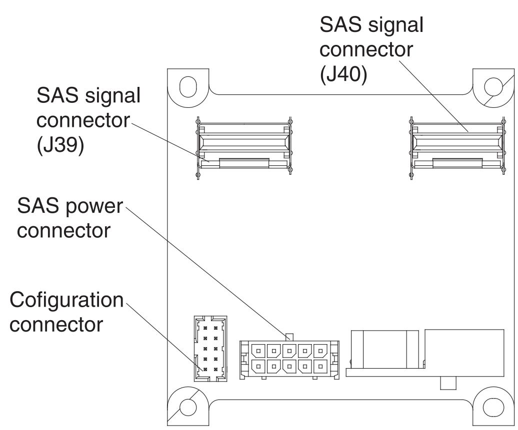


Figure 4-23 8x 1.8-inch SSD backplane (rear view)

Table 4-24 lists the supported 1.8-inch SSDs.

Table 4-24 x3690 X5 hard drive options

Part number	Feature code	Description	Backplane used
43W7735	5314	IBM 50GB SATA 1.8-inch NHS SSD	8x eXFlash SSD

The failure rate of SSDs is low because, in part, the drives have no moving parts. The 50 GB High IOPS SSD is a Single Level Cell (SLC) device with Enterprise Wear Leveling. As a consequence of both of these technologies, the additional layer of protection provided by a RAID controller might not always be necessary in every client environment. In certain cases, RAID-0 might even be an acceptable option.

Table 4-25 on page 151 lists the controllers that support SSDs.

Table 4-25 Controllers supported with the eXFlash SSD backplane option

Part number	Feature code	Description
46M0914	3876	IBM 6Gb SSD Host Bus Adapter (No RAID support)
46M0829	0093	ServeRAID M5015 SAS/SATA Controller[a]
46M0916	3877	ServeRAID M5014 SAS/SATA Controller[a]
46M0969	3889	ServeRAID B5015 SSD
81Y4426	A10C	IBM ServeRAID M5000 Performance Accelerator Key: Adds Cut Through I/O (CTIO) for SSD FastPath optimization on ServeRAID M5014, M5015, and M5025 controllers.

a. Add the Performance Accelerator Key to the ServeRAID M5015 or M5014 for use with SSDs.

Important: When ordering M5000 series controllers for use with SSD drives, you must *not* use the cache battery.

If using M5000 series controllers in a mixed environment, order the cache battery along with Performance Accelerator Key.

If you have already set up the ServeRAID controller that you plan to use and you want to leave the battery attached, you can still disable the write back cache by going into the MegaRAID web BIOS configuration utility and disabling Disk Cache and Default Write, as shown in Figure 4-24.

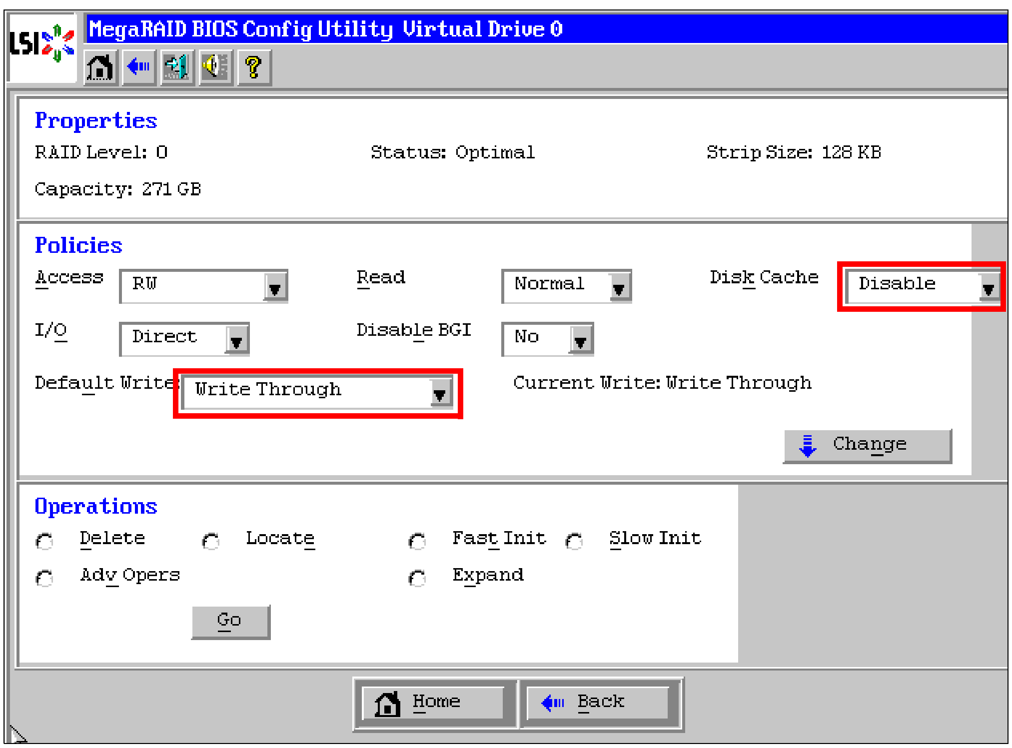


Figure 4-24 Disabling battery cache on controller in MegaRAID web BIOS

ServeRAID M5000 Series Performance Accelerator Key

ServeRAID M5000 Series Performance Accelerator Key for System x enables performance enhancements that are needed by the emerging SSD technologies being used in a mixed SAS and SSD environment. You can enable these performance enhancements by using a seamless field-upgradeable key that works in any M5*xxx* series controller. You gain the following options:

- Performance optimization for SSDs: Improved SAS/SATA controller performance to match an array of SSDs.
- Flash tiering enablement: A data-tiering enabler to support hybrid environments of SSDs and HDDs, realizing higher levels of performance.
- MegaRAID recovery: A data recovery feature that works both in preboot and OS environments.
- Ability to enable RAID-6 and RAID-60 for added data protection.
- Ability to enable SED support for encryption-equipped devices.
- Convenient upgrade with easy-to-use pluggable key.

For more information, see the IBM Redbooks at-a-glance guide *ServeRAID M5000 Series Performance Accelerator Key for IBM System x*, which is available at this website:

http://www.ibm.com/redbooks/abstracts/tips0799.html

4.9.3 SAS and SSD controller summary

In this section, we provide details for the features of each controller card and what they offer. Table 4-26 lists the SAS controllers that are supported in the x3690 X5. Most models of the x3690 X5 have a ServeRAID M1015 installed as standard. See 4.3, “Models” on page 124 for more information.

Table 4-26 lists the disk controllers that are supported in the x3690 X5.

Table 4-26 Disk controllers compatible with the x3690 X5

Part number	Feature code	Name	Supports 2.5-inch SAS backplane	Supports eXFlash SSD backplane	Battery	Cache	RAID support
44E8689	3577	ServeRAID BR10i[a]	Yes	No	No	None	0,1, and 1E
46M0831	0095	ServeRAID M1015	Yes	Yes	No	None	0,1,10,5, and 50[b]
46M0916	3877	ServeRAID M5014	Yes	Yes	Optional	256MB	0,1,10,5,50,6,60[c]
46M0829	0093	ServeRAID M5015	Yes	Yes	Yes[d]	512MB	0,1,10,5,50,6,60[c]
None[e]	3876	IBM 6Gb SSD HBA	No	Yes	No	None	No
46M0969	3889	ServeRAID B5015 SSD	No	Yes	No	None	1 and 5

a. The BR10i is standard on most models. See 4.3, “Models” on page 124.
b. M1015 support for RAID-5 and RAID-50 requires the M1000 Advanced Feature Key (46M0832, fc 9749).
c. M5014 and M5015 support for RAID-6 and RAID-60 requires the M5000 Advanced Feature Key (46M0930, fc 5106).
d. ServeRAID M5015 option part number 46M0829 includes the M5000 battery; however, the feature code 0093 does not contain the battery. Order feature code 5744 if you want to include the battery in the server configuration.

e. The IBM 6Gb SSD Host Bus Adapter is currently not available as a separately orderable option. Use the feature code to add the adapter to a customized order, using the CTO process. Part number 46M0914 is the L1 manufacturing part number. Part number 46M0983 is the pseudo option number, which is also used in manufacturing.

ServeRAID BR10i Controller

The ServeRAID-BR10i has the following specifications:

- LSI 1068e-based adapter
- Two internal mini-SAS SFF-8087 connectors
- SAS 3 Gbps PCIe x8 host bus interface
- Fixed 64 KB stripe size
- Supports RAID-0, RAID-1, and RAID-1E
- No battery and no onboard cache

ServeRAID M5014 and M5015 Controller

The ServeRAID M5014 and M5015 adapter cards have the following specifications:

- Eight internal 6 Gbps SAS/SATA ports.
- Two Mini-SAS internal connectors (SFF-8087).
- Throughput of 6 Gbps per port.
- An 800 MHz PowerPC processor with LSI SAS2108 6 Gbps RAID on Chip (ROC) controller.
- x8 PCI Express 2.0 host interface.
- Onboard data cache (DDR2 running at 800 MHz):
  - ServeRAID M5015: 512 MB
  - ServeRAID M5014: 256 MB
- Intelligent battery backup unit with up to 48 hours of data retention:
  - ServeRAID M5015: Optional for feature code 0093, standard for part 46M0829
  - ServeRAID M5014: Optional

  > **Note:** Battery Cache is not needed when using all SSD drives. If using a controller in a mixed environment with SSD and SAS, you must order and use a battery and the Performance Enablement Key.

- Support for RAID levels 0, 1, 5, 10, and 50 (RAID 6 and 60 support with the optional M5000 Advanced Feature Key).
- Connection of up to 32 SAS or SATA drives.
- SAS and SATA drives are supported, but mixing SAS and SATA in the same RAID array is not supported.
- Up to 64 logical volumes.
- Logical unit number (LUN) sizes up to 64 TB.
- Configurable stripe size up to 1 MB.
- Compliance with Disk Data Format (DDF) configuration on disk (COD).
- Self-Monitoring, Analysis, and Reporting Technology (S.M.A.R.T.) support.
- Support for the optional M5000 Series Performance Accelerator Key, which is recommended when using SSD drives in a mixed environment with SAS and SSD:
  - RAID levels 6 and 60

  - Performance optimization for SSDs
  - LSI SafeStore: Support for self-encrypting drive services, such as instant secure erase and local key management (which requires the use of self-encrypting drives)
- Support for the optional M5000 Advanced Feature Key, which enables the following features:
  - RAID levels 6 and 60
  - LSI SafeStore: Support for self-encrypting drive services, such as instant secure erase and local key management (which requires the use of self-encrypting drives)

Performance Accelerator Key: Performance Accelerator Key uses the same features as the Advanced Feature Key but also includes performance enhancements to enable SSD support in a mixed HDD environment.

For more information, see *ServeRAID M5015 and M5014 SAS/SATA Controllers for IBM System x*, TIPS0738, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0738.html?Open

ServeRAID M1015 Controller

The ServeRAID M1015 SAS/SATA Controller has the following specifications:

- Eight internal 6 Gbps SAS/SATA ports
- SAS and SATA drive support (but not in the same RAID volume)
- SSD support
- Two mini-SAS internal connectors (SFF-8087)
- Throughput of 6 Gbps per port
- LSI SAS2008 6 Gbps RAID on Chip (ROC) controller
- x8 PCI Express 2.0 host interface
- RAID levels 0, 1, and 10 support (RAID levels 5 and 50 with optional ServeRAID M1000 Series Advanced Feature Key)
- Connection of up to 32 SAS or SATA drives
- Up to 16 logical volumes
- LUN sizes up to 64 TB
- Configurable stripe size up to 64 KB
- Compliant with Disk Data Format (DDF) configuration on disk (COD)
- S.M.A.R.T. support

RAID-5, RAID-50, and self-encrypting drive (SED) technology are optional upgrades to the ServeRAID M1015 adapter, with the addition of the ServeRAID M1000 Series Advanced Feature Key, part number 46M0832, feature 9749.

For more information, see *ServeRAID M1015 SAS/SATA Controller for System x,* TIPS0740, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0740.html?Open

IBM 6Gb SSD Host Bus Adapter

The IBM 6Gb SSD Host Bus Adapter is an ideal host bus adapter (HBA) to connect to high-performance SSDs. With two x4 SFF-8087 connectors and a high performance PowerPC I/O processor, this HBA can support the bandwidth that SSDs can generate.

The IBM 6Gb SSD Host Bus Adapter has the following high-level specifications:

- PCI Express 2.0 host interface
- 6 Gbps per port data transfer rate
- MD2 small form factor
- PCI Express 2.0 x8 host interface
- High performance I/O Processor: PowerPC 440 at 533MHz
- UEFI support

For more information, see *IBM 6Gb SSD Host Bus Adapter for IBM System x*, TIPS0744, available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0744.html?Open

> **Important:** Two variants of the 6 Gb Host Bus Adapter exist. The SSD variant has no external port and is part number 46M0914. Do not confuse it with the IBM 6 Gb SAS HBA, part number 46M0907, which is not supported for use with eXFlash.

ServeRAID B5015 SSD Controller

The ServeRAID B5015 is a high-performance RAID controller that is optimized for SSDs. It has the following specifications:

- RAID 1 and 5 support
- Hot-spare support with automatic rebuild capability
- Background data scrubbing
- Stripe size of up to 1 MB
- 6 Gbps per SAS port
- PCI Express 2.0 x8 host interface
- PCI MD2 low profile form factor
- Two x4 internal (SFF-8087) connectors
- SAS controller: PMC-Sierra PM8013 maxSAS 6 Gbps SAS RoC controller
- Up to eight disk drives per RAID adapter
- Performance that is optimized for SSDs
- Three multi-threading MIPS processing cores
- High performance contention-free architecture
- Up to four ServeRAID B5015 adapters supported in a system
- Support for up to four arrays/logical volumes

For more information, see *ServeRAID B5015 SSD Controller*, TIPS0763, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0763.html?Open

> **Important:** This controller does not use Megaraid. This controller is listed in power-on self test (POST) and UEFI as a PMC-SIERRA card. This controller also uses maxRAID Storage Manager for management.

4.9.4 Battery backup placement

When you install RAID adapters that include batteries, the RAID batteries must be remotely located to prevent the batteries from overheating. The batteries must be installed in the RAID battery trays on top of the memory tray or the DIMM air baffle (whichever one is installed in the server).

The battery trays are standard with the server. Each battery tray holds up to two batteries, to support a maximum of four RAID adapters with attached batteries in the x3690 X5.

Table 4-27 lists the kit to order a remote battery cable.

Table 4-27 Remote battery cable ordering

Option	Feature code	Description
44E8837	5862	Remote Battery Cable Kit

The Remote Battery Cable kit, part number 44E8837, contains the following components:

- Remote battery cable
- Plastic interposer
- Plastic stand-off
- Two screws

The screws and stand-off attach the interposer to the RAID controller after the battery is removed. Figure 4-25 shows these components.

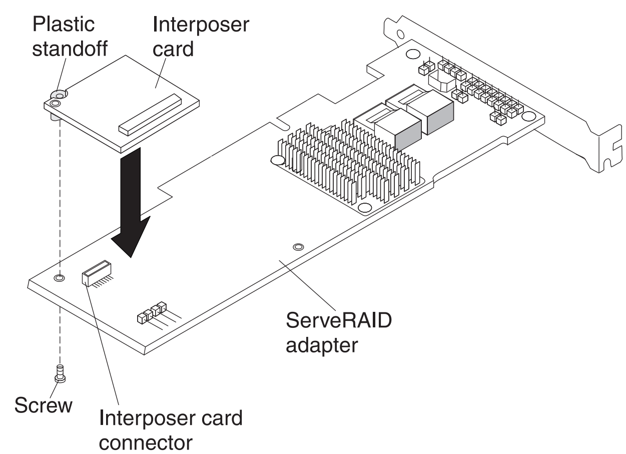


Figure 4-25 Remote battery cable kit

The cable is routed through to the battery that is now installed in the RAID battery tray. This tray is either attached to the memory mezzanine if a memory mezzanine is installed, or the air baffle, which is in place of the mezzanine. Figure 4-26 shows how the battery trays are installed in the memory mezzanine. Each battery tray can hold two batteries.

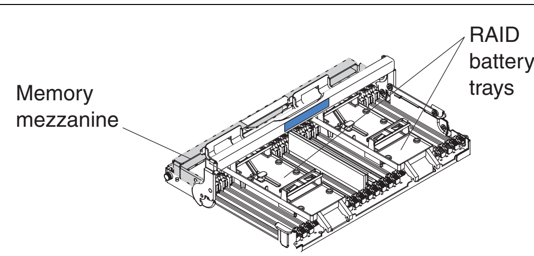


Figure 4-26 RAID battery trays on the memory mezzanine

4.9.5 ServeRAID Expansion Adapter

The ServeRAID Expansion Adapter, which is also known as the IBM x3690 X5 RAID Expansion Adapter or IBM 4x4 Drive Backplane ServeRAID Expansion adapter, is a SAS expander. It allows you to create RAID arrays of up to 16 drives and across up to four backplanes. Table 4-28 shows the ordering information.

Table 4-28 ServeRAID Expansion Adapter ordering

Option	Feature code	Description
60Y0309	4164	ServeRAID Expansion Adapter

The card, which is shown in Figure 4-27, has two input connectors, which you connect to a supported RAID controller, plus four output connectors to go to each backplane, which allows you to connect to up to 16 drives.

Important: You can use only the 2.5-inch hot-swap drive backplanes with this adapter (see Table 4-23).

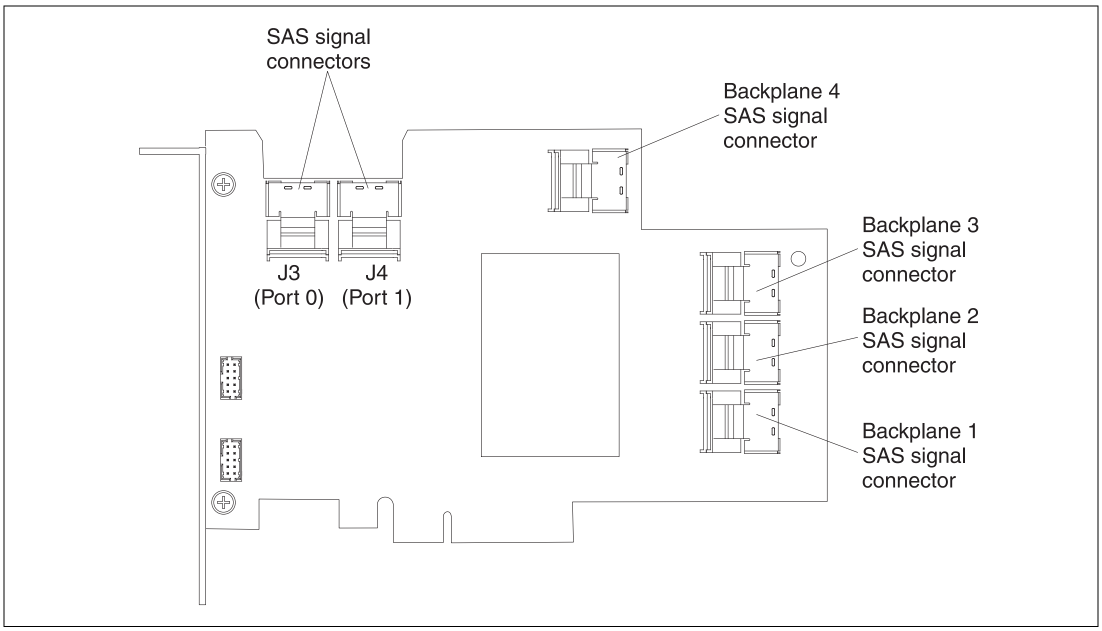


Figure 4-27 ServeRAID Expansion Adapter

You can use the Expansion Adapter only with the following SAS controllers:

- ServeRAID M1015 SAS/SATA adapter
- ServeRAID M5014 SAS/SATA adapter
- ServeRAID M5015 SAS/SATA adapter

The Expansion Adapter must be installed in PCI Slot 1, and the ServeRAID adapter must be installed in PCI Slot 3.

4.9.6 Drive combinations

The x3690 X5 drive subsystem is divided into four backplanes; each backplane can connect to either four 2.5-inch drives or eight 1.8-inch SSDs. This section describes the supported combinations.

Firmware update and installation order:

- You might need a firmware update to the ServeRAID B5015 SSD Controller if you intermix 2.5-inch drives with 1.8-inch SSDs.
- When mixing 2.5-inch backplanes and 1.8-inch backplanes, always install the 2.5-inch backplanes to the left and all 1.8-inch backplanes to the right (as seen when facing the front of the server).
- Not all of these configurations are orderable in a configure-to-order (CTO) configuration.

A configuration with four drives

Figure 4-28 shows a four-drive configuration that uses one 4x HDD backplane. This configuration uses one SAS cable.

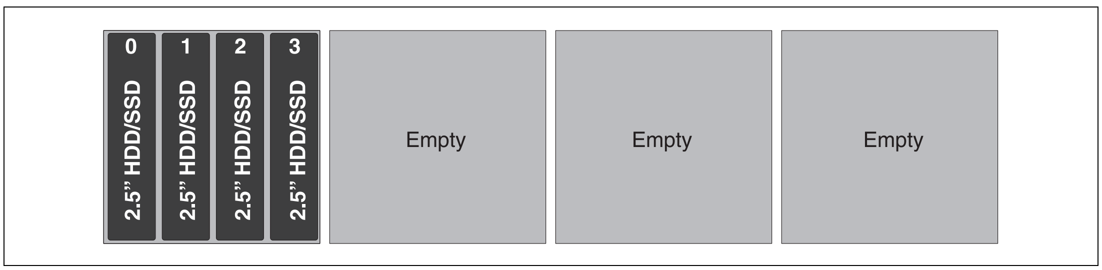


Figure 4-28 x3690 with one IBM 4x 2.5-inch HS SAS HDD backplane

Configurations with eight drives

Figure 4-29 shows two 4x HDD backplanes in use. This configuration requires two SAS cables.

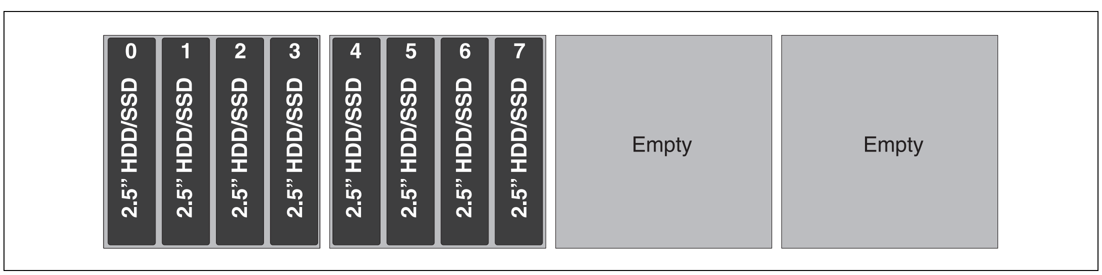


Figure 4-29 x3690 with two IBM 4x 2.5-inch HS SAS HDD backplanes

Figure 4-30 on page 159 shows a configuration that uses one 8x HDD backplane instead of two 4x HDD backplanes. Two SAS cables are needed.

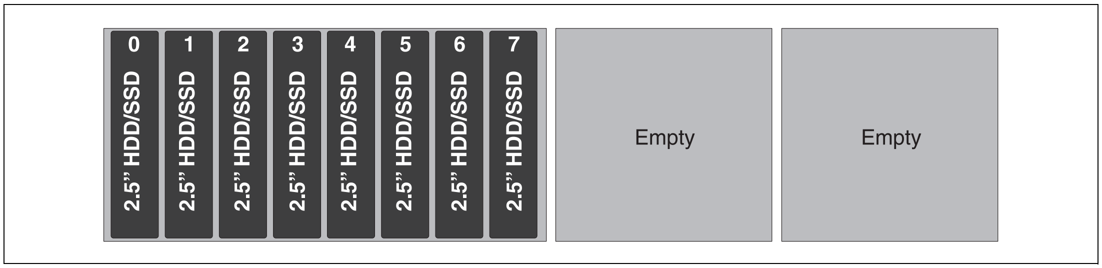


Figure 4-30 x3690 X5 with one IBM 8x 2.5-inch HS SAS HDD backplane

Figure 4-31 illustrates the IBM eXFlash 8x SAS SSD backplane, which requires two SAS cables. With the eXFlash, eight drives can be used in the same space as four 2.5-inch drives.

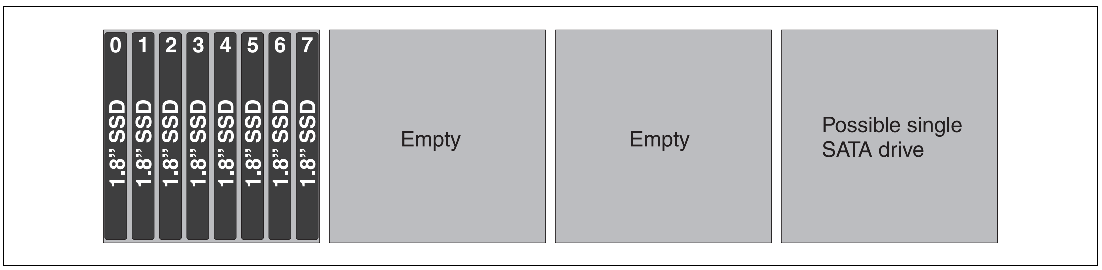


Figure 4-31 x3690 with one IBM eXFlash 8x 1.8-inch HS SAS SSD backplane

Configurations with 12 drives

Figure 4-32 shows three 4x HDD backplanes. This configuration requires three SAS cables.

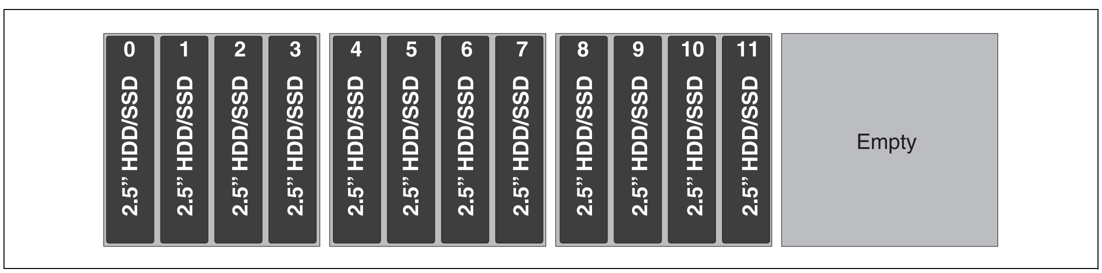


Figure 4-32 x3690 with three IBM 4x 2.5-inch HS SAS HDD backplanes

Figure 4-33 shows one 8x and one 4x HDD backplane resulting in 12 drives. This configuration also requires three SAS cables.

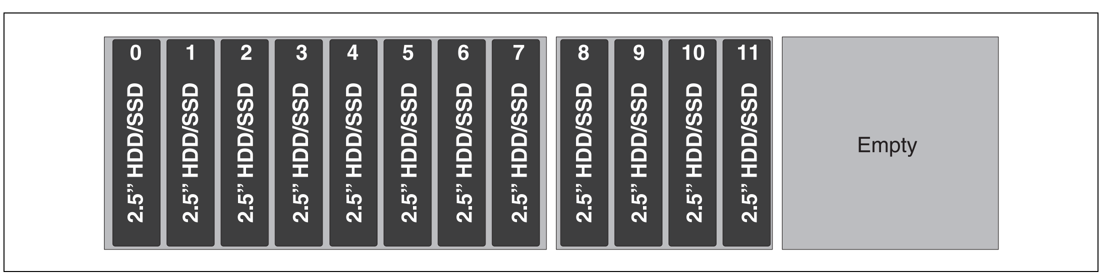


Figure 4-33 x3690 with one 8x 2.5-inch HS SAS HDD and one 4x 2.5-inch HS SAS HDD backplane

Figure 4-34 on page 160 shows a mixture of 2.5-inch HDDs and 1.8-inch SSDs. This configuration requires three SAS cables.

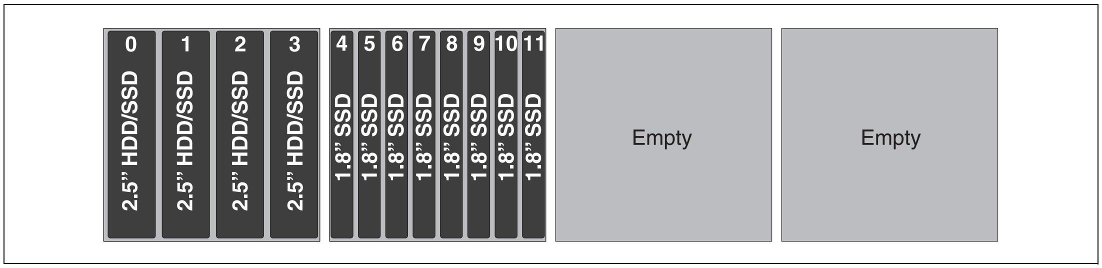


Figure 4-34 x3690 with one 8x 2.5-inch backplane and one eXFlash 8x 1.8-inch SSD backplane

Configurations with 16 drives

Figure 4-35 and Figure 4-36 both show the full sixteen 2.5-inch drive configuration. Both configurations require four SAS cables.

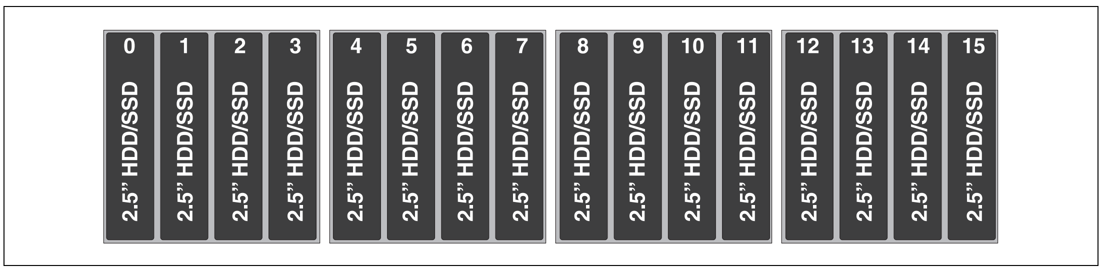


Figure 4-35 x3690 with four IBM 4x 2.5-inch HS SAS HDD backplanes

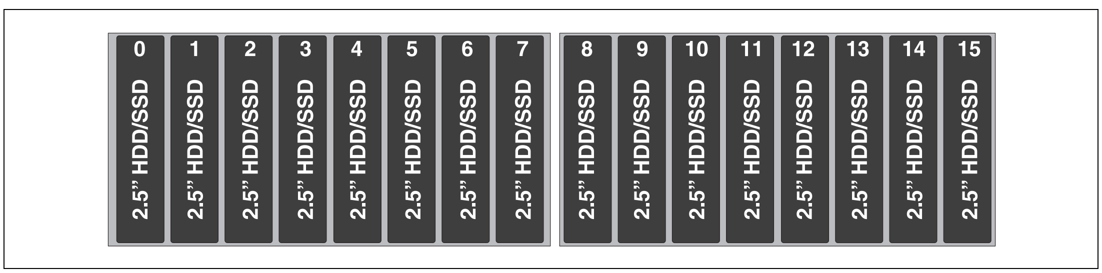


Figure 4-36 x3690 with two IBM 8x 2.5-inch HS SAS HDD backplanes

Figure 4-37 illustrates another 16-drive configuration with one 8x and two 4x backplanes. Also, you can configure this system with the two 4x backplanes for bays 0 - 7 and the 8x backplane for bays 8 - 15. Four SAS cables are required.

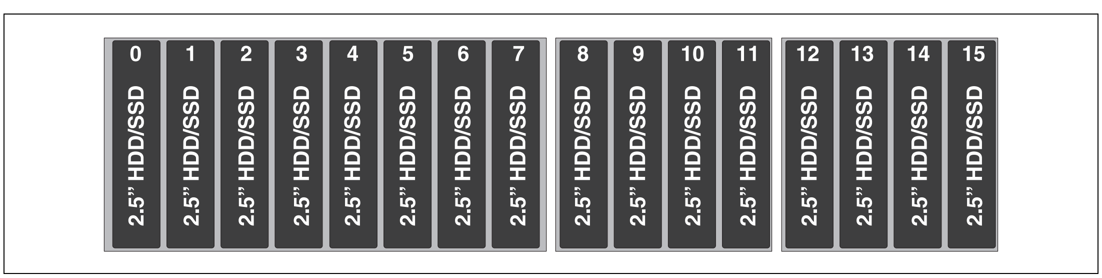


Figure 4-37 x3690 with one 8x 2.5-inch backplane and two 4x 2.5-inch backplanes

Figure 4-38 on page 161 shows two 4x backplanes and one eXFlash backplane. You can use one 8x backplane instead of the two 4x backplanes that are shown here. Four SAS cables are used in this configuration.

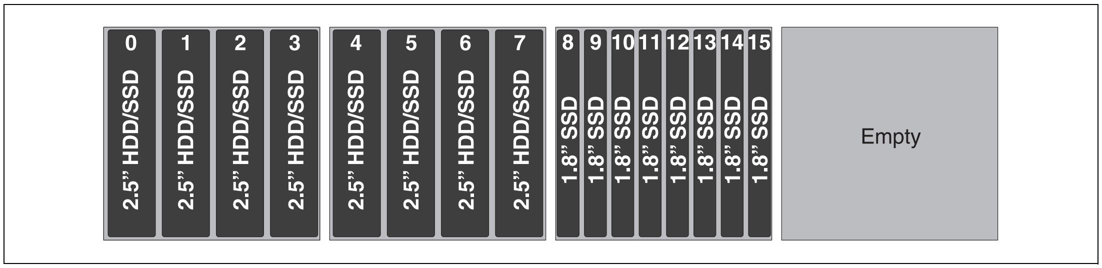


Figure 4-38 x3690 with two 4x 2.5-inch backplanes and one IBM eXFlash 8x 1.8-inch SSD backplane

Figure 4-39 shows two 8x eXFlash backplanes. Using these two backplanes requires four SAS cables. Figure 4-39 also shows the use of a single SATA drive.

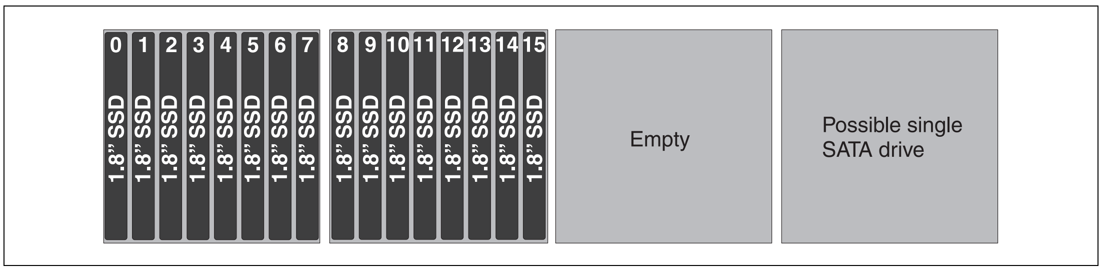


Figure 4-39 x3690 X5 with two IBM eXFlash 8x 1.8-inch SSD backplanes and a single SATA drive

Configurations with 20 drives

Figure 4-40 shows a full complement of drives using three 4x backplanes and one 8x eXFlash backplane. Also, you can achieve this configuration with one 8x backplane, and one 4x and one 8x eXFlash. Either configuration uses five SAS cables.

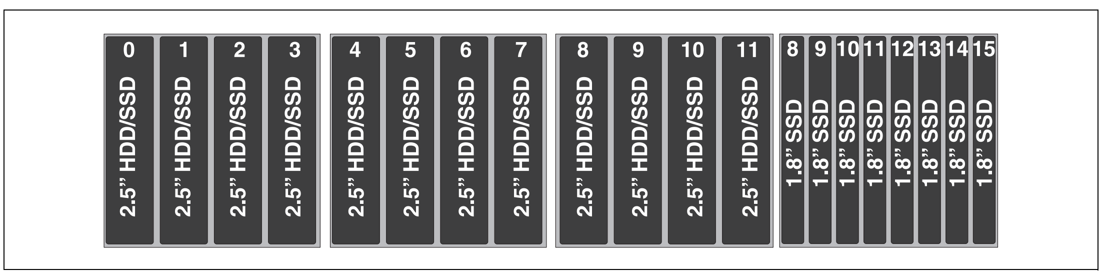


Figure 4-40 x3690 X5 with three 4x 2.5-inch backplanes and one IBM eXFlash 8x 1.8" SSD backplane

Figure 4-41 shows one 4x backplane and two 8x eXFlash backplanes. Five SAS cables are needed.

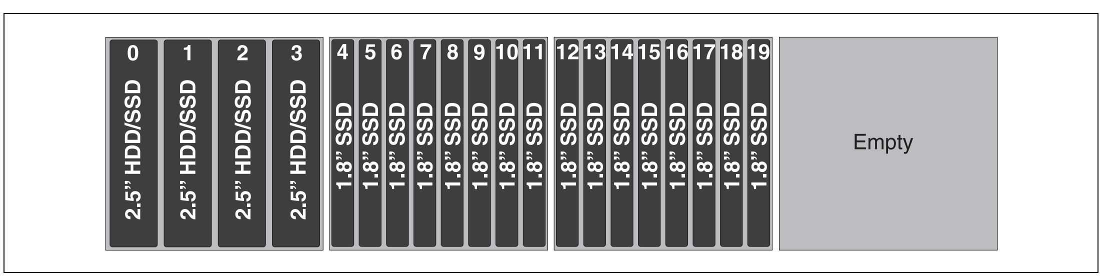


Figure 4-41 x3690 X5 with one 4x 2.5-inch backplane and two IBM eXFlash 8x 1.8" SSD backplanes

Configurations with 24 drives

Figure 4-42 shows two 4x backplanes and 2 8x eXFlash backplanes. The 8x backplane can be used here instead of two 4x backplanes. Six SAS cables are required.

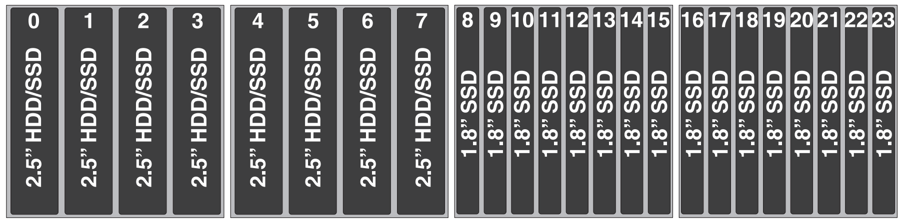


Figure 4-42 x3690 X5 with two 4x 2.5-inch backplanes and two IBM eXFlash 8x 1.8" SSD backplanes

Figure 4-43 shows the maximum number of 8x eXFlash backplanes supported in an x3690 X5. This configuration requires six SAS cables.

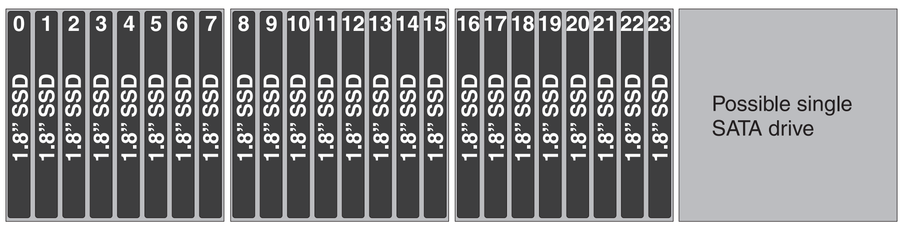


Figure 4-43 x3690 X5 with three IBM eXFlash 8x 1.8-inch SSD backplanes and a possible SATA drive

Important: A configuration of 32 drives is not supported.

4.9.7 External SAS storage

The x3690 X5 supports the use of the ServeRAID M5025 for external SAS storage connectivity. The M5025 offers two external SAS ports to connect to external storage. Table 4-29 lists the cards, support cables, and feature keys.

Table 4-29 External ServeRAID card

Part number	Feature code	Description
46M0830	0094	IBM ServeRAID M5025 SAS/SATA Controller
39R6531	3707	IBM 3m SAS external cable for ServeRAID M5025 to an EXP2512 (1747 HC1) or EXP2524 (1747 HC2)
39R6529	3708	IBM 1m SAS external cable for interconnect between multiple EXP2512 (1747 HC1) or EXP2524 (1747 HC2) units
46M0930	5106	IBM ServeRAID M5000 Advance Feature Key: Adds RAID-6, RAID-60, and SED Data Encryption Key Management to the ServeRAID M5025 controller

The M5025 has two external SAS 2.0 x4 connectors and supports the following features:

- Eight external 6 Gbps SAS 2.0 ports implemented through two four-lane (x4) connectors.
- Two mini-SAS external connectors (SFF-8088).

- 6 Gbps throughput per SAS port.
- 800 MHz PowerPC processor with LSI SAS2108 6 Gbps RAID on Chip (ROC) controller.
- PCI Express 2.0 x8 host interface.
- 512 MB onboard data cache (DDR2 running at 800 MHz).
- Intelligent lithium polymer battery backup unit standard with up to 48 hours of data retention.
- Support for RAID levels 0, 1, 5, 10, and 50 (RAID 6 and 60 support with either M5000 Advanced Feature Key or M5000 Performance Key).
- Connections:
  - Up to 240 SAS or SATA drives.
  - Up to nine daisy-chained enclosures per port.
- SAS and SATA drives are supported, but the mixing of SAS and SATA in the same RAID array is not supported.
- Support for up to 64 logical volumes.
- Support for LUN sizes up to 64 TB.
- Configurable stripe size up to 1024 KB.
- Compliant with Disk Data Format (DDF) configuration on disk (COD).
- S.M.A.R.T. support.
- Support for the optional M5000 Advanced Feature Key, which enables the following features:
  - RAID levels 6 and 60.
  - LSI SafeStore: Support for self-encrypting drive services, such as instant secure erase and local key management (which requires the use of self-encrypting drives).
- Support for SSD drives in a mixed environment with SAS and SSD with the optional M5000 Series Performance Accelerator Key, which enables the following features:
  - RAID levels 6 and 60.
  - Performance optimization for SSDs.
  - LSI SafeStore: Support for self-encrypting drive services, such as instant secure erase and local key management (which requires the use of self-encrypting drives).

For more information, see *ServeRAID M5025 SAS/SATA Controller for IBM System x*, TIPS0739, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0739.html?Open

4.9.8 Optical drives

An optical drive is optional. Table 4-30 on page 163 lists the supported part numbers.

Table 4-30 Optical drives

Part number	Feature code	Description
46M0901	4161	IBM UltraSlim Enhanced SATA DVD-ROM
46M0902	4163	IBM UltraSlim Enhanced SATA Multi-Burner

DVD-ROM: The DVD-ROM drive uses the same connector on the system board as the single SATA drive; therefore, the DVD-ROM drive cannot be installed when using the SATA drive.

4.10 PCIe slots

The x3690 X5 provides five PCIe 2.0 slots for add-in cards. Figure 4-44 shows the location of the slots as viewed from the rear of the server.

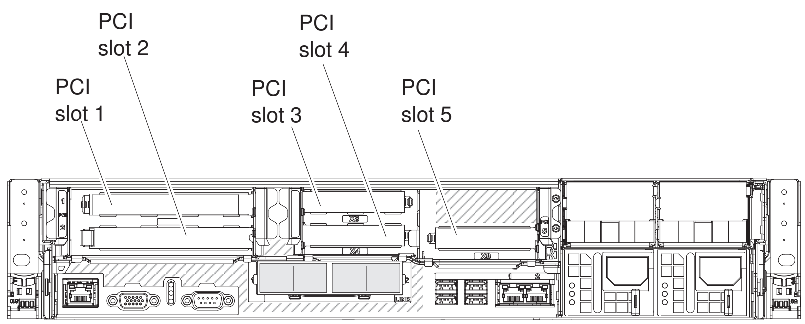


Figure 4-44 x3690 X5 PCIe slots

These slots are connected to the planar through two riser cards, both of which are installed as standard. Figure 4-45 on page 164 shows the locations of the two riser cards in the server.

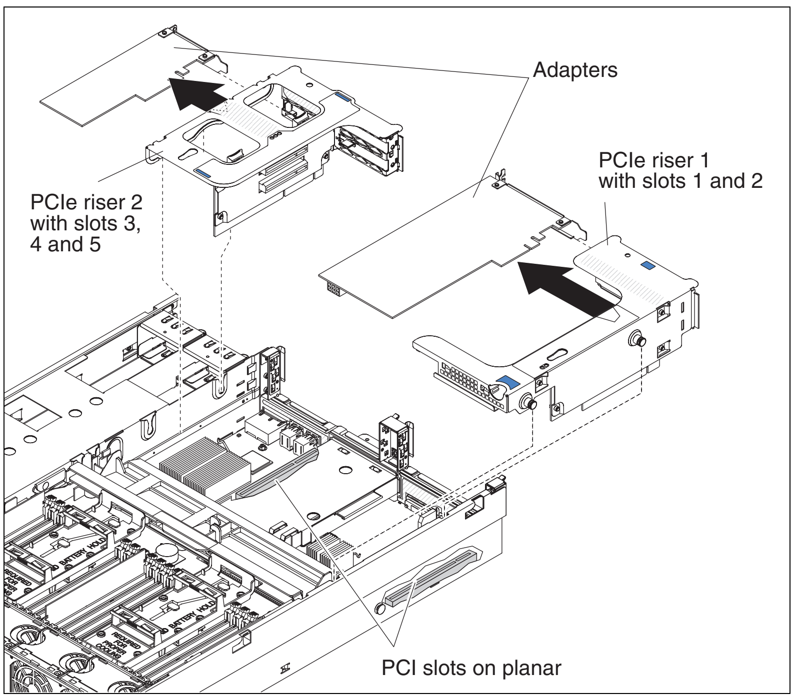


Figure 4-45 Location of the PCIe riser cards in the server

4.10.1 Riser 1

In standard x3690 X5 models, riser slot 1 has the 2x8 riser card installed (60Y0329, feature 9285) which has the following slots:

- Slot 1, PCIe 2.0 x8 full height, full length slot
- Slot 2, PCIe 2.0 x8 full height, half length slot

The 2x8 riser can be replaced by another riser with one PCIe 2.0 x16 slot, which is either a full-length slot or a 3/4-length slot, as listed in Table 4-31. This x16 slot is suitable for graphics processing unit (GPU) adapters. Additional power for the adapter is available from an onboard power connector if needed.

Table 4-31 lists the riser card options for Riser 1. Only one of the risers listed in the table can be installed in the server at a time.

Table 4-31 x3690 X5 PCIe Riser 1 card options

Part number	Feature code	Riser card
60Y0329	9285	IBM System x3690 X5 PCI-Express (2x8) Riser Card[a]
60Y0331	9282	IBM System x3690 X5 PCI-Express (1x16) Riser Card - 3/4 length
60Y0337	9283	IBM System x3690 X5 PCI-Express (1x16) Riser Card - full length[b]

a. The 2x8 riser card is standard in all x3690 X5 models, including 7148-ARx.
b. The 1x16 full-length riser cannot be used if the memory mezzanine is installed in the server.

4.10.2 Riser 2

Riser slot 2 has the 3x8 riser card installed in all standard models, except for model 7148-ARx (see 4.3, “Models” on page 124), and contains the following slots:

- Slot 3, PCIe 2.0 x8 low profile adapter.
- Slot 4, PCIe 2.0 x4 low profile adapter (x8 mechanical).
- Slot 5, PCIe 2.0 x8 low profile adapter. The Emulex 10Gb Ethernet adapter is installed in this slot if the adapter is part of the server configuration.

Full-length adapters: Full-length adapters cannot be installed in any slots if the memory mezzanine is also installed. Instead, adapters up to 3/4 length are supported.

Table 4-32 lists the option.

Table 4-32 x3690 X5 PCIe Riser 2 option

Part number	Feature code	Riser card option
60Y0366	9280	IBM System x3690 X5 PCI-Express (3x8) Riser Card[a]

a. The 3x8 riser card is standard in all x3690 X5 models, *except* 7148-ARx.

Note: The Emulex 10GbE Virtual Fabric Adapter that is standard in most models is installed in slot 5. See 4.10.3, “Emulex 10Gb Ethernet Adapter” on page 166 for details of the adapter.

4.10.3 Emulex 10Gb Ethernet Adapter

As described in 4.3, “Models” on page 124, certain models include the Emulex 10Gb Ethernet Adapter as standard. The card is installed in PCIe slot 5. Slot 5 is a nonstandard x8 slot, which is slightly longer than normal. It accepts both standard PCIe adapters and the Emulex 10Gb Ethernet Adapter.

Tip: The Emulex 10Gb Ethernet Adapter that is standard with specific models is a custom version of the Emulex 10Gb Virtual Fabric Adapter for IBM System x, 49Y4250. However, the features and functions of the two adapters are identical.

The Emulex 10Gb Ethernet Adapter in the x3690 X5 has been customized with a special type of connector called an *extended edge connector*. The card is colored blue instead of green to indicate that it is nonstandard and that it cannot be installed in a standard x8 PCIe slot.

At the time of writing, only the x3850 X5 and the x3690 X5 have slots that are compatible with the custom-built Emulex 10Gb Ethernet Adapter that is shown in Figure 4-46 on page 166.

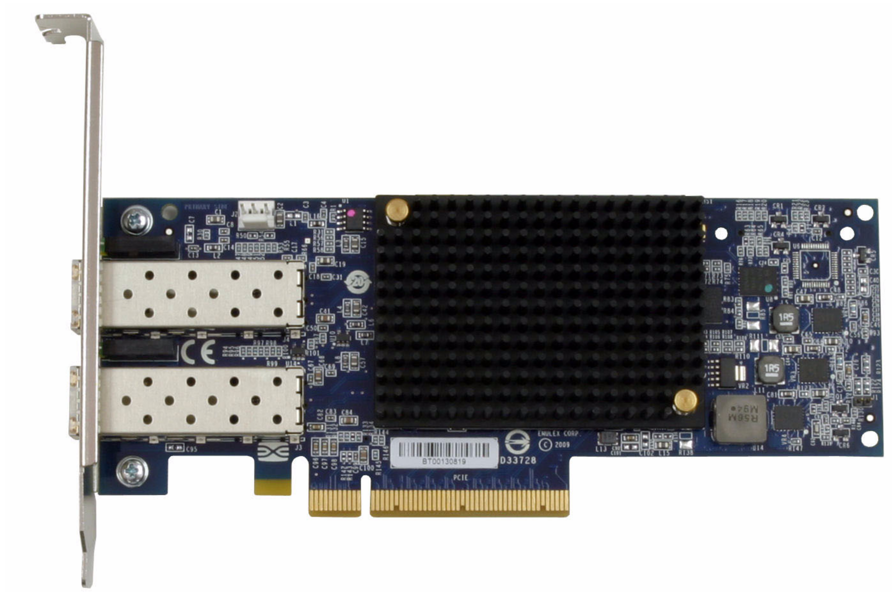


Figure 4-46 The Emulex 10Gb Ethernet Adapter has a blue circuit board and a longer connector

The Emulex 10Gb Ethernet Adapter is a customer-replaceable unit (CRU). To replace the adapter (for example, under warranty), order the CRU number, as shown in Table 4-33. The table also shows the regular Emulex 10Gb Virtual Fabric Adapter (VFA) for IBM System x option, which differs only in the connector type (standard x8) and the color of the circuit board (green).

Redundancy: The standard version of the Emulex VFA and the eX5 extended edge custom version can be used together as a redundant pair. This pair is a supported combination.

Table 4-33 Emulex adapter part numbers

Option description	Part number	Feature code	CRU number
Emulex 10Gb Ethernet Adapter for x3690 X5	None	1648	49Y4202
Emulex 10Gb Virtual Fabric Adapter for IBM System x	49Y4250	5749	Not applicable

General details about this card are in *Emulex 10Gb Virtual Fabric Adapter for IBM System x*, TIPS0762, which is available at the following website:

http://www.redbooks.ibm.com/abstracts/tips0762.html

Important: Although these cards are functionally identical, the availability of iSCSI and Fibre Channel over Ethernet (FCoE) upgrades for one card does not automatically mean availability for both cards. At the time of writing, the target availability of these features is the second quarter of 2011. Check availability of iSCSI and FCoE feature upgrades with your local IBM representative.

The Emulex 10Gb Ethernet Adapter for x3690 X5 includes the following features:

- Dual-channel, 10 Gbps Ethernet controller
- Near line-rate 10 Gbps performance
- Two SFP+ empty cages to support either of the following items:
  - SFP+ SR link is with SFP+ SR Module with LC connectors
  - SFP+ twinaxial copper link is with SFP+ direct attached copper module/cable

  Note: Servers that include the Emulex 10Gb Ethernet Adapter do not include transceivers. You must order transceivers separately if needed, as listed in Table 4-34.

- TCP/IP stateless off-loads
- TCP chimney offload
- Based on Emulex OneConnect technology
- FCoE support as a future feature entitlement upgrade
- Hardware parity, CRC, ECC, and other advanced error checking
- PCI Express 2.0 x8 host interface
- Low-profile form-factor design
- IPv4/IPv6 TCP, User Datagram Protocol (UDP) checksum offload
- VLAN insertion and extraction
- Support for jumbo frames up to 9000 bytes
- Preboot eXecutive Environment (PXE) 2.0 network boot support
- Interrupt coalescing
- Load balancing and failover support
- Deployment and management of this adapter and other Emulex OneConnect-based adapters with OneCommand Manager
- Interoperable with BNT 10Gb Top of Rack (ToR) switch for FCoE functions
- Interoperable with Cisco Nexus 5000, Brocade 10Gb Ethernet switches for NIC/FCoE

SFP+ transceivers are not included with the server and must be ordered separately. Table 4-34 lists compatible transceivers.

Table 4-34 Transceiver ordering information

Option number	Feature code	Description
49Y4218	0064	QLogic 10Gb SFP+ SR Optical Transceiver
49Y4216	0069	Brocade 10Gb SFP+ SR Optical Transceiver
46C3447	5053	BNT SFP+ Transceiver

4.10.4 I/O adapters

Table 4-35 on page 168 shows the current list in the *Configuration and Options Guide* (COG) at the time of writing this paper. See the following website:

http://www.ibm.com/systems/info/x86servers/serverproven/compat/us/

Table 4-35 Available I/O adapters for the x3690 X5

Option	Feature code	Description
Networking		
59Y1887	5763	QLogic QLE7340 single-port 4X QDR IB x8 PCI-E 2.0 HCA
39Y6071	1485	NetXtreme II 1000 Express G Ethernet Adapter- PCIe
49Y4253	5749	Emulex 10GbE Virtual Fabric Adapter
49Y4243	5768	Intel Ethernet Quad Port Server Adapter I340-T4
49Y4233	5767	Intel Ethernet Dual Port Server Adapter I340-T2
49Y4223	5766	NetXtreme II 1000 Express Quad Port Ethernet Adapter
49Y4200	1648	Emulex 10Gb Dual-port Ethernet Adapter
42C1823	1637	Brocade 10Gb CNA
42C1803	5751	QLogic 10Gb CNA
42C1793	5451	NetXtreme II 10 GigE Express Fiber
42C1783	2995	NetXtreme II 1000 Express Dual Port Ethernet Adapter
42C1753	2975	PRO/1000 PF Server Adapter
39Y6139	2974	PRO/1000 PT Quad Port Server Adapter
39Y6129	2944	PRO/1000 PT Dual Port Server Adapter
Storage		
42D0486	3580	Emulex 8Gb FC Single-port HBA
42D0495	3581	Emulex 8Gb FC Dual-port HBA
42D0502	3578	QLogic 8Gb FC Single-port HBA
42D0511	3579	QLogic 8Gb FC Dual-port HBA
46M6051	3589	Brocade 8Gb FC Single-port HBA

Option	Feature code	Description
46M6052	3591	Brocade 8Gb FC Dual-port HBA
59Y1988	3885	Brocade 4Gb FC Single-port HBA
42C2182	3568	QLogic 4Gb FC Dual-Port PCIe HBA
43W7491	1698	Emulex 4GB FC Single-Port PCI-E HBA
43W7492	1699	Emulex 4GB FC Dual-Port PCI-E HBA
Graphics		
49Y6804	1826	NVIDIA Quadro FX 3800

4.11 Standard features

In this section, we describe the standard, onboard features of the x3690 X5. This section covers the following topics:

- 4.11.1, "Integrated management module" on page 169
- 4.11.2, "Ethernet subsystem" on page 170
- 4.11.3, "USB subsystem" on page 170
- 4.11.4, "Integrated Trusted Platform Module" on page 170
- 4.11.5, "Light path diagnostics" on page 170
- 4.11.6, "Cooling" on page 171
- 4.12, "Power supplies" on page 173

4.11.1 Integrated management module

The x3960 X5 contains the Vitesse VSC452 integrated management module (IMM), which combines the baseboard management controller (BMC), video controller, and Remote Supervisor Adapter (RSA) II/CKVM functions into a single chip.

The VSC452 has the following major features:

- 300 MHz 32-bit processor
- BMC I/O, including I^2C and general-purpose I/Os
- Matrox G200 Video core
- DDR2-250 MHz memory controller
- USB 2.0 configurable peripheral
- Avocent digital video compression

The IMM has the following system management features:

- Environmental monitor with fan speed control for temperature, voltages, fan failure, power supply failure, and power backplane failure
- Light path indicators to report fans, power supplies, CPU, voltage regulator module (VRM), and system errors
- System event log
- Automatic CPU disable on failure restart in the two-CPU configuration, when one CPU signals an internal error
- Intelligent Platform Management Interface Specification (IPMI) V2.0 and Intelligent Platform Management Bus (IPMB) support

- Serial Over LAN (SOL)
- Active Energy Manager
- Power/Reset control (power on, hard/soft shutdown, hard/soft reset, and schedule power control)

4.11.2 Ethernet subsystem

The x3690 X5 has an embedded dual 10/100/1000 Ethernet controller. The BCM5709C is a single-chip, high performance, multi-speed dual port Ethernet LAN controller. It contains two standard IEEE 802.3 Ethernet media access controls (MACs), which can operate in either full-duplex or half-duplex mode. Two direct memory access (DMA) engines maximize bus throughput and minimize CPU overhead.

The system has the following features:

- TCP off-load engine (TOE) acceleration
- Shared PCIE interface across two internal Peripheral Component Interconnect (PCI) functions with separate configuration space
- Integrated dual 10/100/1000 MAC and PHY devices able to share the bus through bridge-less arbitration
- Comprehensive nonvolatile memory interface
- IPMI-enabled

4.11.3 USB subsystem

The x3690 X5 contains six external USB 2.0 ports, two on the front of the server as shown in Figure 4-2 on page 119, and four on the rear of the server, as shown in Figure 4-3 on page 120 (rear).

The server also has two internal USB ports, located on riser card 2, as shown in Figure 4-50 on page 175. One of these internal ports is used for the integrated hypervisor key. The other internal port is available.

See 4.13, "Integrated virtualization" on page 174 for more details about the location of the internal USB ports and the placement of the internal hypervisor key.

4.11.4 Integrated Trusted Platform Module

The Integrated Winbond Trusted Platform Module (TPM) Version 1.2 (WPCT201BA0WG) security chip performs cryptographic functions and stores private and public security keys. It provides the hardware support for the Trusted Computing Group (TCG) specification. For more information about the TCG specification, go to the following website:

http://www.trustedcomputinggroup.org/resources/tpm_main_specification

4.11.5 Light path diagnostics

Light path diagnostics is a system of LEDs used to indicate failed components or system errors. When an error occurs, LEDs are lit on the light path diagnostics panel. Figure 4-47 on page 171 shows the location of the light path diagnostics panel on the x3690 X5.

Figure 4-47   x3690 X5 light path diagnostics panel

Light path diagnostics can alert the user to the following errors:

- Over current faults
- Fan faults
- Power supply failures
- PCI errors

You can obtain the full details about the functions and operation of light path diagnostics in this system in the *Installation and User's Guide - IBM System x3690 X5* at the following website:

http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5085206

4.11.6 Cooling

The x3690 X5 has the following fans:

- Five, hot-swappable fans that are located in the front portion of the chassis
- Power supply internal fans that are located at the rear of each power supply

Fans are numbered left to right, if you are looking at the front of the chassis. Fan 1 is nearest the power supplies, and Fan 5 is nearest the operator information panel. Figure 4-48 on page 172 shows the location of the fans. The individual fans are hot-swappable, as denoted by the orange release latches. The complete fan housing unit is not hot swappable.

Fans 1 - 5 are accessible through an opening in the server top cover, the hot-swap fan access panel. You do not have to remove the server top cover to access the fans.

> **Attention**: If you release the cover latch and remove the server top cover while the server is running, the server is automatically powered off immediately. This powering off is required for electrical safety reasons.

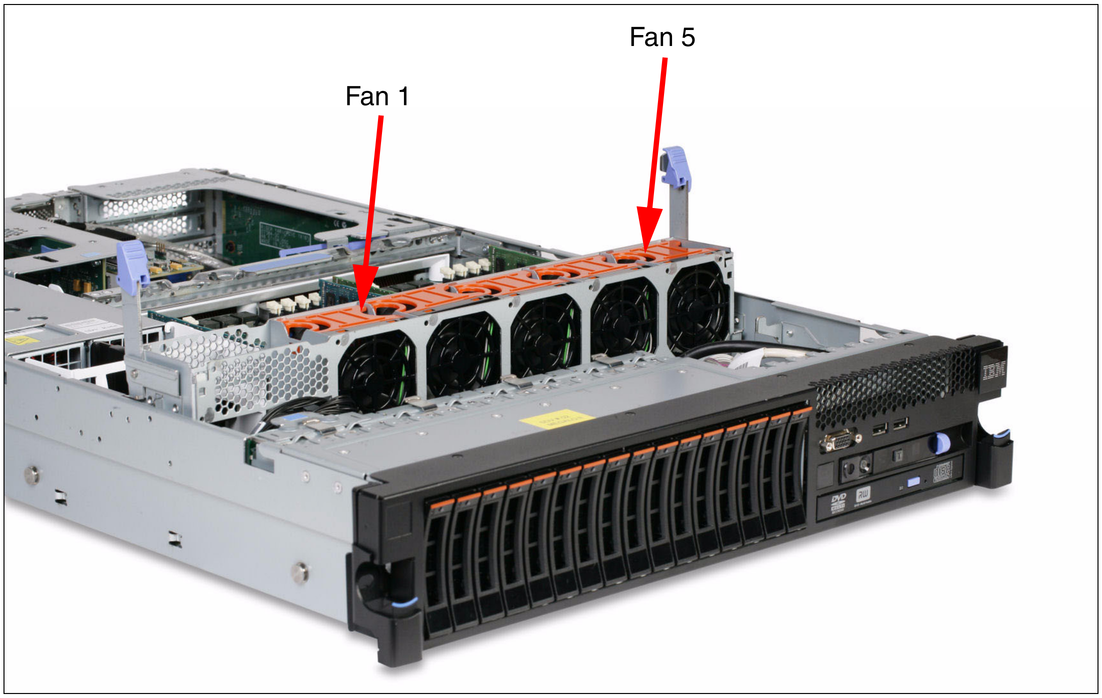


Figure 4-48 x3690 X5 fans

Figure 4-49 shows the top of the server and the hot-swap fan access panel.


Figure 4-49 Hot-swap fan access panel

The following conditions affect system fan-speed adjustments:

- Inlet Ambient temperature
- CPU temperatures

- DIMM temperatures
- Altitude

4.12 Power supplies

This section covers the power subsystem of the x3690 X5 and the MAX5.

4.12.1 x3690 X5 power subsystem

The x3690 X5 power subsystem consists of up to four hot-pluggable 675W auto-sensing power supplies. The modules are independently powered by ac line cords.

Most standard models have one power supply as standard; workload-optimized models have more. See 4.3, “Models” on page 124 for details. One power supply is sufficient when the total power budget is less than 675W. Use the IBM System x and BladeCenter Power Configurator to determine the power requirements of your configuration:

http://www.ibm.com/systems/bladecenter/resources/powerconfig.html

For power budgets under 675W, installing a second power supply provides redundancy. To install a second power supply, use the IBM High Efficiency 675W Power Supply, part number 60Y0332, feature code 4782.

Installing four power supplies ensures redundancy even with a fully loaded server. To install the third and fourth power supplies, use the IBM 675W Redundant Power Supply Kit, part number 60Y0327. The power subsystem is designed for N+N operation and hot-swap exchange. Having four power supplies installed allows for N+N redundancy, where N=2 (that is, a total of four power supplies where two power supplies are redundant backups for the other two).

Table 4-36 shows the part numbers for the power supply options.

Table 4-36 IBM 675W Redundant Power Supply Kit for x3690 X5

Option	Feature code	Description	Use
60Y0332	4782	IBM High Efficiency 675W Power Supply	Power supply 2
60Y0327	Various[a]	IBM 675W Redundant Power Supply Kit	Power supply 3 and 4

a. Use 4782 for the power supplies, 9279 for the power supply interposer, and 6406 for the Y-cable.

The IBM 675W Redundant Power Supply Kit, option 60Y0327, includes the following items:

- Two 675W power supplies
- Two Y-cord power cables (2.8 m, 10A/200-250V, 2xC13 to IEC 320-C14)
- Two power cables (2.8 m, 10A/100-250V, C13 to IEC 320-C14)
- One power interposer card

The Redundant Power Supply Kit includes a power supply interposer (power backplane). The *interposer* is a small circuit board that routes power from the power supply outputs to the system planar.

Table 4-37 on page 174 lists the ac power input requirements.

Table 4-37 Power Supply ac input requirements

	Minimum	Maximum	Nominal	Maximum input current
Low range	90 V ac	137 V ac	100-127 V ac 50/60 Hz	7.8A RMS
High range	180 V ac	265 V ac	200-240 V ac 50/60 Hz	3.8A RMS

4.12.2 MAX5 power subsystem

The MAX5 power subsystem consists of one or two hot-pluggable 675W power supplies. The power subsystem is designed for N+N (fully redundant) operation and hot-swap replacement.

Most standard models of MAX5 have one power supply installed in power supply bay 1, as listed in 4.3, “Models” on page 124. For redundancy, install the second power supply, as listed in Table 4-38.

Table 4-38 Ordering information for the IBM MAX5 for System x

Part number	Feature code	Description
60Y0332	4782	IBM 675W HE Redundant Power Supply

A fan that is located inside each power supply cools the power modules.

MAX5 has five redundant hot-swap fans, all in one cooling zone. The MAX5 fan speed is controlled by the IMM of the attached host, based on altitude and ambient temperature.

Fans also respond to certain conditions and come up to speed accordingly:

- If a fan fails, the remaining fans spin up to full speed.
- As the internal temperature rises, all fans spin up to full speed.

4.13 Integrated virtualization

The VMware ESXi embedded hypervisor software is a virtualization platform that allows multiple operating systems to run on a host system at the same time. An internal USB connector on the x8 low profile PCI riser card, as shown in Figure 4-50 on page 175, is reserved to support one USB flash drive, with hypervisor software preloaded, to enable the embedded hypervisor function. See Table 4-39 on page 175 for details.

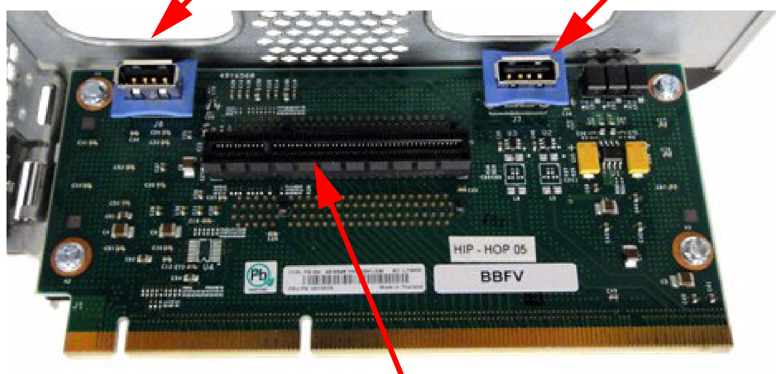


Figure 4-50 Low profile x8 riser card with hypervisor flash USB connector

The IBM USB Memory Key for virtualization is included in the virtualization-optimized models that are listed in 4.3, “Models” on page 124, but it can be added to any x3690 X5 system.

Table 4-39 USB key for embedded hypervisor

Option	Feature code	Description
41Y8278	1776	IBM USB Memory Key for VMware ESXi 4
41Y8287	2420	IBM USB Memory Key for VMware ESXi 4.1 with MAX5

For additional information and setup instructions for VMware ESXi software, see the *VMware ESXi Embedded and vCenter Server Setup Guide* that is available at the following website:

http://www.vmware.com/pdf/vsphere4/r41/vsp_41_esxi_e_vc_setup_guide.pdf

Also, before installing VMware, see the installation guide at 7.9.5, “VMware vSphere ESXi 4.1” on page 358.

4.14 Supported operating systems

The x3690 X5 support for operating systems includes the following items:

- Microsoft Windows Server 2008, Datacenter x64 Edition
- Microsoft Windows Server 2008, Enterprise x64 Edition
- Microsoft Windows Server 2008, Standard x64 Edition
- Microsoft Windows Server 2008, Web x64 Edition
- Windows HPC Server 2008
- Red Hat Enterprise Linux 5 Server x64 Edition
- Red Hat Enterprise Linux 5 Server with Xen x64 Edition
- SUSE LINUX Enterprise Server 11 with Xen for AMD64/EM64T
- SUSE LINUX Enterprise Server 11 for AMD64/EM64T
- SUSE LINUX Enterprise Server 10 with Xen for AMD64/EM64T
- SUSE LINUX Enterprise Server 10 for AMD64/EM64T
- VMware ESX 4.0
- VMware ESXi 4.0
- VMware ESX 4.1

- VMware ESXi 4.1

Memory and processor limits: Certain operating systems have upper limits to the amount of memory that is supported (for example, over 1 TB) or the number of processor cores that are supported (over 64 cores). See the ServerProven Plan for x3690 X5 for details and the full list of supported operating systems at the following website:

`http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/matrix.shtml`

VMware vSphere support: MAX5 requires VMware ESX 4.1 or later.

VMware ESX and two processors: If you plan to install VMware ESX or ESXi on the x3690 X5 with two installed processors, you must also install and populate the memory mezzanine. Failure to do so results in the following error:

“`NUMA node 1 has no memory`”

4.15 Rack mounting

The x3690 X5 is a 2U-high device (1U is one rack unit and is 1.75 inches). The MAX5 memory expansion unit is an additional 1U high unit. Both devices are designed to be installed in standard 19-inch racks. Three slide kits are available for use with the x3690 X5, as listed in Table 4-40.

Table 4-40 Rail kit options

Part number	Feature code	Name	Use
69Y2345	4786	IBM System x3690 X5 Ball Bearing Slide Kit	Required if you plan to attach a MAX5 unit
69Y4403	4178	Universal Slides Kit	Designed to fit telecommunications and short racks
69Y4389	6457	Friction Slide	A low-cost rail kit

Cable management arms (CMAs) are optional but useful because they help prevent cables from becoming tangled and causing server downtime. Table 4-41 lists the available cable management arms.

Table 4-41 Cable management arms

Part number	Feature code	Name	Use with rail kit
69Y2347	6473	IBM System x3690 X5 Cable Management Arm for Ball Bearing Slides	69Y2345
69Y2344	6474	IBM System x3690 X5 2U Cable Management Arm	69Y4403
69Y4390	6458	Friction CMA	69Y4389

5

IBM BladeCenter HX5

The IBM BladeCenter HX5 blade server showcases the eX5 architecture and technology in a blade form factor. This chapter introduces the server and describes its features and options.

This chapter contains the following topics:

5.1 Introduction

The IBM BladeCenter HX5 supports up to two processors using Intel Xeon 6500 (Nehalem EX) 4-core, 6-core, or 8-core processors or four processors using the Intel Xeon 7500 Nehalem EX 4-core, 6-core, or 8-core processors. The HX5 supports up to 40 dual inline memory modules (DIMMs) with the addition of the MAX5 memory expansion blade when using Xeon 7500 Nehalem EX.

Figure 5-1 shows the following three configurations:

- Single-wide HX5 with two processor sockets and 16 DIMM sockets.
- Double-wide HX5 with four Xeon 7500 series processors and 32 DIMM sockets.
- Double-wide HX5 with two Xeon 7500 or 6500 series processors and 40 DIMM sockets: 16 in the HX5 server and 24 in the attached MAX5 memory expansion blade.

MAX5: MAX5 can only connect to a single HX5 server.

HX5 2-socket

HX5 4-socket

HX5 2-socket with MAX5

Figure 5-1 IBM BladeCenter HX5 blade server configurations

Table 5-1 lists the features of the HX5.

Table 5-1 Features of the HX5 type 7872

Features	**HX5 2-socket**	**HX5 4-socket**	**HX5 2-socket with MAX5**
Form factor	30 mm (1-wide)	60 mm (2-wide)	60 mm (2-wide)
Maximum number of processors	Two	Four	Two
Processor options	Intel Xeon 6500 and 7500 Four-core, 6-core, or 8-core	Intel Xeon 7500 Four-core, 6-core, or 8-core	Intel Xeon 7500 Four-core, 6-core, or 8-core
Cache	12 MB, 18 MB, or 24 MB (shared between cores) (processor-dependent)		
Memory speed	978 or 800 MHz (processor SMI link speed dependent)		HX5: Up to 800 MHz MAX5: Up to 1066 MHz
DIMM slots	16	32	40

Features	HX5 2-socket	HX5 4-socket	HX5 2-socket with MAX5
Maximum RAM (using 8 GB DIMMs)	128 GB	256 GB	320 GB
Memory type	DDR 3 error checking and correction (ECC) Very Low Profile (VLP) Registered DIMMs		
DIMMs per channel	1	1	HX5: 1; MAX5: 2
Internal storage	Optional 1.8-inch solid-state drives (SSDs); Non-hot-swap (require an additional SSD carrier)		
Maximum number of drives	Two	Four	Two
Maximum internal storage	Up to 100 GB using two 50 GB SSDs	Up to 200 GB using four 50 GB SSDs	Up to 100 GB using two 50 GB SSDs
I/O expansion	► One CIOv ► One CFFh	► Two CIOv ► Two CFFh	► One CIOv ► One CFFh

Figure 5-2 shows the components on the system board of the HX5.

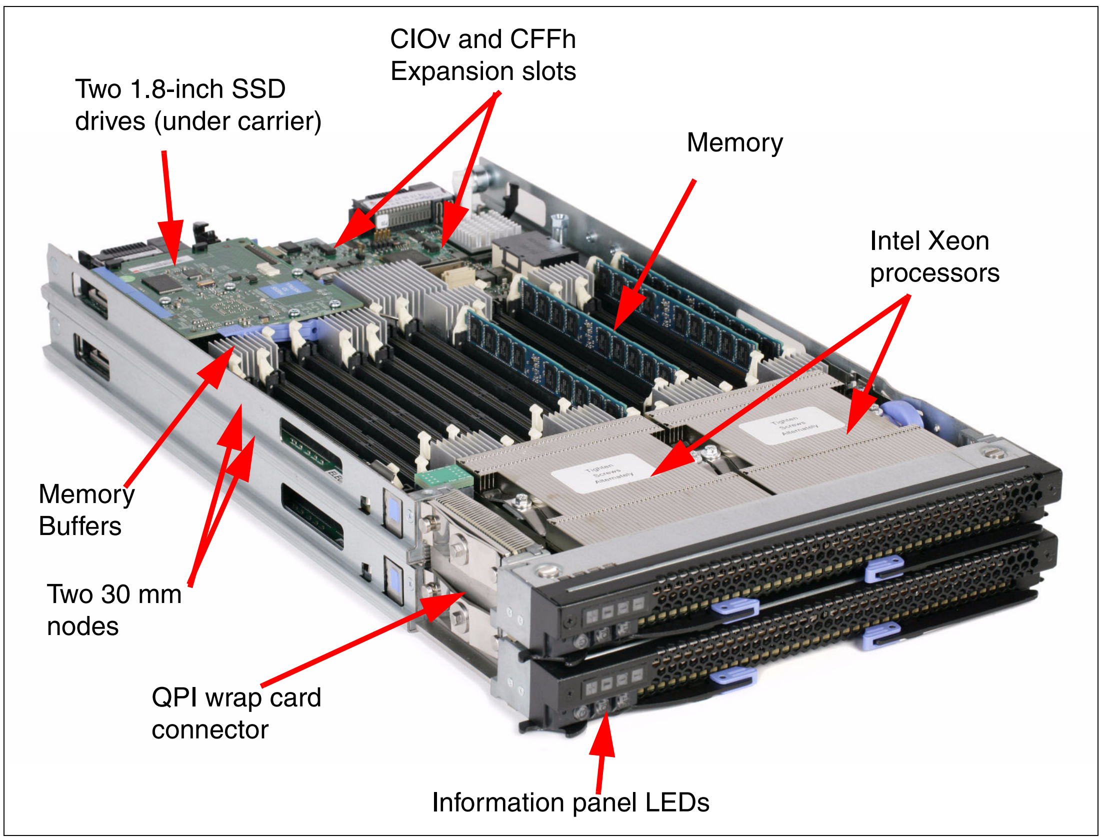
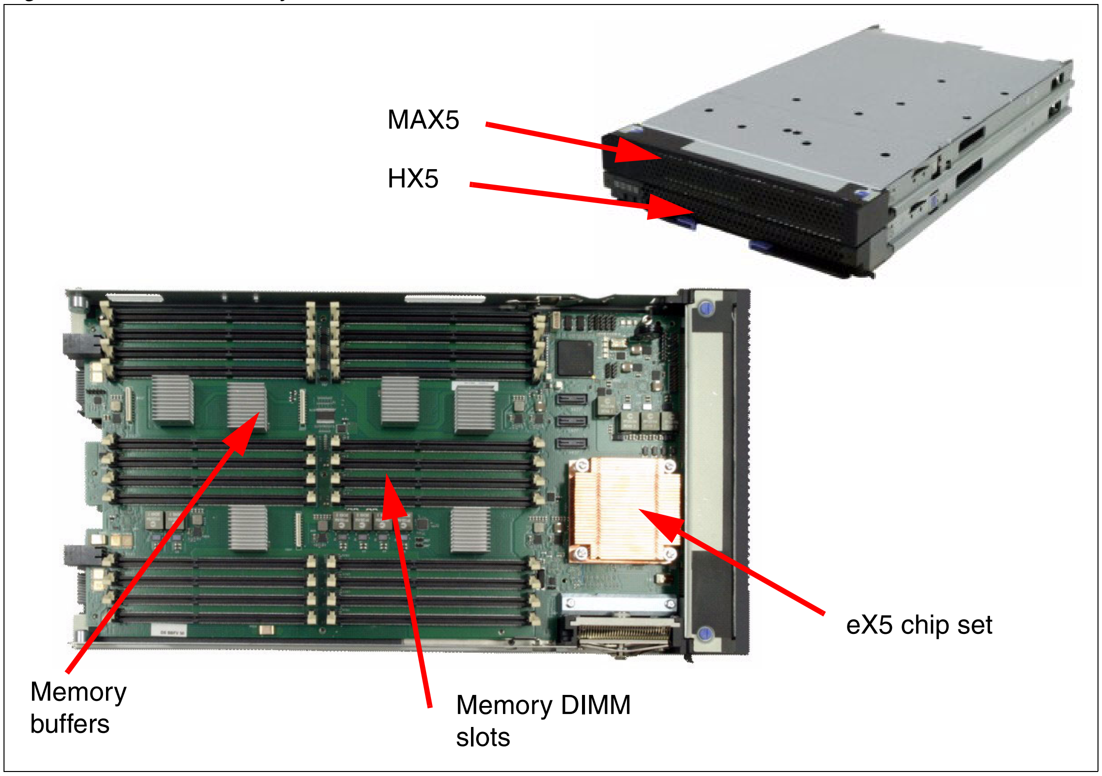


Figure 5-2 Layout of HX5 (showing a 2-node 4-socket configuration)

The MAX5 memory expansion blade, which is shown in Figure 5-3 on page 180, is a device with the same dimensions as the HX5. When the MAX5 is attached to the HX5, the combined unit occupies two blade bays in the BladeCenter chassis. The MAX5 cannot be removed separately from the HX5.

Figure 5-3 MAX5 memory slots



5.1.1 Comparison to the HS22 and HS22V

The BladeCenter HS22 is a general-purpose 2-socket blade server and the HS22V is a virtualization blade offering. Table 5-2 compares the HS22 and HS22V with the HX5 offerings.

Table 5-2 HX5 comparison to HS22 and HS22V

Feature	HS22	HS22V	HX5	HX5 with MAX5
Form factor	30 mm blade (1-wide)	30 mm blade (1-wide)	30 mm blade (1-wide) 60 mm blade (2-wide)	60 mm blade (2-wide)
Processor	Intel Xeon Processor 5500 and 5600	Intel Xeon Processor 5500 and 5600	Intel Xeon Processor 6500 or 7500 (Nehalem EX)	Intel Xeon Processor 6500 or 7500 (Nehalem EX)
Maximum number of processors	Two	Two	30 mm blade: two 60 mm blade: four	Two
Number of cores	2, 4, or 6 cores	2, 4, or 6 cores	4, 6, or 8 cores	4, 6, or 8 cores
Cache	4 MB or 8 MB	8 MB	12 MB, 18 MB, or 24 MB (shared between cores)	
Memory Speed	Up to 1333 MHz	Up to 1333 MHz	978 or 800 MHz (scalable memory interconnects (SMI) link-speed dependent)	HX5: Up to 978 MHz MAX5: Up to 1066MHz (SMI link-speed dependent)

Feature	HS22	HS22V	HX5	HX5 with MAX5
DIMMs per channel	Two	Three	One	HX5: One MAX5: Two
DIMM sockets	12	18	30 mm: 16 60 mm: 32	40
Maximum installable RAM (8 GB DIMMs)	96 GB	144 GB	30 mm: 128 GB 60 mm: 256 GB	320 GB
Memory type	DDR3 ECC VLP RDIMMs	DDR3 ECC and non-ECC VLP RDIMMs	DDR3 ECC VLP RDIMMs	DDR3 ECC VLP RDIMMs
Internal disk drives	2x Hot-swap 2.5” drive SAS, SATA, or SSD	2x Non-hot-swap 1.8” SSD	Two or four non-hot-swap 1.8” SSDs (require the SSD Expansion Card)	Two non-hot-swap 1.8” SSDs (require the SSD Expansion Card)
I/O expansion	► One CIOv ► One CFFh	► One CIOv ► One CFFh	Per 30mm blade: ► One CIOv ► One CFFh	Per 60mm blade: ► One CIOv ► One CFFh
serial-attached SCSI (SAS) controller	Onboard LSI 1064 Optional ServeRAID MR10ie (CIOv)	Onboard LSI 1064 Optional ServeRAID MR10ie (CIOv)	LSI 1064 controller on the optional SSD Expansion Card	LSI 1064 controller on the optional SSD Expansion Card
Embedded Hypervisor	Internal USB socket for VMware ESXi	Internal USB socket for VMware ESXi	Internal USB socket for VMware ESXi	Internal USB socket for VMware ESXi
Onboard Ethernet	Broadcom 5709S	Broadcom 5709S	Broadcom 5709S	Broadcom 5709S
Chassis supported	► BladeCenter E (certain restrictions) ► BladeCenter H ► BladeCenter S ► BladeCenter HT		► BladeCenter H ► BladeCenter S ► BladeCenter HT (ac model only)	

5.2 Target workloads

The HX5 is designed for business-critical workloads, such as database and virtualization.

Virtualization provides many benefits, including improved physical resource utilization, improved hardware efficiency, and reduced power and cooling expenses. Server consolidation helps reduce the cost of overall server management and the number of assets that have to be tracked by a company or department.

Virtualization and server consolidation can provide the following benefits:

- Reduce the rate of physical server proliferation
- Simplify infrastructure
- Improve manageability
- Lower the total cost of IT, including power and cooling costs

The HX5 2-socket and HX5 4-socket are strong database systems. They are ideal upgrade candidates for database workloads that are already on a blade. The multicore processors, large memory capacity, and I/O options make the HX5 proficient at taking on database workloads that are being transferred to the blade form factor.

5.3 Chassis support

The HX5 is supported in BladeCenter chassis S, H, and HT, as listed in Table 5-3.

Table 5-3 HX5 chassis compatibility (BC is BladeCenter)

Description	BC-E 8677	BC-S 8886	BC-H 8852	BC-HT ac 8750	BC-HT dc 8740
HX5 server	No	Yes	Yes[a]	Yes	No[b]
HX5+MAX5 server	No	Yes	Yes[a]	Yes	No[b]

a. One-node and 2-node HX5 configurations with 130W processors are not supported in chassis with standard cooling modules. See Table 5-4.
b. Support for the BC-HT dc model can be granted for specific configurations with the SPORE process.

The number of HX5 servers supported in each chassis depends on the thermal design power of the processors that are used in the HX5 servers. Table 5-4, which uses the following conventions, shows the HX5 servers:

- A green square in a cell means the chassis can be filled with HX5 blade servers up to the maximum number of blade bays in the chassis (for example, 14 blades in the BladeCenter H).
- A yellow square in a cell means that the maximum number of HX5 blades that the chassis can hold is fewer than the total available blade bays (for example, 12 in a BladeCenter H). *All other bays must remain empty*. The empty bays must be distributed evenly between the two power domains of the chassis (bays 1 - 6 and bays 7 - 14).

Table 5-4 HX5 chassis compatibility

Server	Thermal design power (TDP) of the CPUs	Maximum number of servers supported in each chassis						
		BC-S (8886)	BC-H (models other than 4Tx)				BC-H (-4Tx)	BC-HT AC (8750)
			2900W supplies		2980W supplies[a]		2980W	
			Std. blower	Enh. blower[b]	Std. blower	Enh. blower[b]	Enh. blower[b]	
HX5 1-node (30 mm)	95W, 105W	5	14	14	14	14	14	10
	130W	4	None[c]	10	None[c]	12	12	8
HX5 2-node (60 mm)	95W, 105W	2	7	7	7	7	7	5
	130W	2	None[c]	5	None[c]	6	6	4
HX5 1-socket + MAX5 (60 mm)	95W, 105W	2	7	7	7	7	7	5
	130W	2	6	6	7	7	7	5

a. IBM BladeCenter H 2980W AC Power Modules, 68Y6601 (standard in 4Tx, optional with all other BC-H chassis models)
b. IBM BladeCenter H Enhanced Cooling Modules, 68Y6650 (standard in 4Tx, optional with all other BC-H chassis models)
c. Not supported

Network Equipment Building System (NEBS): The HX5 is currently not a NEBS-compliant offering in the BC-HT.

5.4 Models

The base models of the BladeCenter HX5, with and without the MAX5 memory expansion blade, are shown in Table 5-5. In the table, *Opt* indicates optional and *Std* indicates standard.

Table 5-5 Models of the HX5

Model[a]	Intel Xeon model and cores	Clock speed	TDP	HX5 max memory speed	MAX5 memory speed	MAX5	Scale to four socket	10 GbE card[b]	Standard memory[c]
7872-42x	1x E7520 4C	1.86 GHz	95W	800 MHz	800 MHz	Opt	Yes	Opt	2x 4 GB
7872-82x	1x L7555 8C	1.86 GHz	95W	978 MHz	978 MHz	Opt	Yes	Opt	2x 4 GB
7872-61x	1x E7530 6C	1.86 GHz	105W	978 MHz	978 MHz	Opt	Yes	Opt	2x 4 GB
7872-64x	1x E7540 6C	2.00 GHz	105W	978 MHz	1066 MHz	Opt	Yes	Opt	2x 4 GB
7872-65x	1x E7540 6C	2.00 GHz	105W	978 MHz	1066 MHz	Opt	Yes	Std	2x 4 GB
7872-63x	2x E6540 6C	2.00 GHz	105W	978 MHz	1066 MHz	Std	No	Opt	HX5: 4x 4GB MAX5: None
7872-6Dx	2x E6540 6C	2.00 GHz	105W	978 MHz	1066 MHz	Std	No	Std	HX5: 4x 4GB MAX5: None
7872-83x	2x X6550 8C	2.00 GHz	130W	978 MHz	1066 MHz	Std	No	Opt	HX5: 4x 4GB MAX5: None
7872-84x	2x X7560 8C	2.26 GHz	130W	978 MHz	1066 MHz	Std	No	Opt	HX5: 4x 4GB MAX5: None
7872-86x	1x X7560 8C	2.26 GHz	130W	978 MHz	1066 MHz	Opt	Yes	Std	2x 4 GB
7872-E8x	1x X6550 8C	2.26 GHz	130W	978 MHz	1066 MHz	Opt	No	Std	2x 4 GB
7872-E6x	1x X7540 6C	2.00 GHz	105W	978 MHz	1066 MHz	Opt	Yes	Std	2x 4 GB

a. This column lists worldwide, generally available variant (GAV) model numbers. They are not orderable as listed and must be modified by country. The US GAV model numbers use the following nomenclature: *xx*U. For example, the US orderable part number for 7870-A2*x* is 7870-A2U. See the product-specific official IBM announcement letter for other country-specific GAV model numbers.
b. Emulex Virtual Fabric Adapter Expansion Card (CFFh).
c. The HX5 has 16 DIMM sockets and can hold 128 GB using 8 GB memory DIMMs. The MAX5 has 24 DIMM sockets and can hold 192 GB using 8 GB memory DIMMs. A 1-node HX5 + MAX5 support 320 GB total using 8 GB DIMMs.

Also available is a virtualization workload-optimized model of the HX5. This model is a preconfigured, pretested model that is targeted at large-scale consolidation. Table 5-6 on page 184 shows the model.

Table 5-6 Workload-optimized models of the HX5

Model	Intel Xeon model and cores/max	Clock speed	TDP	HX5 max memory speed[a]	MAX5	Scalable to four socket	10GbE card[b]	Standard memory (max 320 GB)[c]
Virtualization workload-optimized models (includes VMware ESXi 4.1 on a USB memory key)								
7872-68x	2x E6540 6C/2	2.00 GHz	105 W	978 MHz	Std	No	Std	160 GB HX5: 16x 4GB MAX5: 24x 4GB

a. Memory speed of the HX5 is dependent on the processor installed; however, the memory speed of the MAX5 is up to 1066 MHz irrespective of the processor installed in the attached HX5.
b. Emulex Virtual Fabric Adapter Expansion Card (CFFh).
c. HX5 has 16 DIMM sockets and can hold 128 GB using 8 GB memory DIMMs. MAX5 has 24 DIMM sockets and can hold 192 GB using 8 GB memory DIMMs. A 1-node HX5 + MAX5 support 320 GB total using 8 GB DIMMs.

Model 7872-68x is a virtualization-optimized model and includes the following features in addition to standard HX5 and MAX5 features:

- Forty DIMM sockets, all containing 4 GB memory DIMMs for a total of 160 GB of available memory.
- VMware ESXi 4.1 on a USB memory key is installed internally in the server. See 5.15, “Integrated virtualization” on page 214 for details.
- Emulex Virtual Fabric Adapter Expansion Card (CFFh).

5.5 System architecture

The Intel Xeon 6500 and 7500 processors in the HX5 have up to eight cores with 16 threads per socket. The processors have up to 24 MB of shared L3 cache, Hyper-Threading, several with Turbo Boost, four QuickPath Interconnect (QPI) links, one integrated memory controller, and up to four buffered SMI channels.

The HX5 2-socket server has the following system architecture features as standard:

- Two 1567-pin land grid array (LGA) processor sockets
- Intel 7500 “Boxboro” chip set
- Intel ICH10 south bridge
- Eight Intel Scalable Memory Buffers, each with two memory channels
- One DIMM per memory channel
- 16 DDR3 DIMM sockets
- One Broadcom BCM5709S dual-port Gigabit Ethernet controller
- One Integrated Management Module (IMM)
- One Trusted Platform Module 1.2 Controller
- One PCI Express x16 CFFh I/O expansion connector
- One PCI Express x16 CFFh-style connector for use with the SSD Expansion Card and one or two solid-state drives
- One CIOv I/O expansion connector
- Scalability connector
- One internal USB port for embedded virtualization

Figure 5-4 shows the HX5 block diagram.

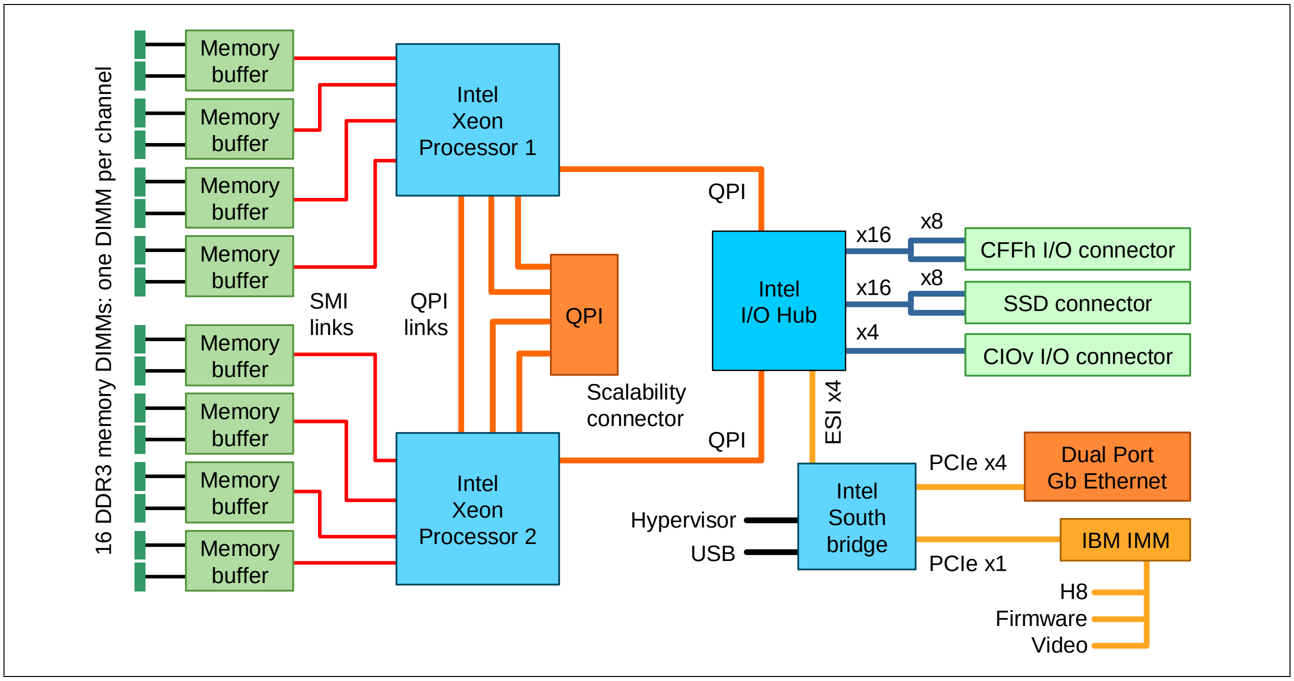


Figure 5-4 HX5 block diagram

5.6 Speed Burst Card

To increase performance in a 2-socket HX5 server (that is, with two processors installed), install the IBM HX5 1-Node Speed Burst Card. The 1-Node Speed Burst Card takes the QPI links that typically are used for scaling two HX5 2-socket blades and routes them back to the processors on the same blade. Table 5-7 lists the ordering information.

Table 5-7 HX5 1-Node Speed Burst Card

Part number	Feature code	Description
59Y5889	1741	IBM HX5 1-Node Speed Burst Card

Figure 5-5 on page 186 shows a block diagram of the Speed Burst Card attachment to the system.

Speed Burst Card: The Speed Burst Card is not required for an HX5 with only one processor installed. It is also not needed for a 2-node configuration (a separate card is available for a 2-node configuration, as described in 5.8, "Scalability" on page 188).

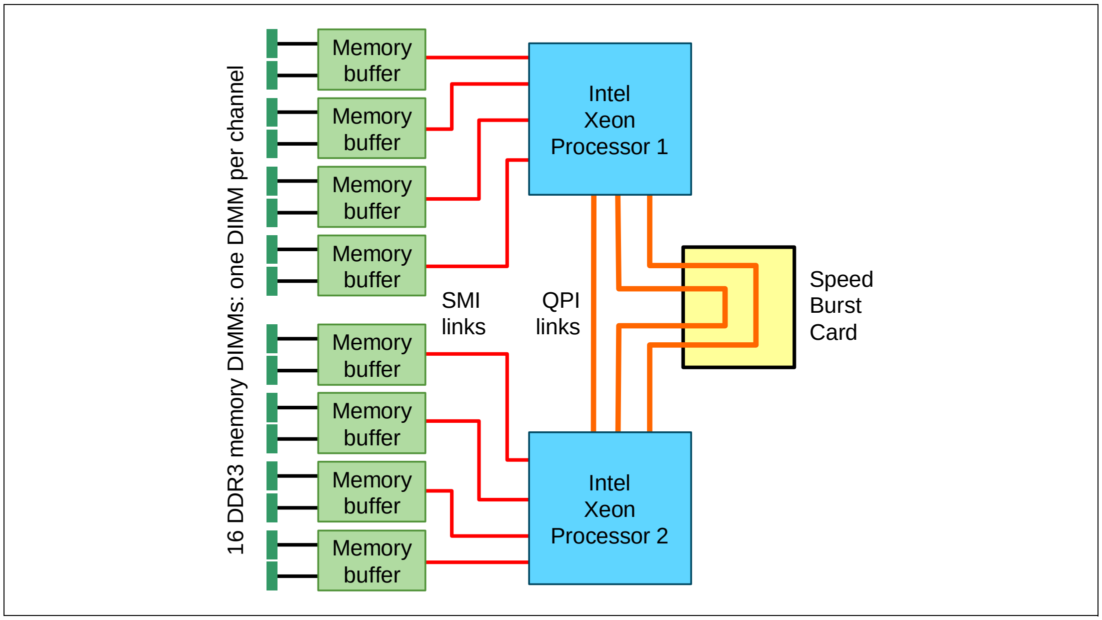


Figure 5-5 HX5 1-Node Speed Burst Card block diagram

Figure 5-6 shows where the Speed Burst Card is installed on the HX5.

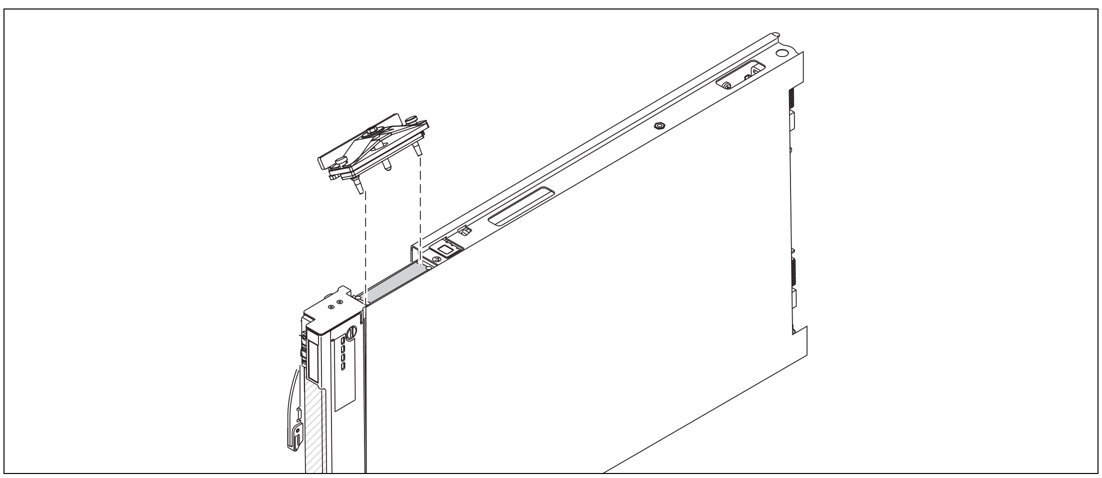


Figure 5-6 Installing the Speed Burst Card

5.7 IBM MAX5 for BladeCenter

IBM MAX5 for BladeCenter, which is shown in Figure 5-3 on page 180, is a memory expansion blade that attaches to HX5 2-socket blade servers. It has the following system architecture features:

- IBM EXA memory controller
- Twenty-four DIMM slots and 6 memory buffers
- Four DIMM slots per memory buffer (two per channel)
- VLP DDR3 memory in 4 GB and 8 GB capacities, 1333 MHz

- Attachment to a single HX5 using the IBM HX5 MAX5 1-node Scalability kit, part number 59Y5877, as described in 5.8.3, “HX5 with MAX5” on page 190
- Communication with the processors on the HX5 using high-speed QPI links

MAX5 is standard with certain models, as listed in 5.4, “Models” on page 183. For other models, MAX5 is available as an option, as listed in Table 5-8.

Table 5-8 IBM MAX5 for BladeCenter

Part number	Feature code	Description
46M6973	1740	IBM MAX5 for BladeCenter
59Y5877	1742	IBM HX5 MAX5 1-node Scalability kit

MAX5 consists of the EX5 node controller chip, six memory buffers, and 24 DIMM sockets. The MAX5 has three power domains: A, B, and C. Each power domain includes two memory controllers and eight DIMM sockets. Figure 5-7 shows the layout of the MAX5.

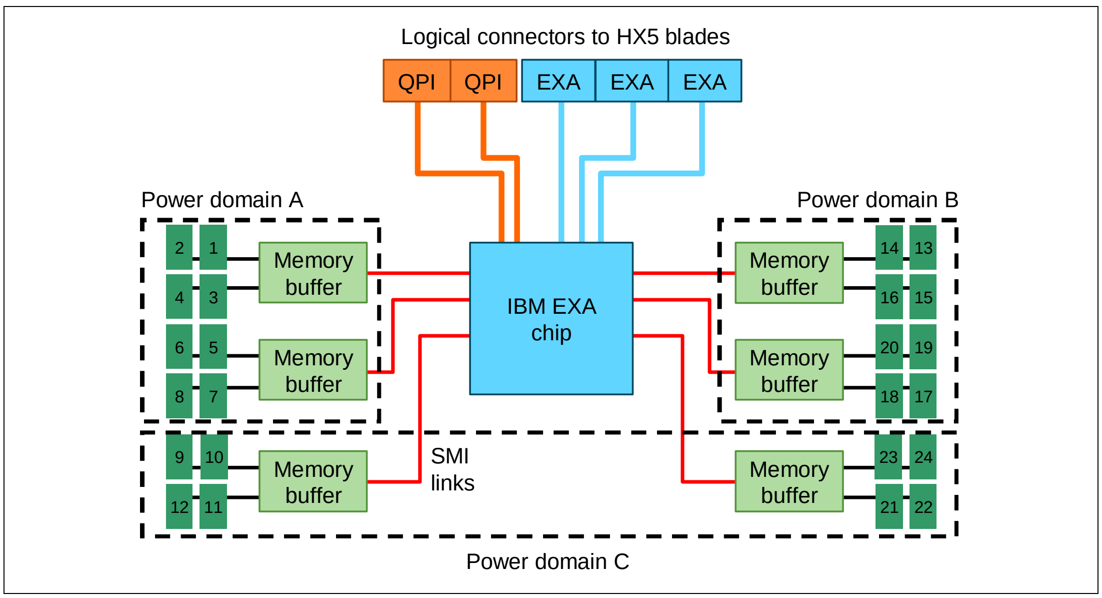


Figure 5-7 MAX5 memory expansion blade

In the next section, 5.8, “Scalability” on page 188, we describe how the MAX5 connects to the HX5. We explain the memory options and rules in 5.10, “Memory” on page 194.

5.8 Scalability

This section explains how the HX5 can be expanded to increase the number of processors and the number of memory DIMMs.

The HX5 blade architecture allows for a number of scalable configurations, including the use of a MAX5 memory expansion blade, but the blade currently supports three configurations:

- A single HX5 server with two processor sockets. This server is a standard 30 mm blade, which is also known as *single-wide* server or *single-node* server.
- Two HX5 servers connected to form a single image 4-socket server. This server is a 60 mm blade, which is also known as a *double-wide* server or *2-node* server.
- A single HX5 server with two processor sockets, plus a MAX5 memory expansion blade attached to it, resulting in a 60 mm blade configuration. This configuration is sometimes referred to as a *1-node+MAX5* configuration.

We describe each configuration in the following sections. We list the supported BladeCenter chassis for each configuration in 5.3, “Chassis support” on page 182.

5.8.1 Single HX5 configuration

This server is the base configuration and supports one or two processors that are installed in the single-wide 30 mm server.

When the server has two processors installed, ensure that the server has the Speed Burst Card installed for maximum performance, as described in 5.6, “Speed Burst Card” on page 185. This card is not required but strongly suggested.

5.8.2 Double-wide HX5 configuration

In the 2-node configuration, the two HX5 servers are physically connected and a 2-node scalability card is attached to the side of the blades, which provides the path for the QPI scaling.

Each node can have one or two processors installed (that is, 2-node configurations with a total of two processors or four processors are supported). All installed processors must be identical, however.

The two servers are connected using a 2-node scalability card, as shown in Figure 5-8 on page 189. The scalability card is immediately adjacent to the processors and provides a direct connection between the processors in the two nodes.

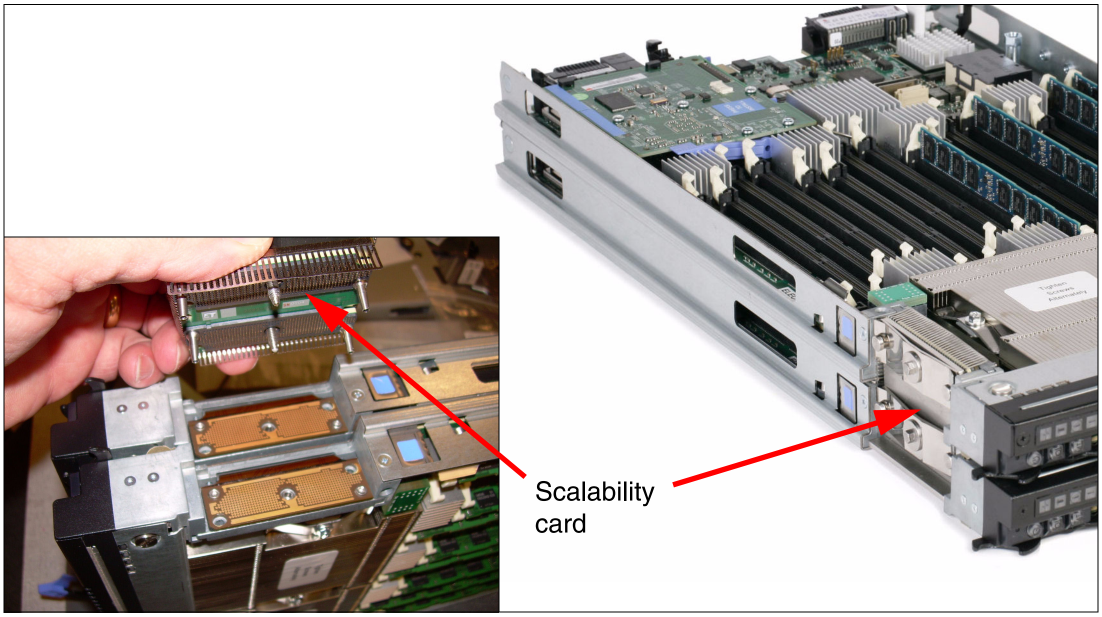


Figure 5-8 2-node HX5 with the 2-node scalability card indicated

The double-wide configuration consists of two connected HX5 servers. This configuration consumes two blade slots and has the 2-node scalability card attached. The scaling is performed through QPI scaling. The 2-node scalability card is not included with the server and must be ordered separately, as listed in Table 5-9.

Table 5-9 HX5 2-Node Scalability Kit

Part number	Feature code	Description
46M6975	1737	IBM HX5 2-Node Scalability Kit

The IBM HX5 2-Node Scalability Kit contains the 2-node scalability card, plus the necessary hardware to physically attach the two HX5 servers to each other.

Figure 5-9 shows the block diagram of a 2-node HX5 and the location of the HX5 2-node scalability card.

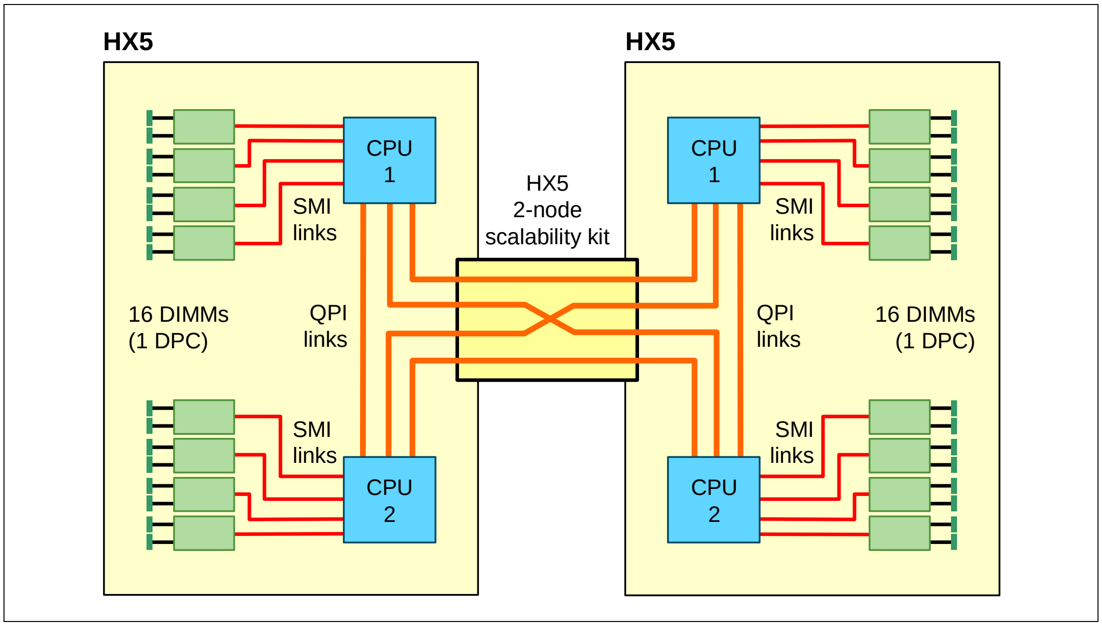


Figure 5-9 Block diagram of a 2-node HX5

Ensure that all firmware is up-to-date before attaching the blades together. For the minimum firmware requirements, see 8.2.3, “Required firmware of each blade and the AMM” on page 379.

For configuring software when using a 2-node for creating the partition in a scalable complex, using Flexnode to toggle between a single blade and a 2-node, and deleting the partition, see 8.6, “Creating an HX5 scalable complex” on page 402.

Important: When a blade is attached, it does not automatically become a 2-node single-image system. You must create the 2-node single-image system by using the scalable complex in the Advanced Management Module (AMM) in the chassis. See 8.6, “Creating an HX5 scalable complex” on page 402 for more information.

5.8.3 HX5 with MAX5

In the HX5 and MAX5 configuration, the HX5 and MAX5 units connect through a 1-node MAX5 scalability card, which provides QPI scaling. See Figure 5-10 on page 191.

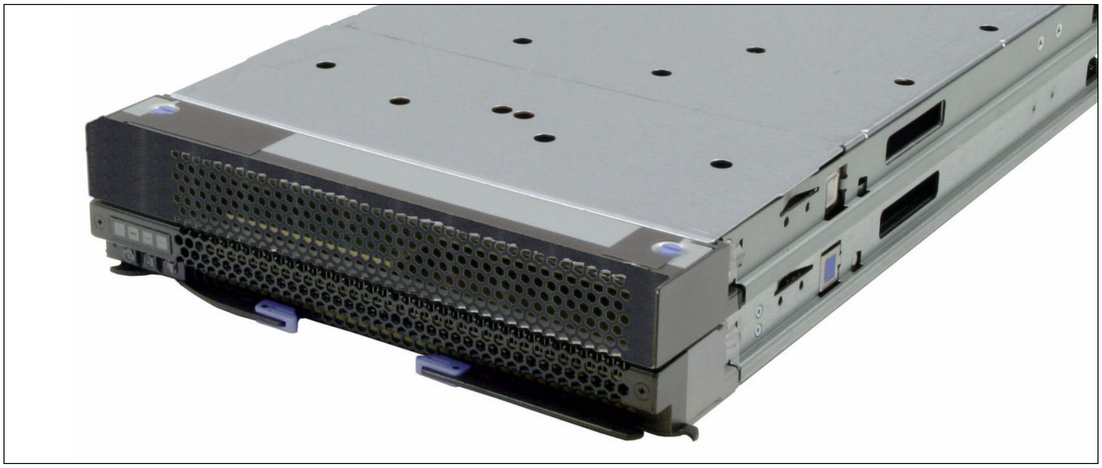


Figure 5-10   Single-node HX5 + MAX5

The card that is used to connect the MAX5 to the HX5 is the IBM HX5 MAX5 1-Node Scalability Kit, which is extremely similar in physical appearance to the 2-Node Scalability Kit that was shown in Figure 5-8 on page 189. Table 5-10 lists the ordering information.

Table 5-10   HX5 1-Node Scalability Kit

Part number	Feature code	Description
59Y5877	1742	IBM HX5 MAX5 1-Node Scalability Kit

Figure 5-11 shows the block diagram of the single-node HX5 with MAX5.

Important: The MAX5 can be connected only to a single HX5 server. A configuration of two MAX5 units connected to a 2-node HX5 is not supported.

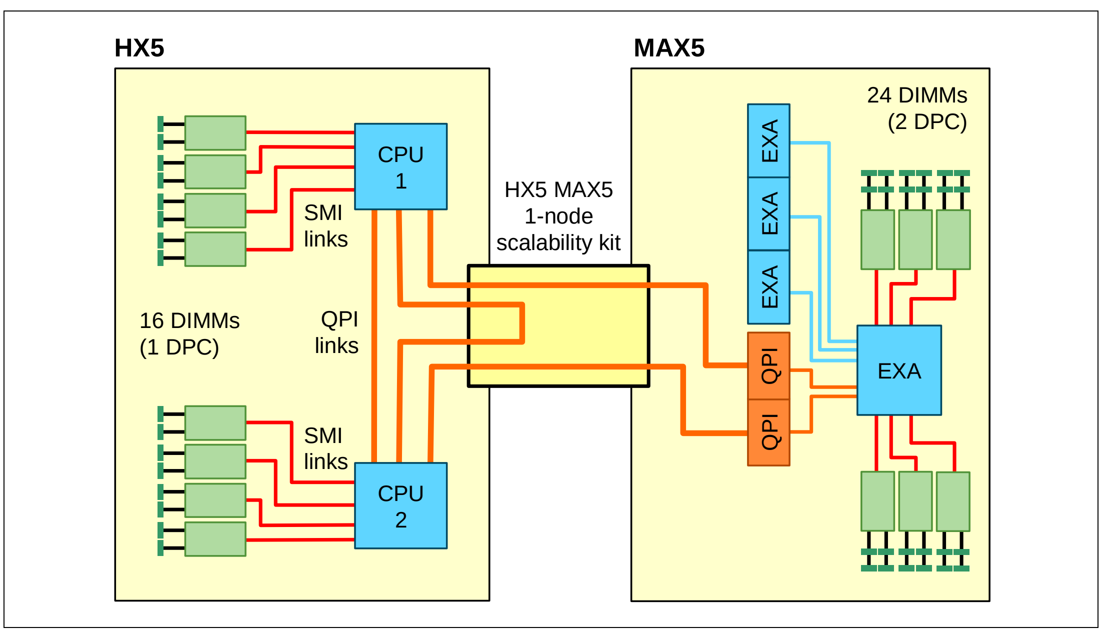


Figure 5-11   HX5 1-node with MAX5 block diagram

Having only one processor installed in the HX5 instead of two processors is supported; however, the recommendation is for two processors to maximize memory performance.

Ensure that all firmware is up-to-date before attaching the MAX5 to the blade. For minimum firmware requirements, see 8.2.3, “Required firmware of each blade and the AMM” on page 379.

When inserting an HX5 with MAX5, there is no partition information to set up in a scalable complex. MAX5 ships ready to use when attached.

5.9 Processor options

The HX5 type 7872 supports Intel Xeon 6500 and 7500 quad-core, 6-core, or 8-core processors. The Intel Xeon processors are available in various clock speeds and have standard and low-power offerings. To see a list of processor features that are included in the Intel Xeon 6500 and 7500 series, see 2.2, “Intel Xeon 6500 and 7500 family processors” on page 16.

Table 5-11 lists the processor options for the HX5 and the supported models.

Table 5-11 Available processor options

Part number	Feature code[a]	Description	Supported model
46M6955	4571/4572	Intel Xeon E6540 Processor, 2.00 GHz, 6C, 105W, 18M, 6.4 GT/s QPI	63x, 6Dx, 68x
46M6863	4558/4559	Intel Xeon E7520 Processor, 1.86 GHz, 4C, 95W, 18M, 4.8 GT/s QPI	42x
59Y5899	4564/4565	Intel Xeon E7530 Processor, 1.86 GHz, 6C, 105W, 12M, 5.86 GT/s QPI	61x
59Y5859	4566/4568	Intel Xeon E7540 Processor, 2.00 GHz, 6C, 105W, 18M, 6.4 GT/s QPI	64x and 65x
46M6873	4562/4563	Intel Xeon L7555 Processor, 1.86 GHz, 8C, 95W, 24M, 5.86 GT/s QPI	82x
46M6995	4573/4574	Intel Xeon X6550 Processor, 2.00 GHz, 8C, 130W, 18M, 6.4 GT/s	83x
59Y5904	4577/4578	Intel Xeon X7542 Processor, 2.66 GHz, 6C, 130W, 18M, 5.86 GT/s QPI	CTO only
59Y5909	4579/4580	Intel Xeon X7550 Processor, 2.00 GHz, 8C, 130W, 18M, 6.4 GT/s QPI	CTO only
46M6960	4575/4576	Intel Xeon X7560 Processor, 2.26 GHz, 8C, 130W, 24M, 6.4 GT/s QPI	84x and 86x

a. The first feature code is for the first processor. The second feature code is for the second processor.

Table 5-12 lists the capabilities of each processor option that is available for the HX5.

Table 5-12 Intel Xeon 6500 and 7500 features

Processor model/cores	Scalable to four socket	Processor frequency	Turbo[a]	HT[b]	L3 cache	Power	QPI speed	HX5 memory speed	MAX5 memory speed
Standard processors (E)									
Xeon E6540 6C	No	2.0 GHz	Yes +2	Yes	18 MB	105 W	6.4 GT/s	978 MHz	1066 MHz
Xeon E7520 4C	Yes	1.86 GHz	No	Yes	18 MB	95 W	4.8 GT/s	800 MHz	800 MHz
Xeon E7530 6C	Yes	1.86 GHz	Yes +2	Yes	12 MB	105 W	5.86 GT/s	978 MHz	978 MHz
Xeon E7540 6C	Yes	2.0 GHz	Yes +2	Yes	18 MB	105 W	6.4 GT/s	978 MHz	1066 MHz

Processor model/cores	Scalable to four socket	Processor frequency	Turbo[a]	HT[b]	L3 cache	Power	QPI speed	HX5 memory speed	MAX5 memory speed
Low-power processors (L)									
Xeon L7555 8C	Yes	1.86 GHz	Yes +2	Yes	24 MB	95 W	5.86 GT/s	978 MHz	978 MHz
Advanced processors (X)									
Xeon X6550 8C	No	2.0 GHz	Yes +3	Yes	18 MB	130 W	6.4 GT/s	978 MHz	1066 MHz
Xeon X7542 6C	Yes	2.66 GHz	Yes +3	No	18 MB	130 W	5.86 GT/s	978 MHz	978 MHz
Xeon X7550 8C	Yes	2.0 GHz	Yes +3	Yes	18 MB	130 W	6.4 GT/s	978 MHz	1066 MHz
Xeon X7560 8C	Yes	2.26 GHz	Yes +3	Yes	24 MB	130 W	6.4 GT/s	978 MHz	1066 MHz

a. Intel Turbo Boost technology. The number that is listed is the multiple of 133 MHz by which the processor base frequency can be increased. For example, if the base frequency is 2.0 GHz and the Turbo value is +2, the frequency can increase to as high as 2.266 GHz.
b. Intel Hyper-Threading technology.

Xeon E6510: As shown in Table 5-12, the Xeon E6510 processor does not support scaling to four sockets, and it does not support the MAX5. This limit is a technical limitation of this particular processor.

Follow these processor configuration rules:

- All installed processors must be identical.
- In a 2-node configuration, both two and four processors are supported. That is, each node can have one or two processors installed. All processors must be identical.
- If you only have two processors installed (two processors in a single-node, or one processor in each node in a 2-node), Xeon 6500 series processors are supported. However, you cannot add two additional processors later to form a 4-processor system, because the Xeon 6500 series processors do not support 4-way.
- A MAX5 configuration (HX5 server with MAX5 attached) supports one or two installed processors.

Memory speed in the HX5 depends on the SMI link of the processor (the SMI link speed is listed in Table 5-12 on page 192 as a GT/s value). It also depends on the limits of the low-power scalable memory buffers that are used in the HX5:

- If the SMI link speed is 6.4 GT/s or 5.86 GT/s, the memory in the HX5 can operate at a maximum of 978 MHz.
- If the SMI link speed is 4.8 GT/s, the memory in the HX5 can operate at a maximum of 800 MHz.

Table 5-12 on page 192 indicates these memory speeds.

If a MAX5 memory expansion blade is installed, the memory in the MAX5 can operate as high as 1066 MHz, depending on the DIMMs installed. The MAX5 memory speed is independent of the HX5 memory speed.

For more information about calculating memory speed, see 2.3.1, “Memory speed” on page 22.

5.10 Memory

The HX5 2-socket has eight DIMM sockets per processor (a total of 16 DIMM sockets), supporting up to 128 GB of memory when using 8 GB DIMMs. With the addition of the MAX5 memory expansion blade, a single HX5 blade has access to a total of 40 DIMM sockets supporting up to 320 GB of RAM when using 8 GB DIMMs.

The HX5 and MAX5 use registered Double Data Rate 3 (DDR3), very low profile (VLP) DIMMs and provide reliability, availability, and serviceability (RAS) and advanced Chipkill memory protection. For more information about Chipkill memory protection, see “Chipkill” on page 29. For information about RAS, see 2.3.6, “Reliability, availability, and serviceability (RAS) features” on page 28.

This section has the following topics:

- 5.10.1, “Memory options” on page 194
- 5.10.2, “DIMM population order” on page 196
- 5.10.3, “Memory balance” on page 199
- 5.10.4, “Memory mirroring” on page 200
- 5.10.5, “Memory sparing” on page 202

To see a full list of the supported memory features, such as Hemisphere Mode, Chipkill, nonuniform memory access (NUMA), and memory mirroring, and an explanation of each memory feature, see 2.3, “Memory” on page 22.

5.10.1 Memory options

Table 5-13 lists the memory options for the HX5 and MAX5. They are the same options for both products *with the exception* of the 2 GB DIMM option.

> **Tip:** Memory must be installed in pairs of two identical DIMMs, or in quads if memory mirroring is enabled. The options in Table 5-13, however, are for single DIMMs.

Table 5-13 Memory options HX5 and MAX5

Part number	FC	Capacity	Support in MAX5	Description	Rank	Speed[a]
44T1486	1916	1x 2 GB	No	2GB PC3-10600 CL9 ECC VLP (2Rx8, 1.5V, 1Gb)	Dual	1333 MHz
44T1596	1908	1x 4 GB	Yes	4GB PC3-10600 CL9 ECC VLP (2Rx8, 1.5V, 2Gb)	Dual	1333 MHz
46C7499	1917	1x 8 GB	Yes	8GB PC3-8500 CL7 ECC VLP (4Rx8, 1.5V, 2Gb)	Quad	1066 MHz
49Y1554	A13Q	1x 8 GB	Yes[b]	8GB PC3-10600 CL9 ECC (1x8GB, 2Rx4, 1.5V 2Gb)	Dual	1333 MHz

a. Although the speed of the supported memory DIMMs is as high as 1333 MHz, the actual memory bus speed is a function of the processor and the memory buffers used in the HX5 server. In the HX5, memory speed is up to 978 MHz and in the MAX5, memory speed is up to 1066 MHz. See Table 5-12 on page 192 for specifics.
b. This DIMM supports redundant bit steering (RBS) when used in the MAX5, as described in “Redundant bit steering” on page 29.

> **Two GB DIMM option:** The 2 GB DIMM option that is listed in Table 5-13 is not supported in the MAX5, because the MAX5 does not support mixing DIMMs with various DRAM technologies, such as 1 Gb versus 2 Gb.

For optimal performance, populate all DIMM slots on the HX5 before filling the MAX5.

Each processor controls eight DIMMs and four memory buffers in the server, as shown in Figure 5-12. To make use of all 16 DIMM sockets, you must install both processors. If only one processor is installed, you can only install eight DIMM sockets.

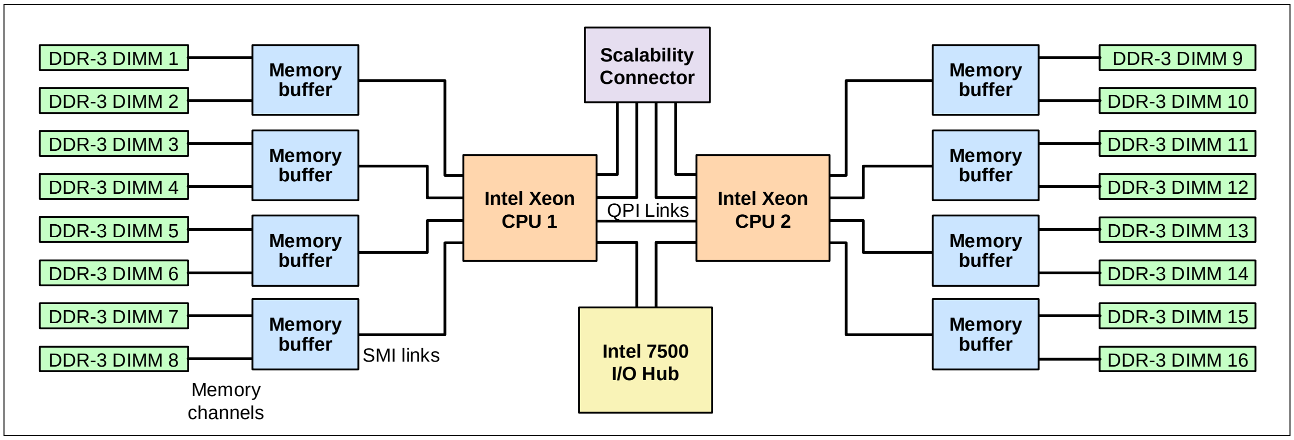


Figure 5-12 Portion of the HX5 block diagram showing the processors, memory buffers, and DIMMs

Figure 5-13 shows the physical locations of the 16 memory DIMM sockets.

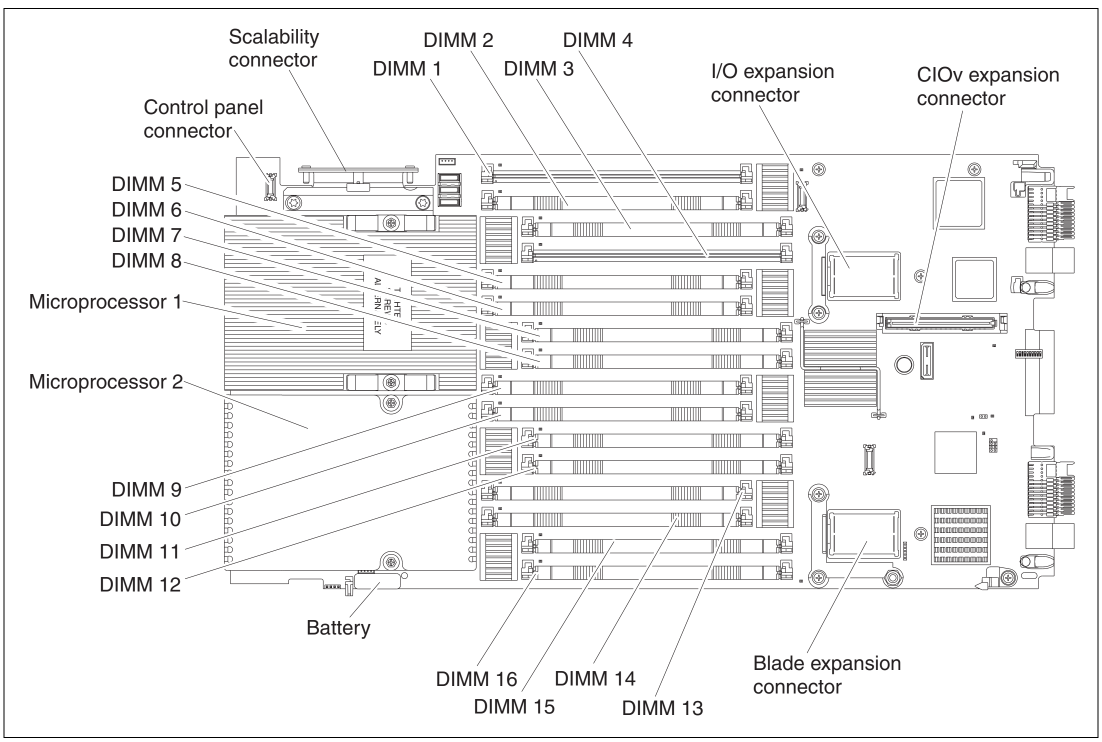


Figure 5-13 DIMM layout on the HX5 system board

The MAX5 memory expansion blade has 24 memory DIMM sockets, as shown in Figure 5-14 on page 196. The MAX5, which must be connected to an HX5 system (only the 1-node HX5 supports the MAX5), has one memory controller and six SMI-connected memory buffers.

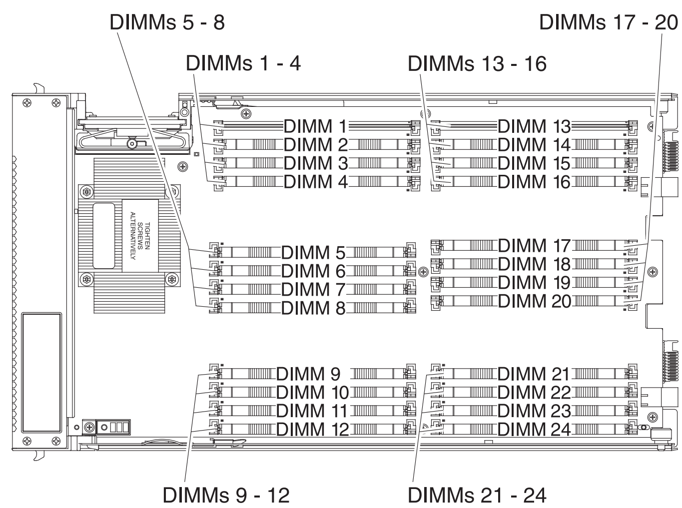


Figure 5-14 DIMM layout on the MAX5 system board

MAX5 memory runs at 1066, 978, or 800 MHz DDR3 speeds. The memory speed is dependent on the processor QPI speed in the HX5:

- A QPI speed of 6.4 GHz means the speed of the MAX5 memory is 1066 MHz.
- A QPI speed of 5.8 GHz means the speed of the MAX5 memory is 978 MHz.
- A QPI speed of 4.8 GHz means the speed of the MAX5 memory is 800 MHz.

Table 5-12 on page 192 indicates these memory speeds for each processor.

To see more information about how memory speed is calculated with QPI, see 2.3.1, "Memory speed" on page 22.

5.10.2 DIMM population order

Installing DIMMs in the HX5 and MAX5 in the correct order is essential for system performance. See 5.10.4, "Memory mirroring" on page 200 for the effects on performance when you do not install the DIMMs in the correct order.

HX5 memory population order

As shown in Figure 5-12 on page 195, the HX5 design has two DIMMs per memory buffer and one DIMM socket per memory channel.

For best performance, install the DIMMs in the sockets, as shown in Table 5-14 on page 197. This sequence spreads the DIMMs across as many memory buffers as possible.

Installation methods: These configurations use the most optimized method for performance. For optional installation methods, see the *BladeCenter HX5 Problem Determination and Service Guide* at the following website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084529

Table 5-14 NUMA-compliant DIMM installation for a single-node HX5

Number of CPUs	Number of DIMMs	Hemisphere Mode[a]	Processor 1								Processor 2							
			Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer	
			DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16
2	4	N	x			x					x			x				
2	8	Y	x			x	x			x	x			x	x			x
2	12	N	x	x	x	x	x			x	x	x	x	x	x			x
2	16	Y	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

a. For more information about Hemisphere Mode and its importance, see 2.3.5, "Hemisphere Mode" on page 26.

In a 2-node (4-socket) configuration with two HX5 servers, follow the memory installation sequence in both nodes. You must populate memory to have a balance for each processor in the configuration.

For best performance, use the following general guidelines:

- Install as many DIMMs as possible. You can get the best performance by installing DIMMs in every socket.
- Each processor needs to have identical amounts of RAM.
- Spread out the memory DIMMs to all memory buffers. That is, install one DIMM to a memory buffer before beginning to install a second DIMM to that same buffer. See Table 5-14 for DIMM placement.
- You must install memory DIMMs in the order of the DIMM size with largest DIMMs first, then next largest DIMMs, and so on. Placement must follow the DIMM socket installation that is shown in Table 5-14.
- To maximize performance of the memory subsystem, select a processor with the highest memory bus speed (as listed in Table 5-12 on page 192).

  The lower value of the processor's memory bus speed and the DIMM speed determine how fast the memory bus can operate. Every memory bus operates at this speed.

Table 5-15 NUMA-compliant DIMM installation for a 2-node HX5

Number of CPUs	Number of DIMMs	Hemisphere Mode[a]	Processor 1								Processor 2							
			Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer	
			DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16
4	8	N	x			x					x			x				
4	16	Y	x			x	x			x	x			x	x			x
4	24	N	x	x	x	x	x			x	x	x	x	x	x			x
4	32	Y	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

a. For more information about Hemisphere Mode and its importance, see 2.3.5, “Hemisphere Mode” on page 26

MAX5 memory population order

With the configuration of an HX5 connected to a MAX5, follow these rules:

- Install at least two DIMMs in the HX5 (four DIMMs if the HX5 has two installed processors).
- For the best memory performance, fully populate the HX5 by using the sequence that is listed in Table 5-14 on page 197, and then populate the MAX5 by using the sequence that is listed in Table 5-16 on page 199.
- The data widths for the following quads must match. For example, DIMMs in each quad must be all 4Rx8 or all 2Rx8. See Figure 5-15 for the block diagram and Figure 5-14 on page 196 for the physical location of these DIMMs.
  - DIMMs 1, 2, 7, and 8
  - DIMMs 3, 4, 5, and 6
  - DIMMs 13, 14, 17, and 18
  - DIMMs 15, 16, 19, and 20
  - DIMMs 9, 10, 21, and 22
  - DIMMs 11, 12, 23, and 24

  Based on the two DIMM options that are currently supported in the MAX5 (listed in Table 5-13 on page 194), this step means that all DIMMs in each of the quads listed here must be either 4 GB or 8 GB. You cannot mix 4 GB and 8 GB DIMMs in the same quad.
- Memory must be installed in matched pairs of DIMMs in the MAX5.
- Memory DIMMs must be installed in the order of DIMM size with largest DIMMs first. For example, if you plan to install both 4 GB and 8 GB DIMMs into the MAX5, use the population order that is listed in Table 5-16 on page 199. Install all 8 GB DIMMs first, and then install the 4 GB DIMMs.

The DIMM sockets in the MAX5 are arranged in three power domains (A, B, and C), as shown in Figure 5-15. Each power domain includes two memory controllers and eight DIMM sockets.

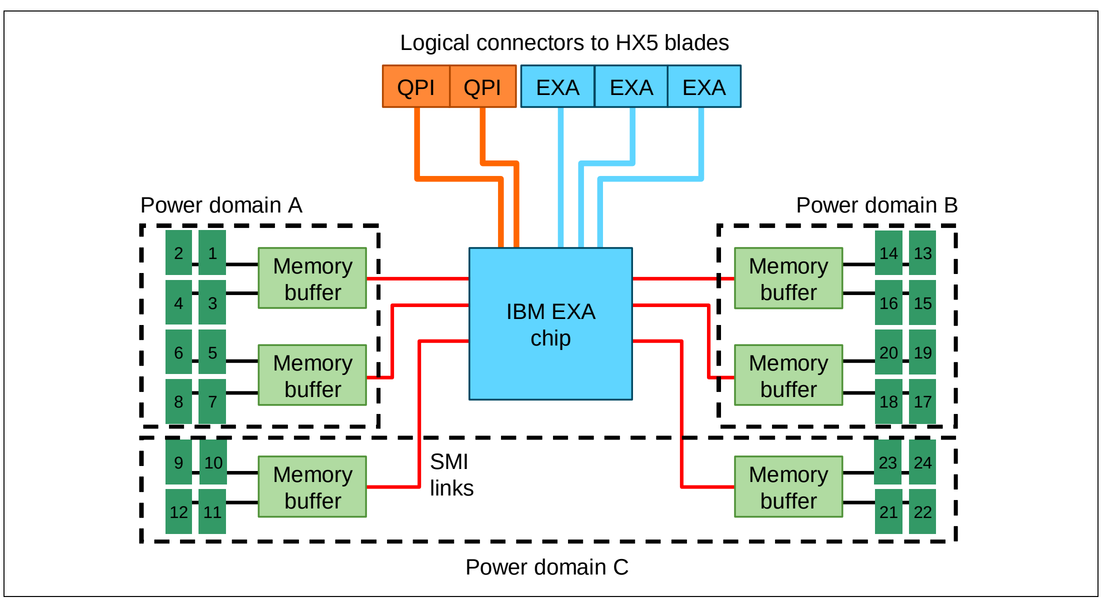


Figure 5-15 Power domains in the MAX5 memory expansion blade

This list shows the correct DIMMs in each power domain:

- Power domain A: 1 - 4 and 5 - 8
- Power domain B: 13 - 16 and 17 - 20
- Power domain C: 9 - 12 and 21 - 24

For the best memory performance, install the DIMMs by spreading them among all six memory buffers and all three power domains. Table 5-16 shows the installation order.

Table 5-16 DIMM installation for the MAX5 for IBM BladeCenter

Number of DIMMs	Power domain A								Domain C (½)				Power domain B								Domain C (½)			
	Buffer				Buffer				Buffer				Buffer				Buffer				Buffer			
	DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16	DIMM 17	DIMM 18	DIMM 19	DIMM 20	DIMM 21	DIMM 22	DIMM 23	DIMM 24
2	x							x																
4	x							x					x				x							
6	x							x				x	x				x							x
8	x		x			x		x				x	x				x							x
10	x		x			x		x				x	x		x		x		x					x
12	x		x			x		x		x		x	x		x		x		x			x		x
14	x	x	x			x	x	x		x		x	x		x		x		x			x		x
16	x	x	x			x	x	x		x		x	x	x	x		x	x	x			x		x
18	x	x	x			x	x	x		x	x	x	x	x	x		x	x	x			x	x	x
20	x	x	x	x	x	x	x	x		x	x	x	x	x	x		x	x	x			x	x	x
22	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x		x	x	x
24	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

MAX5 and VMware ESX: When using a MAX5 with VMware ESX 4.1 or ESXi 4.1, a boot parameter is required to access the MAX5 memory expansion unit. The MAX5 memory expansion unit utilizes NUMA technology, which needs to be enabled within the operating system. Without enabling NUMA technology, you might see the following message:

```
“The system has found a problem on your machine and cannot continue.
Interleaved Non-Uniform Memory Access (NUMA) nodes are not supported.”
```

See the RETAIN® tip H197190 for more information and the necessary parameters:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084842

5.10.3 Memory balance

The Xeon 7500 Series processor uses a non-uniform memory architecture (NUMA), as described in 2.3.4, “Nonuniform memory architecture (NUMA)” on page 26. Because NUMA is used, it is important to ensure that all memory controllers in the system are utilized by configuring all processors with memory. It is optimal to populate all processors in an identical fashion to provide a balanced system and populating all processors identically is also required by VMware.

Looking at Figure 5-16 as an example, Processor 0 has DIMMs populated, but no DIMMs are populated that are connected to Processor 1. In this case, Processor 0 has access to low-latency local memory and high-memory bandwidth. However, Processor 1 has access only to remote or "far" memory. So, threads executing on Processor 1 have a longer latency to access memory as compared to threads on Processor 0. This result is due to the latency penalty incurred to traverse the QPI links to access the data on the other processor's memory controller. The bandwidth to remote memory is also limited by the capability of the QPI links. The latency to access remote memory is more than 50% higher than local memory access.

For these reasons, we advise that you populate all of the processors with memory, remembering the necessary requirements to ensure optimal interleaving and Hemisphere Mode.

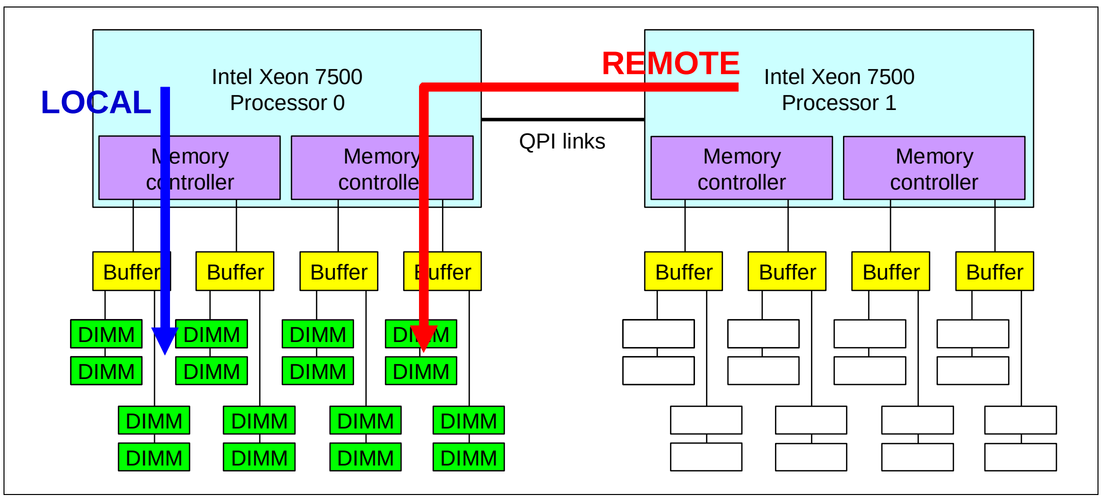


Figure 5-16 Memory latency when not spreading DIMMs across both processors

5.10.4 Memory mirroring

Memory mirroring is supported using HX5 and MAX5. On the HX5, when enabled, the first DIMM quadrant is duplicated onto the second DIMM quadrant for each processor. For a detailed understanding of memory mirroring, see "Memory mirroring" on page 28.

This section contains DIMM placements for each solution.

Important: If using memory mirroring, all DIMMs must be identical in size and rank.

DIMM placement: HX5

Table 5-17 on page 201 lists the DIMM installation sequence for memory-mirroring mode when one processor is installed.

Table 5-17 DIMM installation for memory mirroring: One processor

Number of processors	Number of DIMMs	Processor 1								Processor 2							
		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer	
		DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16
1	8	x	x	x	x	x	x	x	x								

Table 5-18 lists the DIMM installation sequence for memory-mirroring mode when two processors are installed.

Table 5-18 DIMM installation for memory mirroring: Two processors

Number of processors	Number of DIMMs	Processor 1								Processor 2							
		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer	
		DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16
1	8	x	x	x	x	x	x	x	x								
1	16	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

DIMM placement: MAX5

Table 5-19 lists the DIMM installation sequence in the MAX5 for memory-mirroring mode. Only power domains A and B are populated.

Table 5-19 DIMM installation for the MAX5 memory mirroring for IBM BladeCenter

Number of DIMMs	Power domain A								Domain C (½)				Power domain B								Domain C (½)			
	Buffer				Buffer				Buffer				Buffer				Buffer				Buffer			
	DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16	DIMM 17	DIMM 18	DIMM 19	DIMM 20	DIMM 21	DIMM 22	DIMM 23	DIMM 24
4	x							x					x				x							
8	x		x			x		x					x		x		x		x					
12	x	x	x			x	x	x					x	x	x		x	x	x					
16	x	x	x	x	x	x	x	x					x	x	x	x	x	x	x	x				

Domain C must be empty: Memory mirroring is only supported using two domains. You *must* remove all DIMMs from Domain C. If there is memory in Domain C, you get the following error in the AMM logs and all memory in the MAX5 is disabled:

“`Group 1, (memory device 1-40) (All DIMMs) memory configuration error`”

5.10.5 Memory sparing

The HX5 supports DIMM sparing, but only on the DIMMs that are installed in the HX5, not in the MAX5. For more information about memory sparing, see “Memory sparing” on page 29.

Table 5-20 shows the installation order when one processor is installed.

Sparing:

- Rank sparing is not supported on the HX5.
- MAX5 does not support rank sparing or DIMM sparing. Rank sparing or DIMM sparing works on an HX5 with a MAX5, but memory is only spared on the HX5.

Table 5-20 DIMM installation for the HX5 memory sparing: One processor

Number of processors	Number of DIMMs	Processor 1								Processor 2							
		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer	
		DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16
1	4	x	x	x	x												
1	8	x	x	x	x	x	x	x	x								

Table 5-21 shows the installation order when two processors are installed.

Table 5-21 DIMM installation for the HX5 memory sparing: Two processors

Number of processors	Number of DIMMs	Processor 1								Processor 2							
		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer	
		DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16
2	4	x	x	x	x												
2	8	x	x	x	x	x	x	x	x								
2	12	x	x	x	x	x	x	x	x	x	x	x	x				
2	16	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Redundant bit steering: Redundant bit steering (RBS) is not supported on the HX5 because the integrated memory controller of the Intel Xeon 7500 processors does not support the feature. See “Redundant bit steering” on page 29 for details.

The MAX5 memory expansion blade supports RBS, but only with x4 memory and not x8 memory. As shown in Table 5-13 on page 194, the 8 GB DIMM, part number 49Y1554, uses x4 DRAM technology. RBS is automatically enabled in the MAX5 memory port, if all DIMMs installed to that memory port are x4 DIMMs.

Mirroring or sparing effect on performance

To understand the effect on performance of selecting various memory modes, we use a system that is configured with X7560 processors and populated with sixty-four 4 GB quad-rank DIMMs.

Figure 5-17 shows the peak system-level memory throughput for various memory modes, measured using an IBM-internal memory load generation tool. As shown, there is a 50% decrease in peak memory throughput when comparing a normal (non-mirrored) configuration to a mirrored memory configuration.

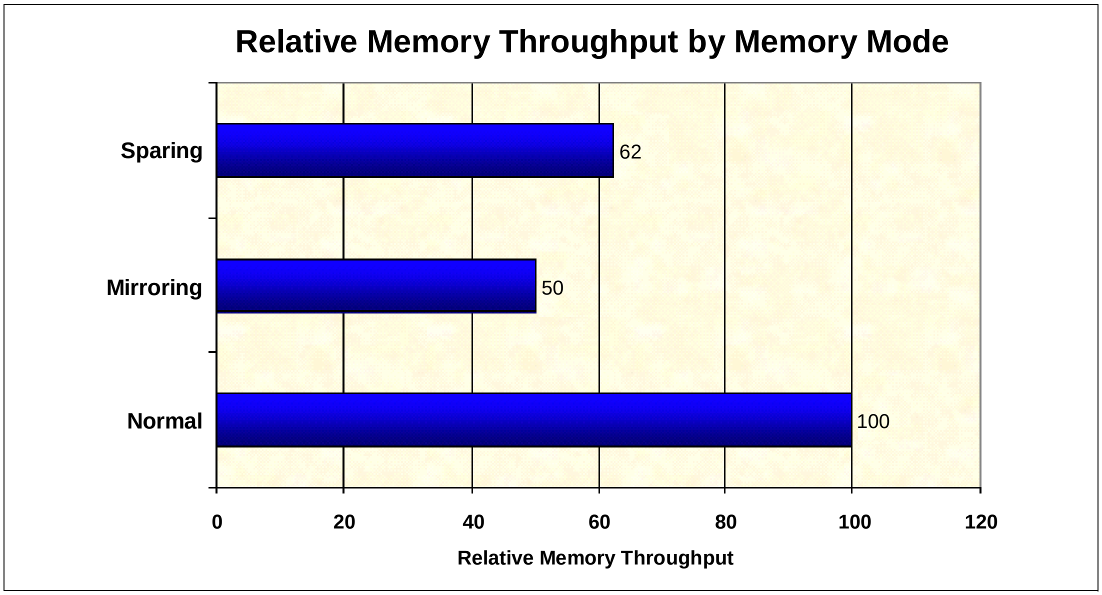


Figure 5-17 Relative memory throughput by memory mode

5.11 Storage

The storage system on the HX5 blade is based on the use of the SSD Expansion Card for IBM BladeCenter HX5, which contains an optional LSI 1064E SAS Controller and two 1.8-inch micro SATA drive connectors. The SSD Expansion Card allows the attachment of two 1.8-inch solid-state drives (SSDs). If two SSDs are installed, the HX5 supports RAID-0 or RAID-1 capability.

Installation of the SSDs in the HX5 requires the SSD Expansion Card for IBM BladeCenter HX5. Only one SSD Expansion Card is needed for either one or two SSDs. Table 5-22 lists the ordering details.

Table 5-22 SSD Expansion Card for IBM BladeCenter HX5

Part number	Feature code	Description
46M6908	5765	SSD Expansion Card for IBM BladeCenter HX5

Figure 5-18 on page 204 shows the SSD Expansion Card.

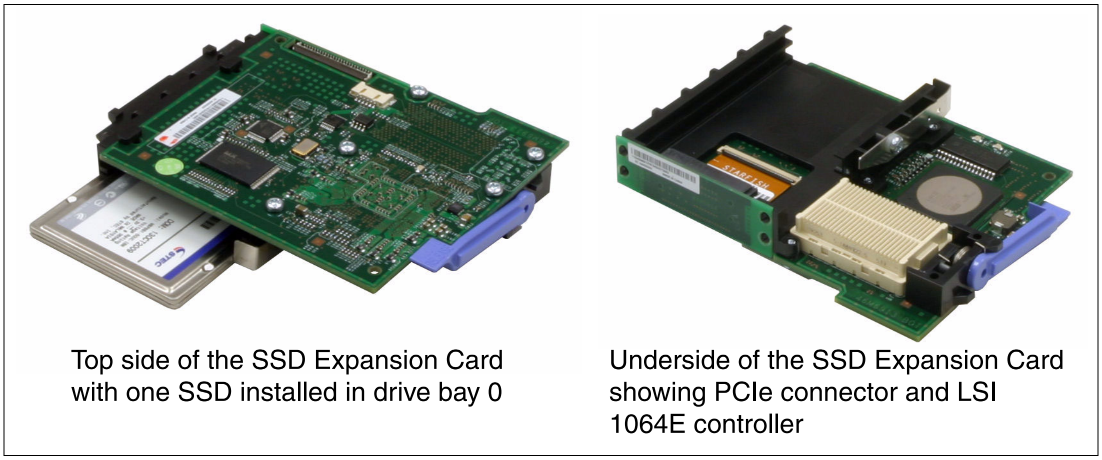


Figure 5-18 SSD Expansion Card for the HX5 (left: top view; right: underside view)

The SSD Expansion Card can be installed in the HX5 in combination with a CIOv I/O expansion card and CFFh I/O expansion card, as shown in Figure 5-19.

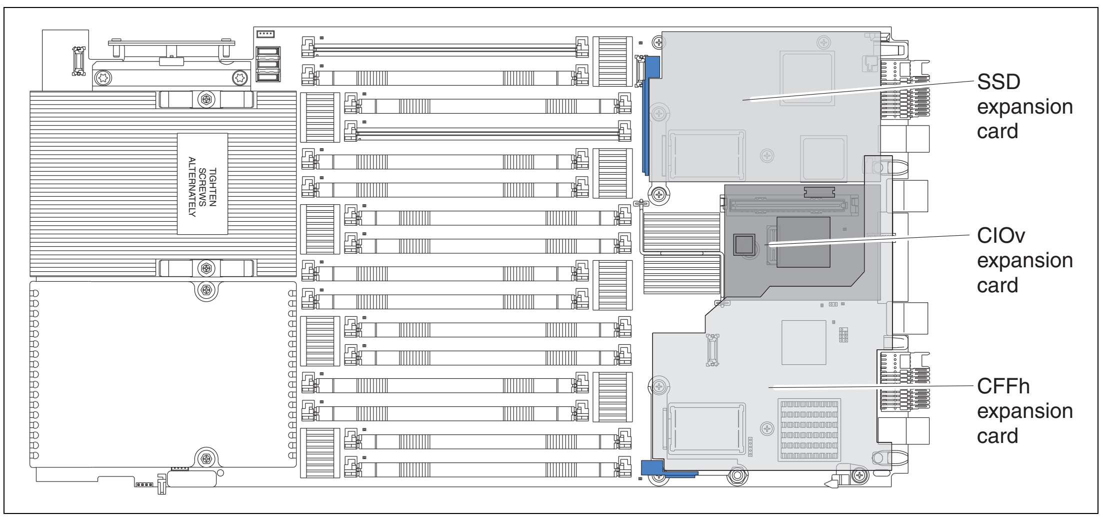


Figure 5-19 Placement of SSD expansion card in combination with a CIOv card and CFFh card

ServeRAID MR10ie support: The HX5 does not currently support the ServeRAID MR10ie RAID controller.

5.11.1 Solid-state drives (SSDs)

SSDs are a relatively new technology in the server world. SSDs are more reliable than spinning hard disk drives. SSDs consume much less power than a standard serial-attached SCSI (SAS) drive, approximately 0.5 W (SSD) versus 11 W (SAS).

Target applications for SSDs include video surveillance, transaction-based database (DB), and other applications that have high performance but moderate space requirements. Table 5-23 on page 205 lists the supported SSDs.

Table 5-23 Supported SSDs

Part number	Feature code	Description
43W7734	5314	IBM 50GB SATA 1.8-inch NHS SSD

For more information about SSD drives and their advantages, see 2.8.1, “IBM eXFlash price-performance” on page 49.

5.11.2 LSI configuration utility

Figure 5-20 shows the LSI SAS Configuration Utility window running on a 2-node HX5 with one controller in each node. The SAS1064 that is listed first is always the primary node controller and the SAS1064 that is listed second is the secondary node’s controller.

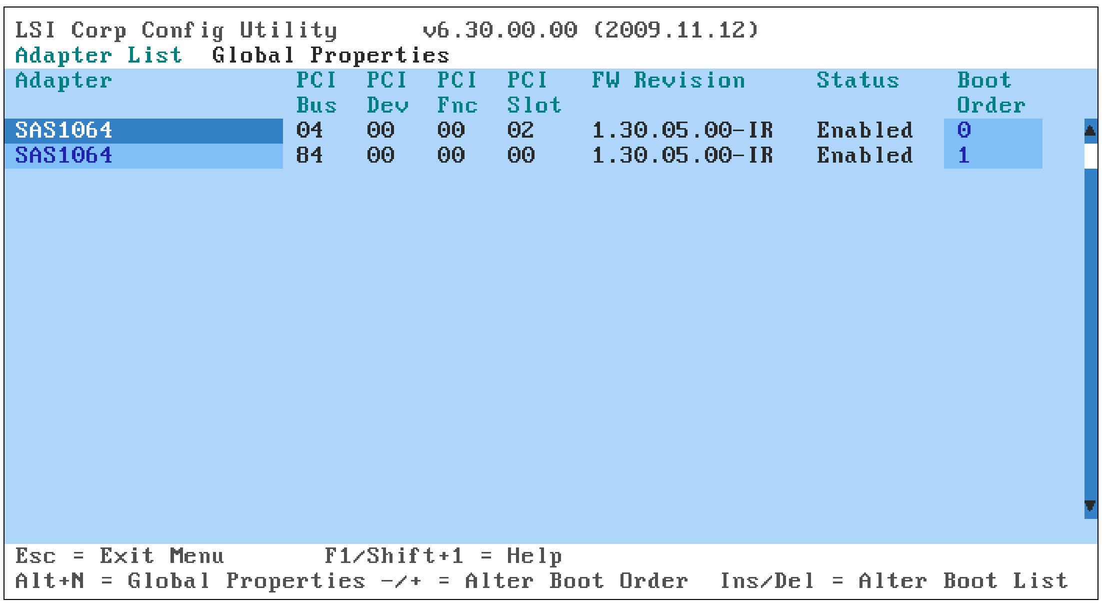


Figure 5-20 LSI Configuration Utility

In a 2-node configuration, each controller operates independently, and each controller maintains its own configuration for the sets of drives that are installed in that node. One controller cannot cross over to the other node to perform a more complex RAID solution.

Using independent controllers allows for several configuration options. Each LSI 1064 controller has an option of RAID-1, RAID-0, or JBOD (just a bunch of disks). No redundancy exists in a JBOD configuration, and each drive runs independently. The blade uses JBOD, by default, if no RAID array is configured.

Figure 5-21 on page 206 shows the three options in the LSI 1064 setup page, LSI Logic MPT Setup Utility. Only two options are supported with this blade, because there are only two drives maximum that are installable in the HX5 and an Integrated Mirroring Enhanced (IME) volume requires three drives minimum.

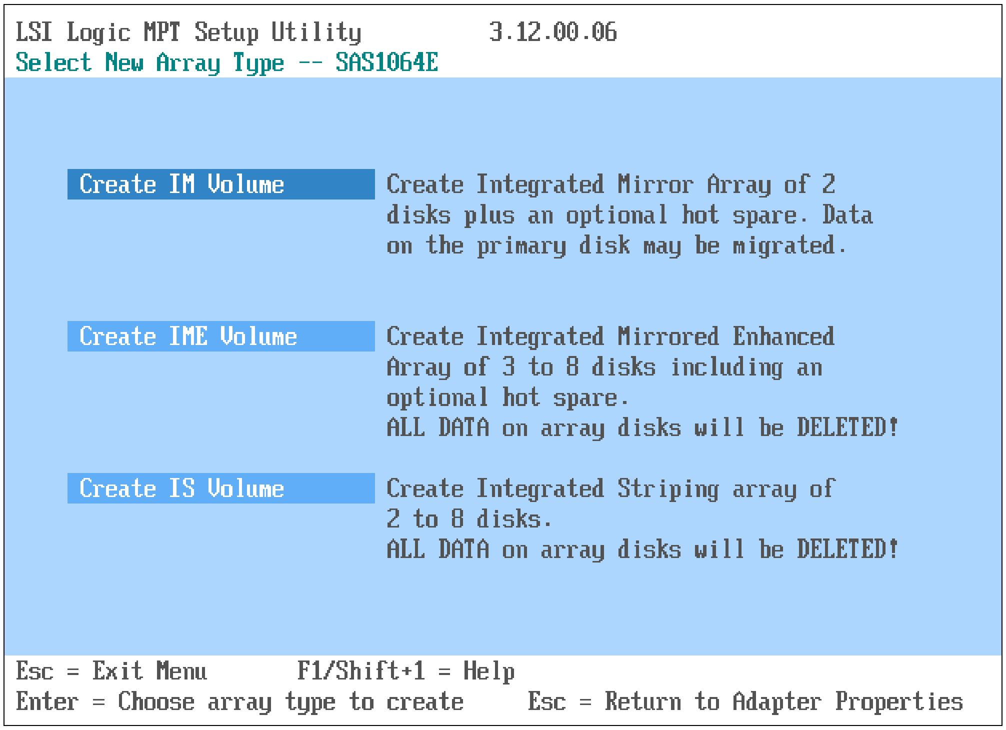


Figure 5-21 RAID choices using the LSI configuration utility

The following options are presented:

- Create IM Volume: Creates a RAID-1 array

  RAID 1 drives are mirrored on a 1 to 1 ratio. If one drive fails, the other drive takes over automatically and keeps the system running. However, in this configuration, you lose 50% of your disk space with one of the drives being used as a mirrored image.

  The stripe size is 64 Kb and cannot be altered.

  This option also affects the performance on the drives, because all data has to be written twice, once per drive. See the performance chart in Figure 5-22 on page 207 for details.

- Create IME Volume: Creates a RAID-1E array

  This option requires three drives, so it is not available in the HX5.

- Create IS Volume: Creates a RAID-0 array

  RAID 0 or the Integrated Striping (IS) volume, as shown in LSI, is one of the faster performing disk arrays because read and write sectors of data are interleaved between multiple drives. The downside to this configuration is immediate failure if one drive fails. There is no redundancy.

  In a RAID-0, you also keep the full size of both drives. Identical size drives are recommended for performance, as well as data storage efficiency.

  The stripe size is 64 Kb and cannot be altered.

We provide the instructions to create an array in 8.4, “Local storage considerations and array setup” on page 385.

5.11.3 Determining which SSD RAID configuration to choose

Using an industry standard IO tool that measures hard drive performance, we tested each available configuration type. We used two separate tests: one test using 50% sequential reads and 50% random writes (Figure 5-22) and the other test using 90% sequential reads and 10% random writes (Figure 5-23). We tested each group using 16 KB, 512 KB, and 1 MB transfer request sizes, which are the most common transfer request sizes that are used in server environments today.

50% Sequential / 50% Random									
RAID Type	16k Transfer Request Size			512k Transfer Request Size			1mb Transfer Request Size		
	i/o per Second	Total MB per	Average I/O Response	i/o per Second	Total MB per	Average I/O Response	i/o per Second	Total MB per	Average I/O Response
RAID 0	836.34	13.07	1.1943ms	119.46	59.73	8.3699ms	67.43	67.43	14.8296ms
RAID 1	888.44	13.88	1.1242ms	118.06	59.03	8.4689ms	62.27	62.27	16.0559ms
JBOD	1440.69	22.51	0.6928ms	141.06	70.76	7.0698ms	73.61	73.61	13.5844ms

Figure 5-22 Showing 50% Sequential/50% Random test results

90% Sequential / 10% Random									
RAID Type	16k Transfer Request Size			512k Transfer Request Size			1mb Transfer Request Size		
	i/o per Second	per second	Response Time	i/o per Second	per second	Response Time	i/o per Second	per second	Response Time
RAID 0	879.87	13.75	1.1352ms	132.39	66.19	7.5521ms	70.53	70.53	14.1774ms
RAID 1	914.68	14.29	1.0919ms	116.92	58.46	8.5514ms	62.59	62.59	15.9745ms
JBOD	1521.63	23.78	0.06559ms	142.68	71.34	7.0072ms	74.02	74.02	13.5085ms

Figure 5-23 Showing 90% Sequential/10% Random test results

> **Tests:** These tests are not certified tests. We performed these tests to illustrate the performance differences among the three configuration options.

We ran these tests on an HX5 2-node system running four 2 GHz Intel Xeon 7500 series 6-core processors and 16 GB of memory. The SSDs were IBM 50 GB SATA 1.8-inch NHS SSDs. Results might vary depending on the size and type of installed drives.

The results show that when choosing a configuration for setting up hard drives, the performance difference between a RAID-1, RAID-0, and JBOD configuration is minimal. JBOD was always the fastest performer, followed by RAID-1, and then RAID-0.

5.11.4 Connecting to external SAS storage devices

The SAS Connectivity Card (CIOv) for IBM BladeCenter is an expansion card that offers the ideal way to connect the supported BladeCenter servers to a wide variety of SAS storage devices. The SAS Connectivity Card connects to two SAS Controller Modules in the BladeCenter chassis. You can then attach these modules to the IBM System Storage DS3200 from the BladeCenter H or HT chassis, or Disk Storage Modules in the BladeCenter S.

SAS signals are routed from the LSI 1064E controller on the SSD Expansion Card to the SAS Connectivity Card, as shown in Figure 5-24 on page 208.

Two of the SAS ports (SAS 0 and SAS 1) from the LSI 1064E on the SSD Expansion Card are routed to the 1.8-inch SSD connectors. The other SAS ports (SAS 2 and SAS 3) are routed from the LSI 1064E controller through the server planar to CIOv connector, where the SAS Connectivity Card (CIOv) is installed.

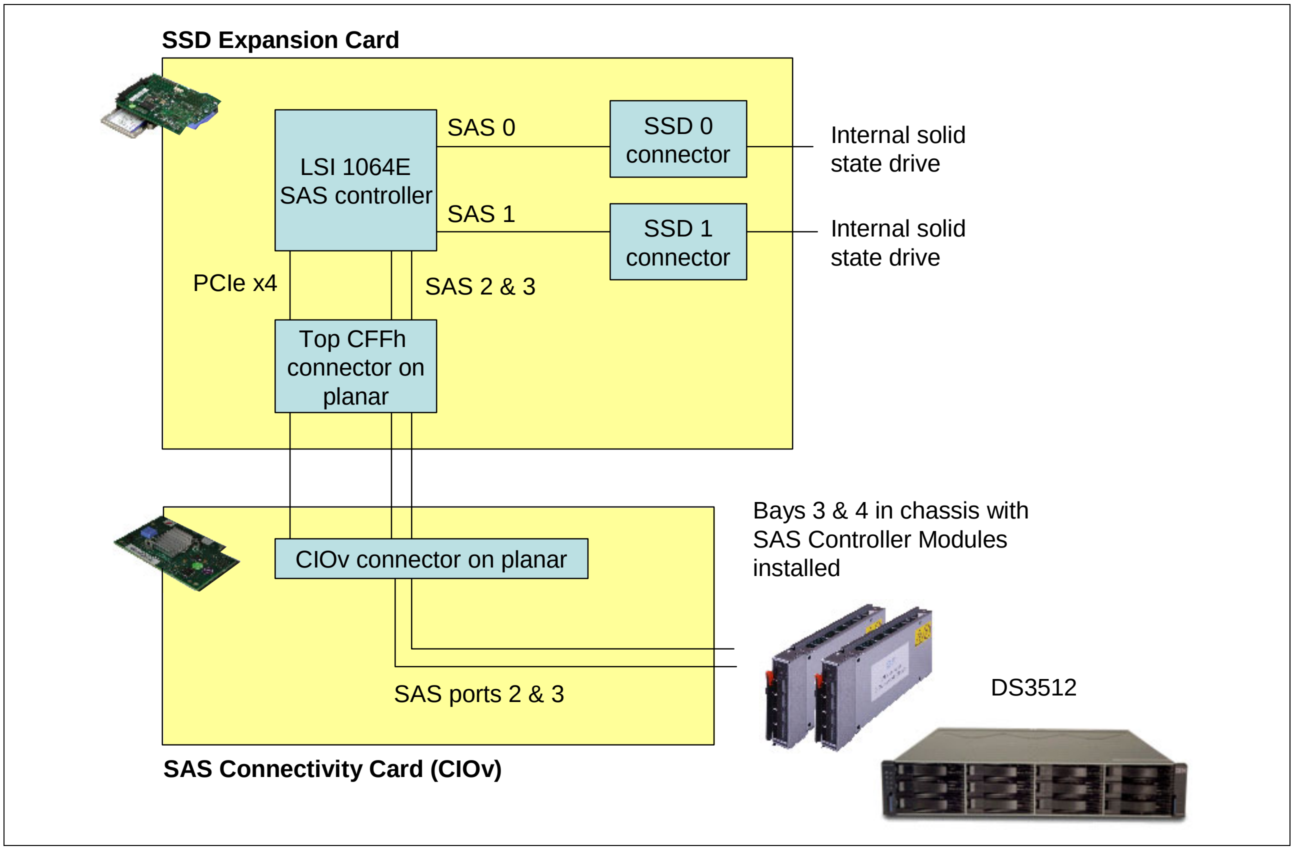


Figure 5-24 Connecting SAS Connectivity Card and external SAS solution

5.12 BladeCenter PCI Express Gen 2 Expansion Blade

The IBM BladeCenter PCI Express Gen 2 Expansion Blade provides the capability to attach selected PCI Express cards to the HX5. This capability is ideal for many applications that require special telecommunications network interfaces or hardware acceleration using a PCI Express card.

The expansion blade provides one full-height and full-length PCI Express slot and one full-height and half-length PCI Express slot with a maximum power usage of 75 W for each slot. It integrates the PCI Express card support capability into the BladeCenter architecture.

You can attach up to three expansion blades to a single-node HX5. You can attach up to two expansion blades to a 2-node HX5.

See Table 5-24 for ordering information.

Table 5-24 PCI Express Gen 2 Expansion Blade

Part number	Feature code	Description
46M6730	9295	IBM BladeCenter PCI Express Gen 2 Expansion Blade

The expansion blade has the following features:

- Support for PCIe 2.0 adapters in an expansion blade

  The expansion blade lets you install one or two standard form factor PCIe 2.0 adapter cards in a BladeCenter environment, enabling the use of specialized adapters or adapters that otherwise are not available to BladeCenter clients. Each of the two adapters can consume up to 75 W.

- Ability to stack up to four expansion blades on a single base blade

  You can attach up to two, three, or four expansion blades (depending on the attached server), therefore, maintaining the BladeCenter density advantage while still giving you the option to install PCIe cards as needed without the need for each expansion blade having to be attached to a server and the added complexity and cost that brings. The first expansion blade connects to the server blade using the CFFh expansion slot of the server blade. The second expansion blade attaches to the CFFh connector on the first expansion blade, and so on.

- The following maximums are for the attached expansion blades:

  Single-node HX5: Up to three expansion blades

  Two-node HX5: Up to two expansion blades

  > **MAX5 support:** The HX5 with an attached MAX5 does not also support the attachment of the PCI Express Gen 2 Expansion Blade.

- CFFh slot still available

  The CFFh expansion connector is accessible on the topmost expansion blade, even with four expansion blades attached. This design lets you maintain the integrated networking capabilities of the blade server when it is installed in a BladeCenter S, H, or HT chassis.

For details about the supported PCI Express adapter cards, see the IBM Redbooks at-a-glance guide, *IBM BladeCenter PCI Express Gen 2 Expansion Blade*, TIPS0783, which is available at this website:

http://www.ibm.com/redbooks/abstracts/tips0783.html

5.13 I/O expansion cards

The HX5 type 7872 connects to a wide variety of networks and fabrics by installing the appropriate I/O expansion card. Supported networks and fabrics include 1 Gb and 10 Gb Ethernet, 4 Gb and 8 Gb Fibre Channel, SAS, and InfiniBand.

The HX5 blade server with an I/O expansion card is installed in a supported BladeCenter chassis complete with switch modules (or pass-through) that are compatible with the I/O expansion card in each blade. The HX5 supports two types of I/O expansion cards: the CIOv and the CFFh form factor cards.

5.13.1 CIOv

The CIOv I/O expansion connector provides I/O connections through the midplane of the chassis to modules located in bays 3 and 4 of a supported BladeCenter chassis. The CIOv slot is a second-generation PCI Express 2.0 x8 slot. A maximum of one CIOv I/O expansion card is supported per HX5. A CIOv I/O expansion card can be installed on a blade server at the same time that a CFFh I/O expansion card is installed in the blade.

Table 5-25 lists the CIOv expansion cards that are supported in the HX5.

Table 5-25 Supported CIOv expansion cards

Part number	Feature code	Description
44X1945	1462	QLogic 8Gb Fibre Channel Expansion Card (CIOv)
46M6065	3594	QLogic 4Gb Fibre Channel Expansion Card (CIOv)
46M6140	3598	Emulex 8Gb Fibre Channel Expansion Card (CIOv)
43W4068	5093	SAS Connectivity Card (CIOv)
44W4475	5477	Ethernet Expansion Card (CIOv)

See the IBM ServerProven compatibility website for the latest information about the expansion cards that are supported by the HX5:

http://ibm.com/servers/eserver/serverproven/compat/us/

CIOv expansion cards are installed in the CIOv slot in the HX5 2-socket, as shown in Figure 5-25.

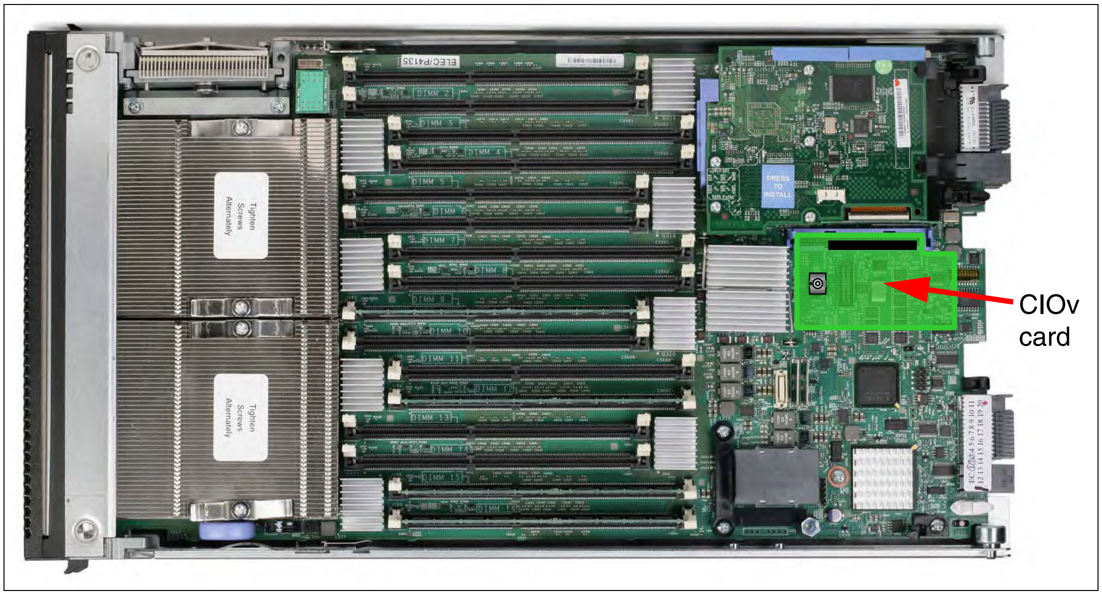


Figure 5-25 The HX5 type 7872 showing the CIOv I/O expansion card position

5.13.2 CFFh

The CFFh I/O expansion connector provides I/O connections to high-speed switch modules that are located in bays 7, 8, 9, and 10 of a BladeCenter H or BladeCenter HT chassis, or to switch bay 2 in a BladeCenter S chassis.

The CFFh slot is a second-generation PCI Express x16 (PCIe 2.0 x16) slot. A maximum of one CFFh I/O expansion card is supported per blade server. A CFFh I/O expansion card can be installed on a blade server at the same time that a CIOv I/O expansion card is installed in the server.

Table 5-26 lists the supported CFFh I/O expansion cards.

Table 5-26 Supported CFFh expansion cards

Part number	Feature code	Description
39Y9306	2968	QLogic Ethernet and 4Gb Fibre Channel Expansion Card[a]
44X1940	5485	QLogic Ethernet and 8Gb Fibre Channel Expansion Card
46M6001	0056	2-Port 40Gb InfiniBand Expansion Card
42C1830	3592	QLogic 2-port 10Gb Converged Network Adapter
44W4465	5479	Broadcom 4-port 10Gb Ethernet CFFh Expansion Card[a]
44W4466	5489	Broadcom 2-Port 10Gb Ethernet CFFh Expansion Card[a]
46M6164	0098	Broadcom 10Gb 4-port Ethernet Expansion Card
46M6168	0099	Broadcom 10Gb 2-port Ethernet Expansion Card
49Y4235	5755	Emulex Virtual Fabric Adapter
44W4479	5476	2/4 Port Ethernet Expansion Card
46m6003	0056	2-port 40Gb InfiniBand Expansion Card (CFFh)

a. IBM System x has withdrawn this card from marketing.

See the IBM ServerProven compatibility website for the latest information about the expansion cards that are supported by the HX5:

http://ibm.com/servers/eserver/serverproven/compat/us/

CFFh expansion cards are installed in the CFFh slot in the HX5, as shown in Figure 5-26.

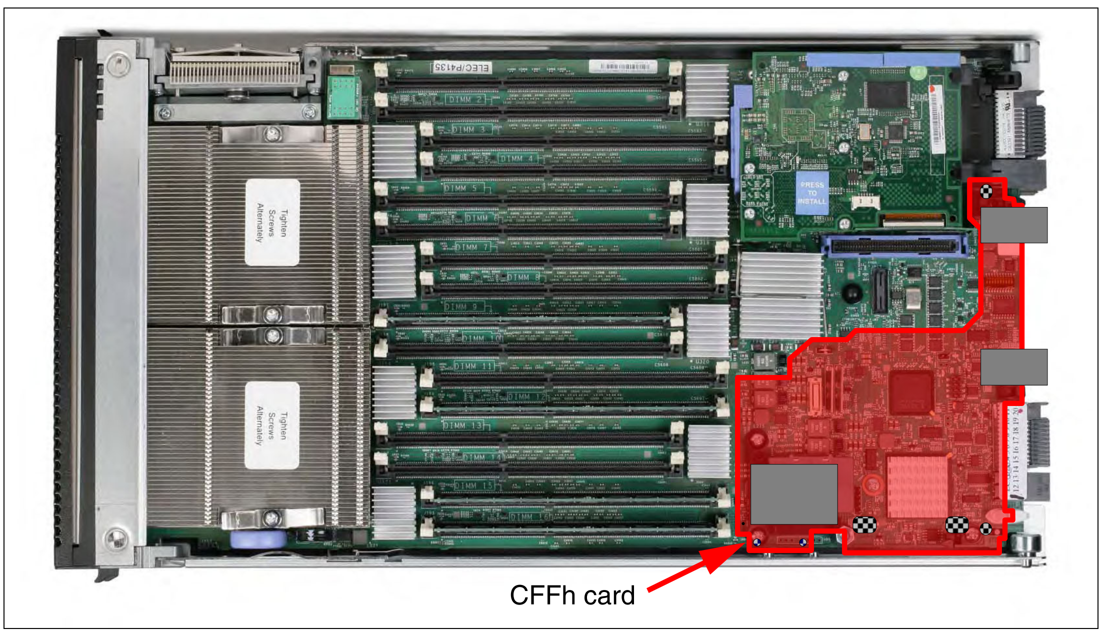


Figure 5-26 The HX5 type 7872 showing the CFFh I/O expansion card position

A CFFh I/O expansion card requires that a supported high-speed I/O module or a Multi-switch Interconnect Module is installed in bay 7, 8, 9, or 10 of the BladeCenter H or BladeCenter HT chassis.

In a BladeCenter S chassis, the CFFh I/O expansion card requires a supported switch module in bay 2. When used in a BladeCenter S chassis, a maximum of two ports are routed from the CFFh I/O expansion card to the switch module in bay 2.

See the *IBM BladeCenter Interoperability Guide* for the latest information about the switch modules that are supported with each CFFh I/O expansion card at the following website:

http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5073016

5.14 Standard onboard features

This section describes the standard onboard features of the HX5 blade server:

- UEFI
- Onboard network adapters
- Integrated systems management processor called Integrated Management Module (IMM)
- Video controller
- Trusted Platform Module (TPM)

5.14.1 UEFI

The HX5 2-socket uses an integrated Unified Extensible Firmware Interface (UEFI) next-generation BIOS.

The UEFI provides the following capabilities:

- Human readable event logs; no more beep codes
- Complete setup solution by allowing adapter configuration function to be moved to UEFI
- Complete out-of-band coverage by Advanced Settings Utility to simplify remote setup

Using all of the features of UEFI requires an UEFI-aware operating system and adapters. UEFI is fully backward-compatible with BIOS.

For more information about UEFI, see the IBM white paper, *Introducing UEFI-Compliant Firmware on IBM System x and BladeCenter Servers*, which is available at the following website:

http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5083207

For the UEFI menu setup, see 8.5, “UEFI settings” on page 396.

5.14.2 Onboard network adapters

The HX5 2-socket includes a dual-port Gigabit Ethernet controller with the following specifications:

- Broadcom BCM5709S dual-port Gigabit Ethernet controller
- Support for TCP Offload Engine (TOE)
- Support for failover and load balancing for better throughput and system availability
- Support for highly secure remote power management using Intelligent Platform Management Interface (IPMI) 2.0
- Support for Wake on LAN and Preboot Execution Environment (PXE)
- Support for IPv4 and IPv6

5.14.3 Integrated Management Module (IMM)

The HX5 blade server includes an IMM to monitor server availability, perform predictive failure analysis and so forth, and trigger IBM Systems Director alerts. The IMM performs the functions of the baseboard management controller (BMC) of earlier blade servers, and adds the features of the Remote Supervisor Adapter (RSA) in System x servers, and also remote control and remote media.

For more information about the IMM, see the IBM white paper, *Transitioning to UEFI and IMM*, which is available at the following website:

http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5079769

The IMM controls the service processor LEDs and the light path diagnostics capability. The IMM controls the LEDs, which can indicate an error and the physical location of the error. To enable illumination of the LEDs after the blade is removed from the chassis, the LEDs have a backup power system. The LEDs are related to DIMMs, CPUs, battery, CIOv connector, CFFh connector, scalability, the system board, non-maskable interrupt (NMI), CPU mismatch, and the SAS connector.

5.14.4 Video controller

The video subsystem in the HX5 supports an SVGA video display. The video subsystem is a component of the IMM and is based on a Matrox video controller. The HX5 has 128 MB of video memory. Table 5-27 lists the supported video resolutions.

Table 5-27 Supported video resolutions

Resolution	Maximum refresh rate
640 x 480	85 Hz
800 x 600	85 Hz
1024 x 768	75 Hz

5.14.5 Trusted Platform Module (TPM)

Trusted computing is an industry initiative that provides a combination of secure software and secure hardware to create a trusted platform. It is a specification that increases network security by building unique hardware IDs into computing devices. The HX5 implements TPM Version 1.2 support.

The TPM in the HX5 is one of the three layers of the trusted computing initiative, as shown in Table 5-28.

Table 5-28 Trusted computing layers

Layer	Implementation
Level 1: Tamper-proof hardware, used to generate trustable keys	Trusted Platform Module
Level 2: Trustable platform	► UEFI or BIOS ► Intel processor
Level 3: Trustable execution	► Operating system ► Drivers

5.15 Integrated virtualization

The HX5 offers an IBM 2 GB USB flash drive option that is preloaded with either VMware ESXi 4.0 Update 1 or VMware ESXi 4.1:

- ESXi 4.0 is for 1-node HX5 configurations only.
- ESXi 4.1 is required for configurations with the MAX5 memory expansion blade or for 2-node HX5 configurations.

The virtualization-optimized models include ESXi 4.1, as listed in Table 5-6 on page 184.

ESXi is an embedded version of VMware ESX and the hypervisor is fully contained on the flash drive. Table 5-29 lists the ordering information for the IBM USB Memory Key for VMware Hypervisor.

Table 5-29 USB Hypervisor option

Part number	Feature code	Description
41Y8278	1776	IBM USB Memory Key for VMware Hypervisor ESXi 4.0 Update 1
41Y8287	3033	IBM USB Memory Key for VMware ESXi 4.1 (required for MAX5)

As shown in Figure 5-27, the flash drive plugs into the Hypervisor Interposer, which, in turn, attaches to the system board near the processors. The Hypervisor Interposer is included as standard with the HX5.

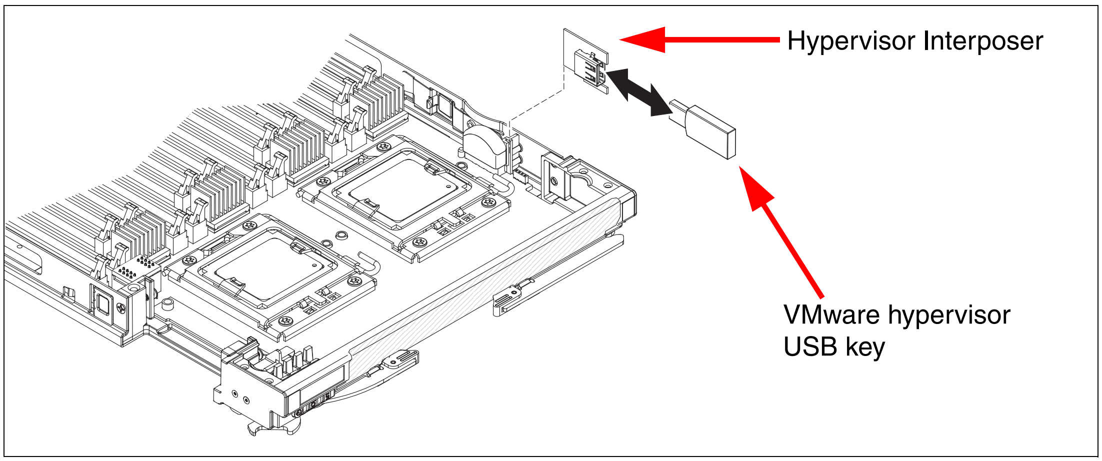


Figure 5-27 Placement of VMware USB key in HX5

See 5.17, “Operating system support” on page 215 for details about VMware ESX and other operating system support.

5.16 Partitioning capabilities

When you have a 4-socket HX5 that consists of two HX5 blade servers, you use scalable complex within the Advanced Management Module to create, delete, and switch between stand-alone mode and scaled mode, as shown in Figure 5-28 on page 215.

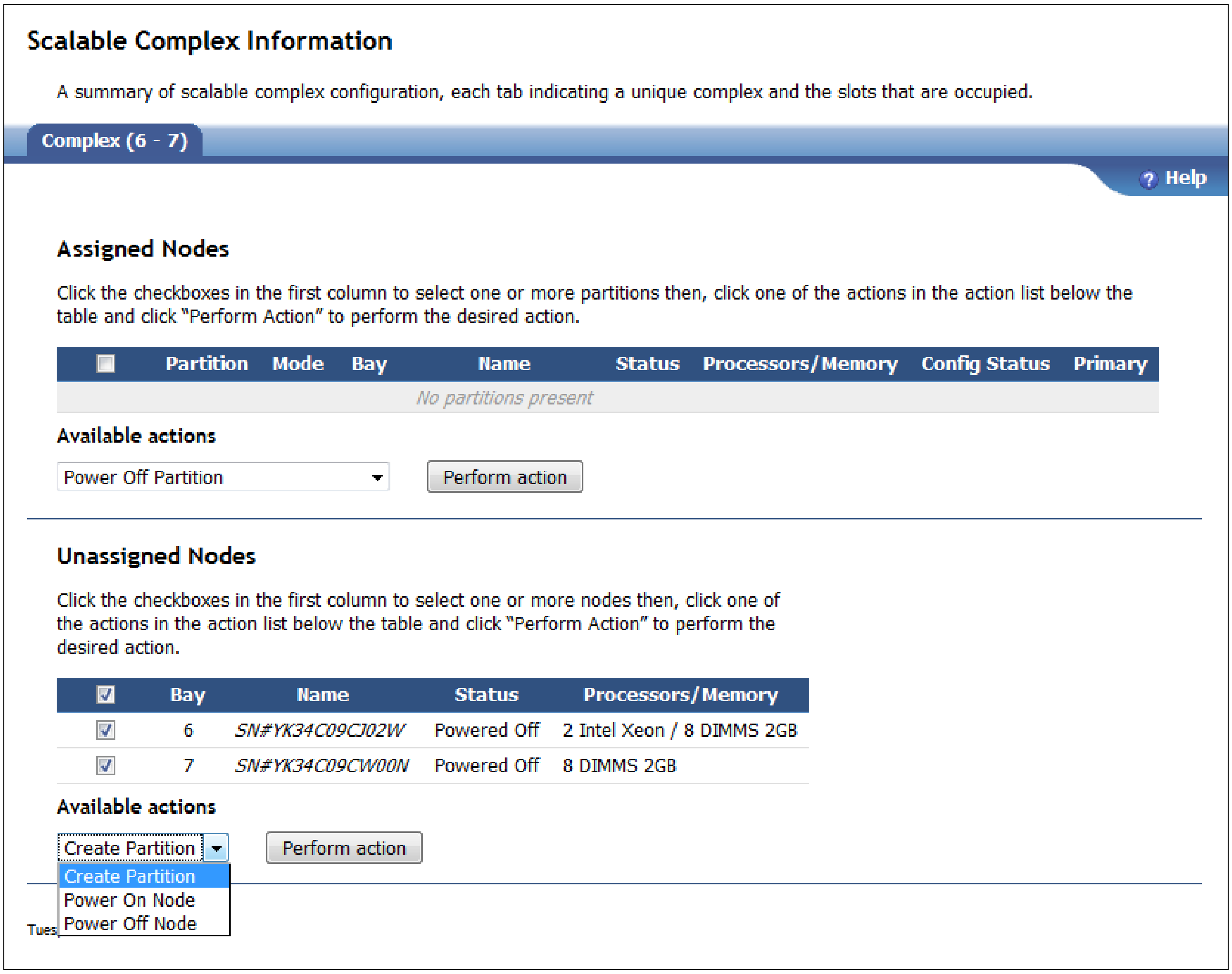


Figure 5-28 Two unpartitioned HX5s

For information about setup options and common errors, see 8.6, “Creating an HX5 scalable complex” on page 402.

5.17 Operating system support

The HX5 supports the following operating systems:

- Microsoft Windows Server 2008 R2
- Microsoft Windows Server 2008 Datacenter x64 Edition
- Microsoft Windows Server 2008 Enterprise x64 Edition
- Microsoft Windows Server 2008 HPC Edition
- Microsoft Windows Server 2008 Standard x64 Edition
- Microsoft Windows Server 2008 Web x64 Edition
- Red Hat Enterprise Linux 5 Server x64 Edition
- Red Hat Enterprise Linux 5 Server with Xen x64 Edition
- Red Hat Enterprise Linux 6 Server x64 Edition
- SUSE Linux Enterprise Server 10 for AMD64/EM64T
- SUSE Linux Enterprise Server 10 with Xen for AMD64/EM64T
- SUSE Linux Enterprise Server 11 for AMD64/EM64T

- SUSE Linux Enterprise Server 11 with Xen for AMD64/EM64T
- VMware ESX 4.0 Update 1
- VMware ESXi 4.0 Update 1
- VMware ESX 4.1
- VMware ESXi 4.1

Key information regarding VMware ESX:

- ESXi 4.0 support is single-node HX5 only.
- ESX 4.0 supports single-node and 2-node HX5.
- ESXi 4.1 and ESX 4.1 are required for MAX5 support.

See the ServerProven website for the most recent information:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/ematrix.shtml

Part 2

Implementing scalability

In this second part of the book, we provide detailed configuration and setup information to get your servers operational. We particularly focus on setting up MAX5 configurations of all three eX5 servers, as well as 2-node configurations of the x3850 X5 and HX5.

This part consists of the following chapters:

6

IBM System x3850 X5 and x3950 X5

The System x3850 X5 and x3950 X5 are enterprise-class Intel processor-based servers for mission-critical applications. Scaling two of these servers together and virtualizing the resources allow this server to replace an entire rack of conventional servers.

This chapter provides assistance for making configuration, monitoring, and maintenance decisions when implementing an x3850 X5 or x3950 X5 configuration. The information provided is meant to help you make informed decisions. Do not consider this document as the absolute implementation process.

This chapter contains the following topics:

6.1 Before you apply power for the first time after shipping

Before you begin populating your server with all of its processors, memory, and PCI adapters and before you install an operating system, perform the recommendations described in the following sections.

6.1.1 Verify that the components are securely installed

Perform the following tasks to ensure that all of the electrical components of your server have proper connectivity:

- Inspect heat sinks to ensure that they are secure.
- Verify that dual inline memory modules (DIMMs) are mounted in the correct locations and are fully plugged in with their retain clips in the closed position and the memory cards are in the correct location.
- Inspect PCIe adapters to ensure that they are securely plugged into their slots. The blue retaining clips for the PCIe adapter can come loose during shipping.
- Check all of the cable connections on the hard drive backplane and all internal disk controllers to ensure that they are properly snapped into place. If a cable can be unplugged from its connector easily by tugging on the cable, plug it back in until it clicks or clips into place.

6.1.2 Clear CMOS memory

When a server is shipped from one location to another location, you have no idea what the server has been exposed to. For all you know, it might have been parked next to a large magnet or electric motor and everything in the server that stores information magnetically, including CMOS memory, has been altered. IBM does not indicate on the shipping carton that magnetic material is enclosed because the information is readily recoverable.

Booting the server to the F1 system configuration panel and selecting **Load Default Settings** restore the default values for the items that you can change in configuration. These steps do not change the settings of the registers that are used by the Integrated Management Module (IMM) and the Unified Extensible Firmware Interface (UEFI). These registers define the system state of the server, and if they become corrupt, they can affect the server in the following ways:

- Fail to power on
- Fail to complete the power-on self test (POST)
- Turn on amber light path diagnostics lights that describe conditions that do not exist
- Reboot unexpectedly
- Fail to detect all of the installed CPU, memory, PCIe adapters, or physical disks

These internal registers cannot be modified or restored to the defaults by the F1 system configuration panel; however, they can be restored to the defaults by clearing the CMOS memory.

With ac power removed from the server, you can clear CMOS memory by following these steps:

1. Remove the top cover of the server.
2. Locate the CMOS battery on the right side of the server, behind the disk fans, to the right of processor 4. See Figure 6-1 on page 221.

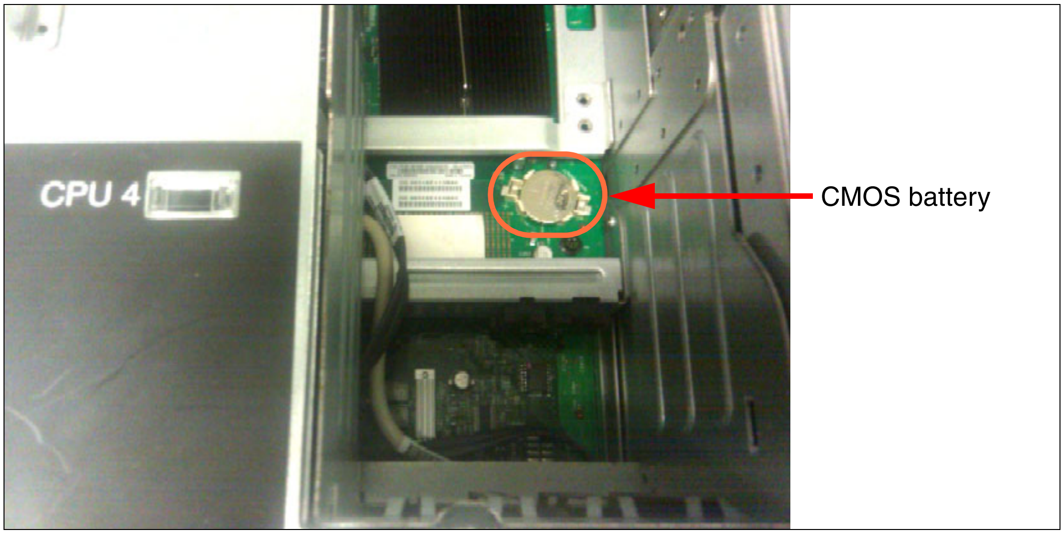


Figure 6-1 CMOS battery location

3. Use your finger to pry up the battery on the side closest to the neighboring IC chip. The battery easily lifts out of the socket.

> **Tip:** The light path diagnostics (LPD) lights are powered from a separate power source (a capacitor) than the CMOS memory. LPD lights remain lit for a period of time after ac power and the CMOS memory battery have been removed.

4. After 30 seconds, insert one edge of the battery, with the positive side up, back into the holder.
5. Push the battery back into the socket with your finger and it clips back into place, as shown in Figure 6-1.

6.1.3 Verify that the server completes POST before adding options

When you have ordered options for your server that have not yet been installed, it is a good idea to ensure that the server completes the POST properly before you start to add your options. Performing this task makes it easier to compartmentalize a potential problem with an installed option rather than having to look at the entire server to try to identify a good starting point during problem determination.

6.2 Processor considerations

> **Tip:** To understand the information in this section, thoroughly read the information in 3.7, “Processor options” on page 74.

This server supports a total of four matched processors. The required match refers to the family of processors, number of cores, size of level 2 cache, core, and front side bus speeds.

As a matter of standard manufacturing, the producing vendor might alter the method of manufacturing the processor, which results in separate stepping levels, but does not affect the overall functionality of the processor in its ability to communicate with other processors in the same server. Any operational differences in stepping levels are handled by the microcode of the processor, Integrated Management Module (IMM), and UEFI.

6.2.1 Minimum processors required

The minimum number of processors required for the server to boot into any operational configuration is one. The processor that must be installed is the processor in socket 1 with at least two DIMMs installed in memory card 1 for local memory. In this configuration, PCIe slots 1 through 4 are not functional. For all of the PCIe slots to be available, the server must have two processors installed with one processor in processor socket 1 and the other processor installed in socket 4.

6.2.2 Processor operating characteristics

Table 6-1 describes the operating characteristics of the server, based on the number of processors and how the memory is installed. The table also describes how the server reacts in the unlikely event of a failure. This server, in a stand-alone configuration, also has the option of installing QuickPath Interconnect (QPI) wrap cards to establish a QPI link between processors 1 and 2, or processors 3 and 4, depending on the ports in which the QPI wrap cards are inserted. While the server can function without the QPI wrap plug installed, a significant performance boost occurs for memory-intensive tasks that share memory between processors 1 and 2 and between processors 3 and 4. Table 6-1 describes the effects of not having the cards installed or what can occur if the cards fail.

Table 6-1 Operating characteristics of processor and memory installation options

Memory Installed	Operational characteristic
Only processor 1 is installed as a minimum configuration for testing	
Regardless of the amount of memory installed.	PCIe slots 1 - 4 are not functional.
Memory installed only on memory cards 1 and 2. Minimum of 2 DIMMs on memory card 1.	Performance improves as DIMMs are added to evenly populate all ranks on each memory card.
Memory installed on memory cards other than memory cards 1 and 2.	None of the memory installed on other memory cards is accessible by processor 1.
Memory installed on both the system board and the MAX5.	Performance improves significantly when more active memory calls use local memory on the system board.
Both processors 1 and 4 are installed	
Memory installed only on memory cards 1 and 2. Minimum of two DIMMs on memory card 1.	► Performance improves as DIMMs are added to evenly populate all ranks on each memory controller. ► Processing threads assigned to processor 4 always have a significant drop in performance for memory-intensive tasks. ► This configuration is not an operational configuration for operating systems, such as VMware.

Memory Installed	Operational characteristic
Memory is installed on memory cards 1, 2, 7, and 8.	► Performance improves as DIMMs are added to evenly populate all ranks on each memory card. ► All memory on memory cards 7 and 8 is not visible if processor 4 fails. ► The operational configuration for VMware requires that the memory cards of the installed processor have the exact same total memory installed. For ease of maintenance, the best practice is to have an identical memory configuration on each of the installed processor's memory cards.
Memory is installed on memory cards 1, 2, 7, and 8 and MAX5.	► Performance improves as DIMMs are added to evenly populate all ranks on each memory controller for local memory. ► Performance improves significantly when more active memory calls use local memory on the system board instead of the memory that is located in the MAX5. ► In the rare instance of a memory card failure, VMware operational requirements can be satisfied if all of the memory is installed in the MAX5 or each processor has the same amount of memory disabled. Memory access is significantly slower if all of the memory is in the MAX5.
Processors 2 and 3 are added to the existing processors 1 and 4	
Memory is installed only on memory cards 1, 2, 7, and 8. Minimum of two DIMMs on memory card 1.	► Performance improves as DIMMs are added to evenly populate all ranks on each memory controller. ► Processing threads assigned to processors without memory cards have a significant drop in performance for memory-intensive tasks. ► This configuration is not an operational configuration for operating systems, such as VMware.
Memory is installed on all memory cards for installed processors.	► Performance improves as DIMMs are added to evenly populate all ranks on each memory card. ► Memory that is installed on memory cards that are associated with a failed or uninstalled processor is not seen by the other processors. ► The operational configuration for VMware requires the memory cards of the installed processors have the exact same total memory installed. For ease of maintenance, the best practice is to have an identical memory configuration on each of the installed processor's memory cards. ► Memory latency time increases by 50% for memory access calls between processors 1 and 2 when the QPI wrap card when the QPI wrap card that is installed across QPI slot 1 and 2 has failed or is not installed. The same effect occurs with memory access calls between processors 3 and 4 when the QPI wrap card installed across QPI slot 3 and 4 has failed or is not installed. Failure of a QPI wrap card will be represented in the hardware event log as a loss of a redundant QPI lane for two of the processors.

Memory Installed	Operational characteristic
Memory installed on all memory cards for installed processors and the MAX5.	► QPI wrap cards must be removed from the server to accept QPI cables that are used to link the MAX5 memory expansion unit to the server. ► Performance improves as DIMMs are added to evenly populate all ranks on each memory controller for local memory. ► Performance improves significantly when more active memory calls use local memory on the system board instead of the memory that is located in the MAX5. ► In the rare instance of a memory card failure, VMware operational requirements can be satisfied if all of the memory is installed in the MAX5 or each processor has the same amount of memory disabled. Memory access is significantly slower if all of the memory is in the MAX5.

6.2.3 Processor installation order

The following list shows the processor installation requirements:

- The server ships with a minimum of two processors, which are installed in sockets 1 and 4.
- You can install additional processors in sockets 2 and 3 in either order.
- You need to install a QPI wrap card in QPI slot 3 and 4 after processor 2 is installed.
- You need to install a QPI wrap card in QPI slot 1and 2 after processor 3 is installed.
- The server functions in a slightly reduced capacity if the QPI wrap cards are not installed and the server is running as a stand-alone server. When processor 2 is installed and the QPI wrap card is not installed across slots 3 and 4, processor 2 must make a memory request through processor 3 or 4 to access the memory that is attached to processor 1. This type of memory call through two processors doubles the memory latency to access the memory that is attached to processor 1.

6.2.4 Processor installation tool

A processor installation tool ships with all processor options for this server. This tool assists with both installing and removing processors. The processor pins are extremely fragile so it is extremely easy to bend one or more of these pins while attempting to install or remove a processor. The process is more difficult on this type of server due to the deep well in which the processors are located between the power supply cage and the memory cage.

Bent processor pins can result in a number of possible problems:

- A complete processor failure
- Repeat failures of DIMMs at the same location within the two memory cards that are supported by that processor
- Repeat PCIe failures on the same slots that are associated with either processors 1 and 2 or processors 3 and 4
- Redundant QPI link failures between two processors, between a processor and the I/O hub, or between a processor and the MAX5

Figure 6-2 on page 225 shows the processor tool that is used to install a processor in this system safely or remove a processor from this system.

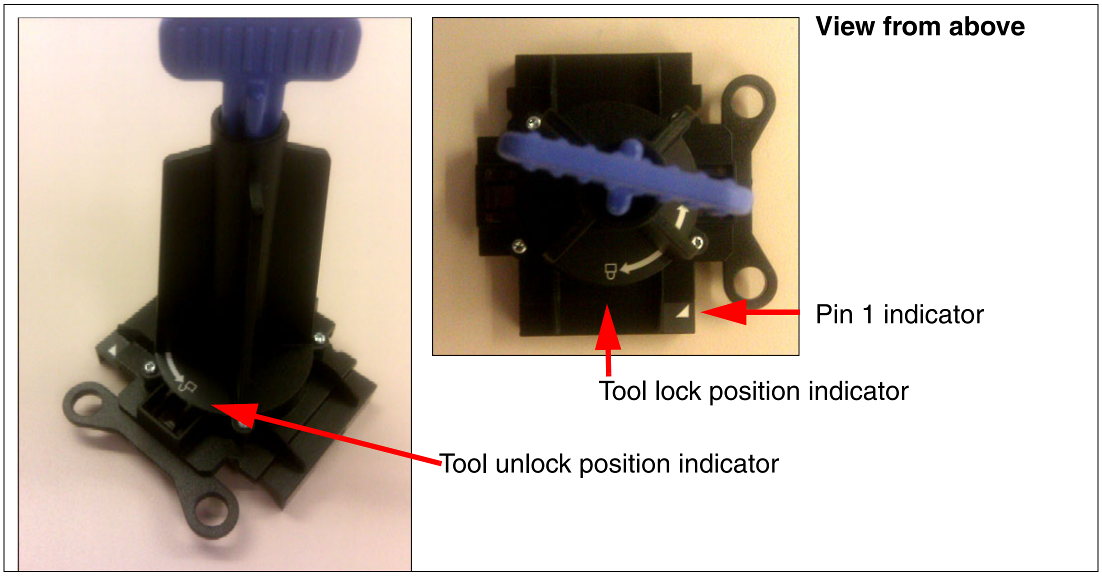


Figure 6-2 Processor installation and removal tool

To use the tool, perform the following steps:

1. When the processor is not installed in the server, ensure that you place the processor flat on an antistatic bag or mat, with the contact pads facing down.
2. Rotate the blue handle at the top of the tool so that the tool is in the unlocked position.
3. Place the tool over the top of the processor with the pin 1 indicator of the processor lined up with the pin 1 indicator on the tool.
4. Rotate the blue handle to the locked position to securely hold the processor inside the processor removal tool.
5. When the processor is being removed from the processor socket, lift the processor straight up out of the processor socket.
6. When placing the processor into the processor socket, ensure the processor socket has its protective cover removed and is open to receive the processor. Line up the pin 1 indicator on the tool with the pin 1 indicator on the processor socket and carefully lower the processor into the socket. The two holes on the processor removal tool need to line up with the two bolt heads that secure the processor socket to the system board.
7. When you have placed the processor in the desired location, rotate the blue handle back to the open position to release the tool from the placed processor. Carefully lift the tool from the processor.

6.3 Local memory configuration

Section 3.8, “Memory” on page 76 covers all of the various technical considerations regarding memory configuration for the System x3850 X5. You have a great deal of flexibility when configuring the memory for this server. However, as a result, you might configure a less than optimal memory environment.

Before you begin this section, understand the purpose of this server. File servers that are used to provide access to disk storage for other servers or workstations are affected less by poor memory latency speed than a database or mail server. Servers that are used as processing nodes in a high-performance cluster, database servers, or print servers for graphics printers require the best memory performance possible.

Remember these considerations when installing memory in the System x3850 X5:

- Due to the fact that no single processor has direct access to all PCIe slots, the QPI links that are used to communicate between the two processors are not only used for memory calls, but also for interfacing to PCIe adapters.
- When installing memory for multiple processors and not all memory cards are installed, only processing threads that are assigned to processor without local memory experience a 50% increase in memory latency. For a server with heavy I/O processing, the additional memory traffic also affects the efficiency of addressing PCIe adapters.
- Any memory that is installed on memory cards for an uninstalled or defective processor is not seen by the other processors on the server.
- For nonuniform memory access (NUMA)-aware operating systems when multiple processors are installed, you must install the same amount of memory on each memory card of each installed processor. See 3.8.2, “DIMM population sequence” on page 79 for details.
- You can achieve the best processor performance when memory is installed to support *Hemisphere Mode.* Hemisphere Mode is the required memory configuration to allow two 3850 X5 servers to scale. It is possible to have processor 1 and 4 in Hemisphere Mode, and not processor 2 and 3, to permit scaling. In this type of installation, having processor 1 and 4 in Hemisphere Mode improves the memory access latency for all processors. To determine the DIMM population for Hemisphere Mode, see 3.8.2, “DIMM population sequence” on page 79.
- You might consider installing all DIMMs of the same field-replaceable unit (FRU) to avoid conflicts that might prevent NUMA compliance or Hemisphere Mode support. These problems are most likely to occur when multiple DIMM sizes are installed. The server supports installing multiple-sized DIMMs in each rank, but the configuration becomes complex and difficult to maintain. Operating systems that depend on NUMA compliance inform you when the server is not NUMA-compliant. However, nothing informs you that the processors are not in Hemisphere Mode.
- It is better to install more smaller DIMMs than fewer larger DIMMs to ensure that all of the memory channels and buffers of each processor have access to the same amount of memory. This approach allows the processors to fully use their interleave algorithms to access memory faster by spreading access over multiple paths.

6.3.1 Testing the memory DIMMs

The DIMM components on a server are the most sensitive to static discharge. Improper handling of memory is the most likely cause of DIMM failures. Working in an extremely dry location dramatically increases the possibility that you will build up a static charge. Always use an electrostatic discharge (ESD) strap connected between you and a grounding point to reduce static buildup.

The best practice when installing memory is to run a memory quick test in diagnostics to ensure that all of the memory is functional. The following reasons describe why memory might not be functional:

- Wrong DIMM for the type of server that you have. Ensure that only IBM-approved DIMMs are installed in your server.
- The DIMM is not fully installed. Ensure that the DIMM clips are in the locked position to prevent the DIMM from pulling out of its slot.

- The DIMM configuration is invalid. See the DIMM placement tables in “x3690 X5 memory population order” on page 133 for details.
- A nonfunctional DIMM, failed DIMM slot, bent processor pin, or resource conflict with a PCIe adapter. Swap the DIMM with a functional DIMM, reactivate the DIMM by pressing F1-Setup, and retest:
  - When the problem follows the DIMM, replace the DIMM.
  - When the problem stays with the memory slot location, remove any non-IBM PCIe adapters, reactivate, and retest.
  - When the failure is in the same slot on the memory board, verify that the memory board is completely seated or swap memory boards with a known functional memory board location, reactivate, and retest.
  - When the problem remains with the memory board location, a bent processor pin might be the cause.
  - Contact IBM support for replacement parts.

The firmware of the server contains a a powerful diagnostic tool. Use the following steps to perform a simple test of a new memory configuration before placing the server into production:

1. During POST, at the IBM System x splash panel, press F2-Diagnostics, as shown in Figure 6-3.

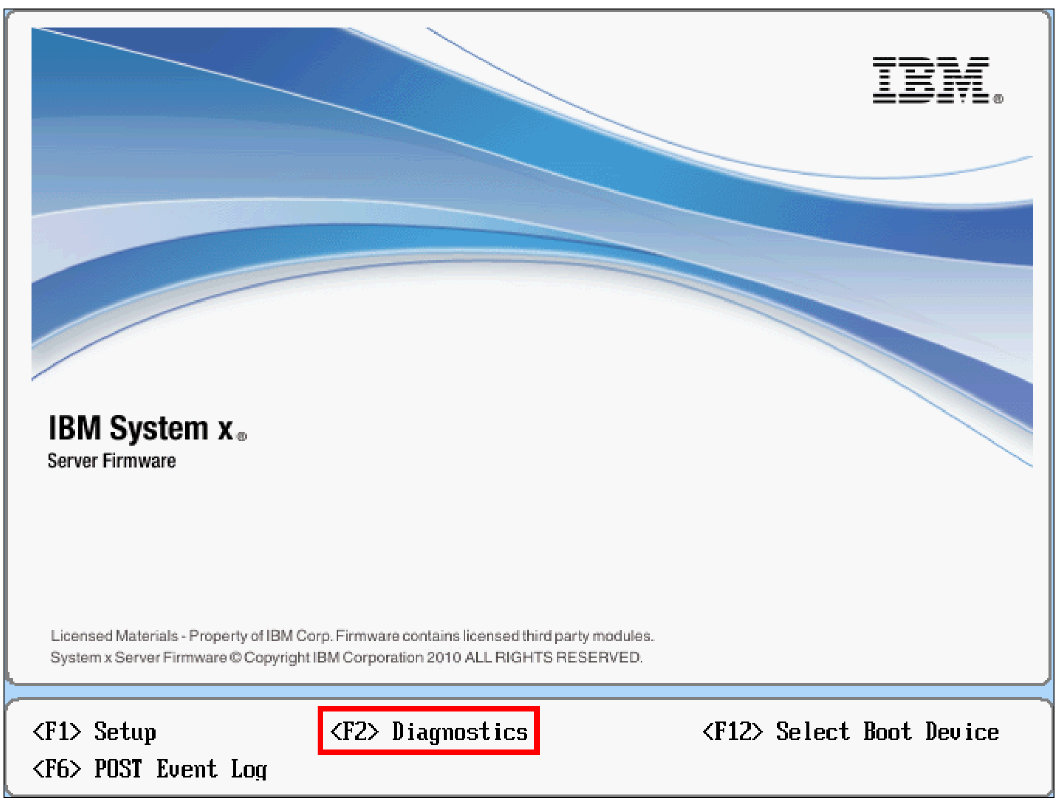


Figure 6-3 How to access diagnostics in POST

2. When the built-in diagnostics start, they start the Quick Memory Test for all of the memory in the server, as shown in Figure 6-4 on page 228. You can stop the Quick Memory Test at any time and run a Full Memory Test, which runs the same test patterns multiple times and takes five times longer than the Quick Memory Test. The only time that you want to run the

Full Memory Test is if you have an intermittent memory problem that you are trying to isolate.

Because the server identifies which specific DIMMs are experiencing excessive single-bit failures, it is far more efficient to swap reported DIMMs with similar DIMMs inside the server to see if the problem follows the DIMMs, stays with the memory slots, or simply goes away because the DIMMs were reseated.

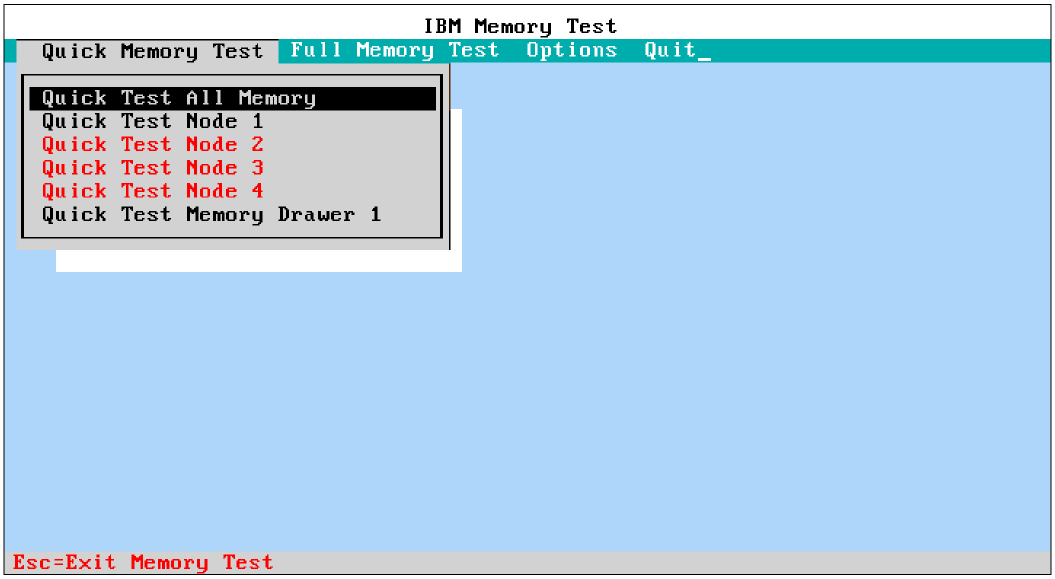


Figure 6-4 Start-up panel for the built-in diagnostics

3. The quick diagnostics continue to run, reporting the test that the quick diagnostics are performing currently and the length of time that it will take to complete that phase. If an error occurs, the quick diagnostics stop and indicate the memory errors that were encountered before progressing into more advanced diagnostics, as shown in Figure 6-5.

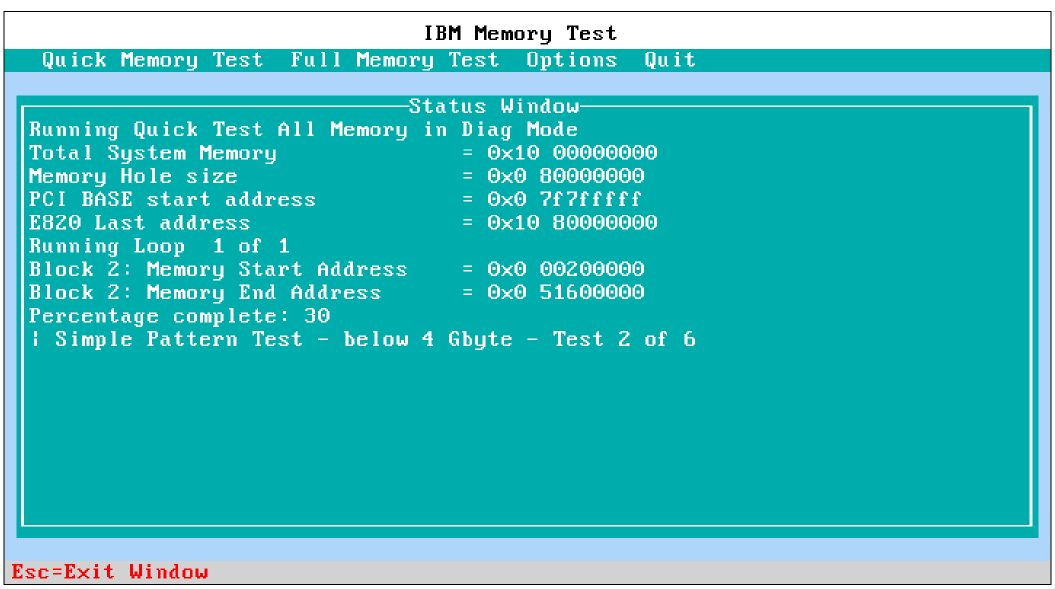


Figure 6-5 Quick Memory Test progress panel

You can terminate the diagnostics at any point by pressing Esc.

Important: Never warm-boot the server while running the built-in diagnostics. Several built-in functions that are used in the normal operation of the server are disabled during diagnostics to get direct results from the hardware. Only a normal exit from diagnostics or a cold boot of the server will re-enable those functions. Failure to perform this task correctly causes the server to become unstable. To correct this problem, simply power off the server and power it back on.

6.3.2 Memory fault tolerance

For servers with high availability requirements, using the memory mirroring or memory sparing configuration allows the server to continue to function normally in the rare event of a memory failure. See 2.3.6, “Reliability, availability, and serviceability (RAS) features” on page 28 for an explanation of memory mirroring and the sparing function.

Considering that memory is a solid-state device (SSD), it is unlikely that a DIMM failure will occur outside of the first 90 days of operation. Statistically, there is a higher risk for failure with power supplies, copper network adapters, storage devices, or processors. High availability is almost always a desired goal, but for true high availability, consider a cluster of host computers using common storage that allows virtual servers to be defined and moved from one host server to another host server.

If high availability is the most important aspect of your server (with cost and performance as secondary concerns), memory mirroring or sparing are good features to enable.

Use the following steps to establish memory mirroring or memory sparing:

1. Boot the server into F1-Setup by pressing F1.
2. From the System Configuration panel, select **System Settings** → **Memory**. See Figure 6-6.

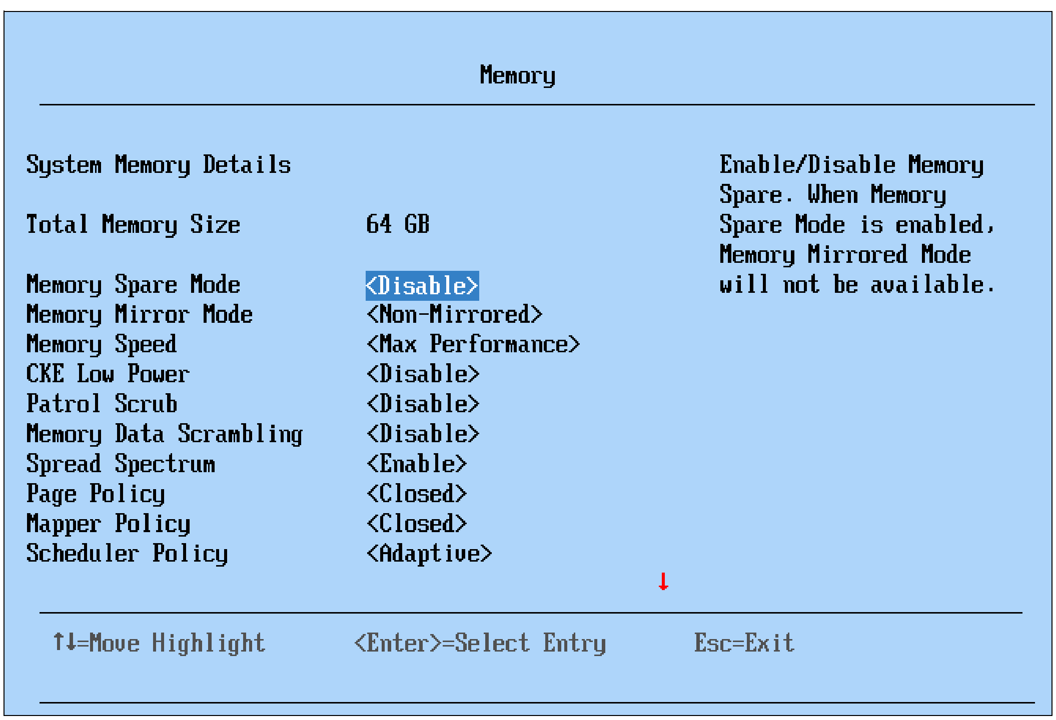


Figure 6-6 The memory configuration panel in F1-Setup

Figure 6-6 on page 229 shows the Memory configuration panel with your choice of performing memory sparing (Memory Spare Mode) or memory mirroring (memory mirror mode), but not both options at the same time. Select the desired option and reboot the server. If your memory population order does not support the requested option, the server reports a memory configuration error during the next reboot.

See 3.8.4, "Memory mirroring" on page 87 for the correct memory population order to support memory mirroring or sparing.

6.4 Attaching the MAX5 memory expansion unit

On top of the 1 terabyte (TB) of memory that you can configure on System x3850 X5, you can also configure an additional half of a TB of memory in the MAX5 memory expansion unit and attach the MAX5 to the server to increase the overall memory access performance and capacity.

It is best to use MAX5 with applications that can benefit from the increase in overall memory capacity. The most significant performance gains are in applications that *require* the additional memory that MAX5 can provide. If an application does not need the extra memory, the potential for performance gains is reduced.

The best way to populate memory DIMMs in the server and the MAX5 depends on how the applications being run address memory and the total amount of memory that is needed by the application. However, consider the following general rules:

- Always populate the DIMM sockets in the server first before installing DIMMs in the MAX5. Local memory has higher bandwidth and lower latency than MAX5 memory. MAX5 memory is limited by the speed of the QPI link.
- Where possible, install DIMMs so that Hemisphere Mode is enabled. Without Hemisphere Mode, performance can suffer considerably.
- You can use the MAX5 on an x3850 X5 that has only two processors installed. However, you get the best performance by having all four processors installed, installing all memory cards, and populating those memory cards fully with memory DIMMs.
- MAX5 adds an additional path to memory through dedicated QPI ports, resulting in potentially greater memory bandwidth. There can be instances where it is better to reserve DIMMs for use in the MAX5.

6.4.1 Before you attach the MAX5

Before you can attach and use the MAX5, the System x3850 X5 requires that you install the firmware that is shown in Table 6-2.

Table 6-2 Minimum firmware levels to support the MAX5 memory expansion unit

Type	Version	Build
IMM	1.19	YUOO75X
UEFI	1.32	G0E131A
FPGA	1.02	G0UD43A

The MAX5 also has Field Programmable Gate Array (FPGA) firmware that resides on it. The MAX5 firmware is updated at the same time that the server to which it is attached has its

FPGA firmware flashed. The recommendation is to reflash the FPGA on your system x3850 X5 after connecting the MAX5 to ensure that the FPGA levels on the MAX5 and server match.

The server might fail to complete POST because of a significant difference in FPGA code between the MAX5 and the server. Correct this situation by flashing the FPGA through the IMM while the server is plugged into power but not powered off. An entry in the hardware error log reports when the FPGA levels of the MAX5 and the server do not match.

The order of the flash types in Table 6-2 on page 230 is the sequence in which to perform the flashes. To update the server to this firmware or higher, you can use one of the following choices:

- When you have a compatible operating system installed, you can use the UpdateXpress System Packs Installer (UXSPI) tool that is described in 9.11, "UpdateXpress System Pack Installer" on page 511.
- Regardless of the version of operating system or when an operating system is not installed, you can use the Bootable Media Creator (BOMC) tool that is described in 9.12, "Bootable Media Creator" on page 514.
- After establishing a network connection and logging into the IMM, use the firmware update function, as described in 9.13, "MegaRAID Storage Manager" on page 521.

> **UEFI update is a two-step process:** The UEFI flash is a two-step process. When the UEFI flash completes, warm-reboot the server to at least the F1-Setup panel to allow the second half of the UEFI flash to complete. Failure to perform the required warm-reboot results in a corrupt version of the UEFI, which forces the server to boot into the recovery page of the UEFI until the situation is corrected by following the UEFI recovery process in the *Problem Determination and Service Guide* (PDSG).

For the latest firmware requirements for using the MAX5, see RETAIN tip H197572:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5085756

6.4.2 Installing in a rack

The QPI cables that are used to connect the MAX5 to the x3850 X5 are extremely short, stiff cables that can be easily damaged. For this reason, hardware ships with the MAX5 to allow the MAX5 to be mounted to the x3850 through a series of brackets and rail kits.

For cabling instructions, see the product publication *IBM eX5 MAX5 to x3850 X5 and x3950 X5 QPI Cable Kit and IBM eX5 MAX5 2-Node EXA Scalability Kit Installation Instructions*, which is available at this website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084861

6.4.3 MAX5 cables

The QPI cables that are used to cable the x3850 X5 and the MAX5 are extremely short and stiff. To plug the cables in, you must start the cable insertion on both sides of the cable into the correct receptors on both the x3850 X5 and MAX5 at the same time.

Use the following tips when plugging in the cables:

- QPI cables come packaged with reusable plastic boots that protect the fragile outside edges of the cable connectors, as shown in Figure 6-7 on page 232. It is a good idea to keep a set of these plastic boots for times when you want to remove the QPI cables for the movement of equipment from rack to rack or for servicing the unit.

Figure 6-7 Reusable QPI cable connector protective boot

- There are four QPI cables connecting all of the QPI ports of the x3850 X5 to all of the QPI ports on the MAX5. Figure 6-8 shows how to connect the cables between the QPI ports of the two units.

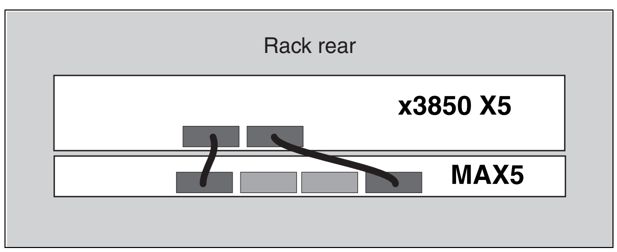


Figure 6-8 QPI cable installation

- The QPI cables are keyed to only be inserted one way. A quick visual check for cable orientation is to look for the 4U QPI or 1U QPI labels on the cable. The labels, along with the blue retainer release tab, are placed on what becomes the visible top of the cable when the cables are installed correctly.
- The ends of the cables are labeled to indicate which end to insert into the correct equipment. The 4U QPI end of the cables plugs into the x3850 X5. The 1U QPI end of the cables plugs into the MAX5.
- The cable end slides into the port until it clicks into place. You can disengage the retainer that holds the cable in place by pressing on the blue tab on the top of the QPI cable connection.
- QPI cables for connecting QPI ports 1 to 1 and 4 to 4 must be installed, even when only one processor is installed, to allow the MAX5 to be controlled by the server. If one of the cables is detached, the server does not power on or complete POST.
- You must install QPI cables for each of the processors installed to ensure full memory access to the MAX5. Table 6-3 on page 233 describes the QPI port on the back of the server with the corresponding processor socket in the server.

Table 6-3 QPI port relationship to the processor socket

QPI port number	Processor socket number
1	4
2	3
3	2
4	1

Important information about FPGA firmware: When attaching the MAX5 for the first time, reapply the FPGA firmware using the IMM firmware update tool after the server is plugged into ac power but prior to powering up the server.

Do not use Bootable Media Creator or USXPI to update the FPGA until the FPGA firmware is a match between the server and the MAX5. Mismatched FPGA makes the server unstable and the server can power off during the flash, corrupting the FPGA. This situation results in hardware replacement, because there is no recovery for corrupt FPGA.

Figure 6-9 shows the back of the x3850 X5 with the attached MAX5.

Figure 6-9 Lab photo of an x3850 X5 (top) attached to a MAX5 (bottom)

Important power considerations: With the server or MAX5 plugged into ac power and the scaled units' power turned off, the QPI cables still have active dc power running through them. Only unplug or plug in the QPI cable when neither the MAX5 nor the server is still plugged into ac power. Failure to follow this rule results in damaged circuits on either the MAX5 memory board or the processor board of the server.

6.4.4 Accessing the DIMMs in the MAX5

To add or subtract memory on the MAX5, you can remove the memory tray from the front of the chassis. Use the following steps to access the DIMMs:

1. Remove ac power from all of the server's power supplies and from the two MAX5 power supplies. Because the QPI cables are already held in alignment by the memory expansion

chassis, it is not a requirement to remove the QPI cables before removing the memory board.

2. Remove the front bezel of the MAX5 by pressing in on the tab buttons on both sides of the bezel. The bezel can be pulled away from the MAX5.
3. As shown in Figure 6-10, there are two blue release tabs that, when pressed to the sides of the enclosure, allow you to pull out the cam levers that are used to begin to pull out the memory tray.

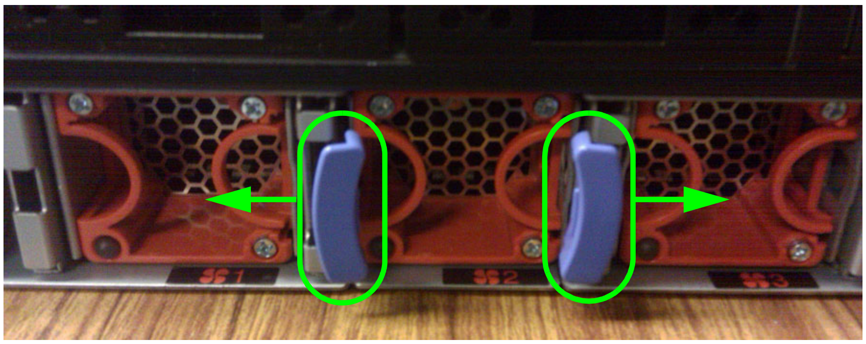


Figure 6-10 Blue release tabs for the memory tray cam levers

4. You can pull the memory tray out about 30% before it stops. This design allows you to get a better grip of the tray on either side and then use a finger to push in another set of blue release tabs on either side of the tray, as shown in Figure 6-11.

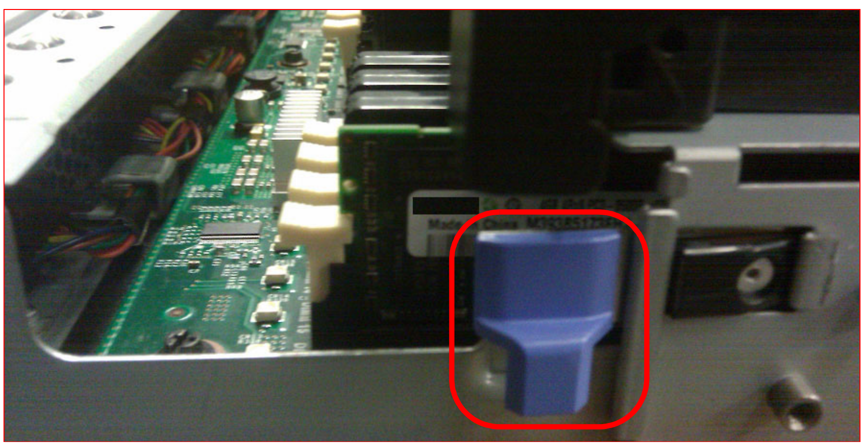


Figure 6-11 Final release tab for removing the MAX5 memory tray

5. Slide the memory tray completely out and place it on a flat work surface to work on the memory.

> **Important**: You must remove the ac power from both the server and the attached MAX5 before removing the memory tray. The FPGA of both the server and the MAX5 is still an active component when the server is powered down but not removed from utility power. FPGA components exist on both the server and the MAX5. Removing the memory board with the ac power still active damages the FPGA components of both the server and the MAX5.

For memory population order, see 4.9, “Storage” on page 145.

After you install and configure the MAX5 properly, a successful link between the server and the MAX5 can be confirmed when both units power on and off as one complete unit and the memory in the MAX5 can be seen by the server. You also see the following message during POST:

```
“System initializing memory with MAX5 memory scaling”
```

6.5 Forming a 2-node x3850 X5 complex

The x3850 X5 server is the easiest server to scale of any of the predecessors of this class of server. When all of the prerequisites of scaling to X3850 X5 servers have been met and the servers are scaled, you can apply all of the updates that are provided by IBM for the stand-alone chassis to the scale chassis as if it were a single unit. The only exception to this rule is the replacement of either the I/O shuttle or the processor board on either node.

For rack and cable installation instructions for scaling two x3850 X5 servers, see the product publication *Installation Instructions for the IBM eX5 MAX5 to x3850 X5 and x3950 X5 QPI Cable Kit and IBM eX5 MAX5 2-Node EXA Scalability Kit,* which is available at this website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084861

6.5.1 Firmware requirements

Before you attempt to scale to x3850 X5, the firmware of the two servers needs to match to a code level that meets or exceeds the minimum firmware levels that are shown in Table 6-4. You must apply the necessary firmware to get the two servers to matched levels while the two servers are not connected by QPI cables.

Table 6-4 Minimum firmware levels to support two scaled x3850 X5 servers

Type	Version	Build
IMM	1.14	YUOO73K
UEFI	1.23	G0E122D
FPGA	1.01	G0UD29C

You might experience these situations if code levels do not match:

- Severe FPGA mismatch: Neither node can power on. The hardware event log, accessed through the IMM management port, reports an FPGA mismatch.
- Minor FPGA mismatch: The system powers on, but it reports an FPGA mismatch during POST, prior to the F1 Splash panel, and in the hardware event log of the IMM. Expect spontaneous reboots, memory failures, processor failures, QPI link failures, and machine check errors.
- Severe IMM mismatch: The secondary node does not power up. The IMM of the primary node is unable to access events occurring in the secondary node.
- Minor IMM mismatch: The primary IMM is unable to collect all of the hardware information from the secondary node. A newer version of the IMM prevents both nodes from powering up.
- UEFI mismatch: When two scaled x3850 X5 servers are powered up, the IMM loads the UEFI from only the primary node. The UEFI of the secondary node is never used. The firmware level of the primary node becomes the running firmware level of both nodes.

6.5.2 Processor requirements

To maximize performance, both nodes must have all four processors and all memory cards installed and populated. However, the x3850 X5 does support 2-node configurations with only two processors installed in each node.

The following list shows the processor requirements to perform scaling:

- Functioning processors must exist in processor sockets 1 and 4 of both servers.
- The processor specification must match among all processors in both servers.

6.5.3 Memory requirements

For scaling to be permitted, you must install the memory in such a way so as to support *Hemisphere Mode* on the processors of both nodes. See Table 3-15 on page 81 to see which memory configurations enable Hemisphere Mode.

To maximize performance, all four processors in both nodes must be in Hemisphere Mode. If needed, the servers can be scaled with only processors 1 and 4 in Hemisphere Mode in both nodes, but this type of configuration cannot support operating systems, such as VMware. For other less sensitive operating environments, if the unlikely failure of a memory card takes processor 1 or 4 out of Hemisphere Mode, you can restore the scaled environment by swapping memory cards from processor 2 or 3.

6.5.4 Cabling the servers together

The QPI cables that are used to cable the two x3850 X5s together are extremely short and stiff. To plug the cables in, you must start the cable insertion on both sides of the cable into the correct receptors on both x3850 X5s at the same time.

Use the following tips when plugging in the cables:

- QPI cables ship packaged with reusable plastic boots that protect the fragile outside edges of the cable connectors, as shown in Figure 6-12. It is a good idea to keep a set of these plastic boots if you want to remove the QPI cables for the movement of equipment from rack to rack or for servicing the unit.

Figure 6-12 Reusable QPI cable connector protective boot

There are four QPI cables connecting all of the QPI ports of both x3850 X5 servers. Figure 6-13 shows the cable routing.

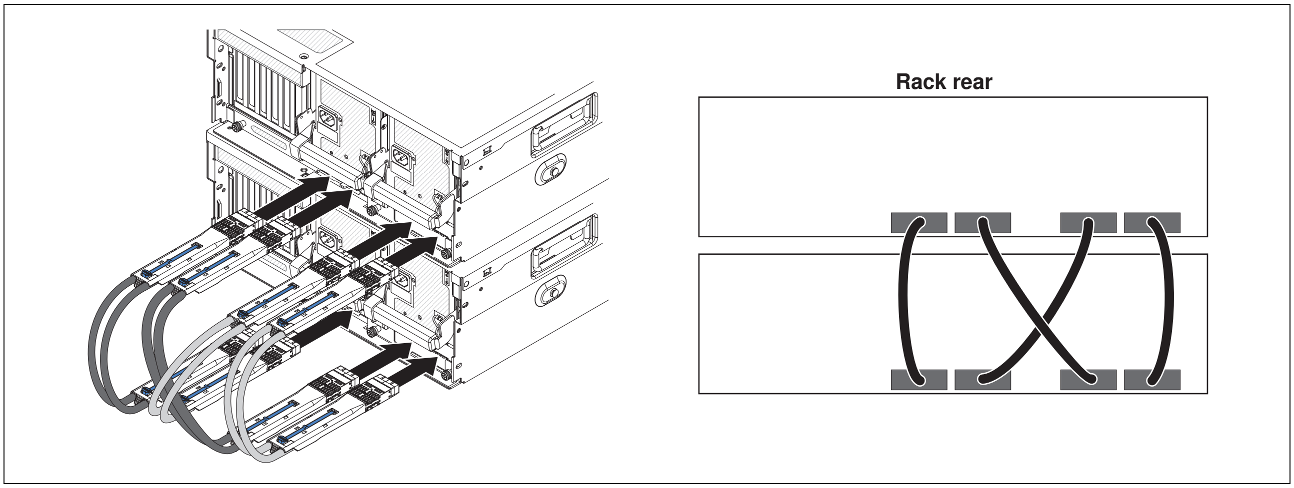


Figure 6-13 Cabling diagram for 2-node x3850 X5

Consider this important information:

- The QPI cables are keyed to be inserted only one way. A quick visual check for cable orientation is to look for the blue retainer release tab on the top of each cable end. The blue retainer release tab is placed on what becomes the visible top of the cable when the cable is installed correctly.
- The ends of the cable are labeled to indicate which end can be inserted into the correct equipment. The server that becomes the primary node has the end of the QPI plugged into it with the labels J2 and J1, as shown in Figure 6-14. The secondary node has the end of the QPI cable plugged into it with the labels J4 and J3. All four QPI cables must be oriented in the exact same manner. Neither server can power on if the cables are not consistent in their orientation. Figure 6-14 shows how the cable ends are labeled.

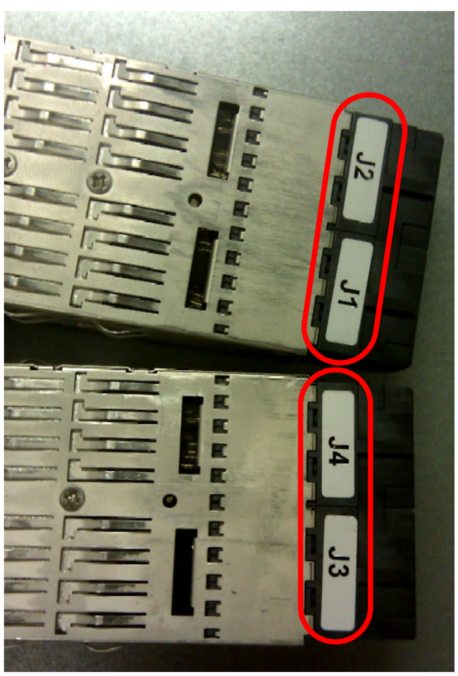


Figure 6-14 QPI cable end labels: J1 and J2 attach to the primary node

- The cable end must slide into the port until it clicks into place. You can disengage the retainer that holds the cable in place by pressing on the blue tab on the top of the QPI cable connection.
- QPI cables for connecting QPI ports 1 to 1 and 4 to 4 must be installed, regardless of whether processors 2 and 3 are installed, to allow the secondary node to be controlled by the primary node. If one of the cables is detached, the server cannot power on.
- QPI cables must be installed for each of the processors installed. Table 6-5 describes the QPI port on the back of the server with the corresponding processor socket in the server.

Table 6-5 QPI port relationship to the processor socket

QPI port number	Processor socket number
1	4
2	3
3	2
4	1

After you attach the cables, connect the power to both nodes of the complex.

When the power status light changes from a rapidly blinking light to a slow blinking green light, press the power button on the primary node. Both nodes power up. If both nodes do not power up, double check that the required matching firmware is on both nodes of the server. You must disconnect the QPI cables to determine if the required matching firmware is on both nodes of the server.

If neither node powers up or the secondary node does not power up, consider swapping QPI cables between port 1 on both nodes to one of the cross-port cables between ports 2 and 3. If this action proves unsuccessful, do the same task with the QPI cables between port 4 of both nodes and again one of the two crossed QPI cables between ports 2 and 3. There are additional communication lanes on port 1 for synchronization of the CPU frequencies and additional communication lines on port 4 for communication between the FPGA of the primary node and the secondary node. If this approach works, check the IMM event log for any QPI link failures between processors to see if one of the cross-linked QPI cables between ports 2 and 3 needs to be replaced. Frequently, one of the critical QPI cables was not completely seated.

After the complex is up and running, double check the IMM event log for any QPI link failures. The error message reports which processors experience the error. Based on Table 6-5, the processor reporting the QPI link error points to a QPI cable on a specific port. Provided the two nodes are running matched IMM and FPGA, the problem can be the processor or QPI port on either end of the QPI cable or the cable itself. The cable is the most likely point of failure.

After the two nodes boot as a single entity, any firmware flashes applied to the primary node are automatically applied to the secondary node.

6.6 PCIe adapters and riser card options

This section describes considerations to remember for determining how to use your PCIe slots, depending on the types of PCIe riser cards that you have installed. The x3850 X5 is an enterprise server that is designed to function in a high availability cluster or powerful stand-alone database server with built-in fault tolerance for power and cooling, processors, memory, and PCIe buses.

6.6.1 Generation 2 and Generation 1 PCIe adapters

All of the PCIe slots in the x3850 X5 are at Generation 2 (*Gen 2*) specification. Besides additional error correction and addressing advancements, Gen 2 means that all of the slots of this server exchange data twice as fast as servers with Gen 1 PCIe slots.

Table 6-6 describes the theoretical limits of each of the common types of PCIe adapters.

Remember that theoretical limits are based on the mathematics of the frequency and data width of the bits that are transmitted over the interface. Theoretical limits do not include the necessary communications to maintain the required protocols to interface between intelligent devices. Theoretical limits also do not consider the inability to maintain a steady flow of data in full duplex.

Table 6-6 Theoretical data transfer limits of Gen1 PCIe slot types versus Gen2 PCIe slot types

PCIe adapter/slot type	Generation 1 Limit	Generation 2 Limit
x1	500 MBs	1 GBs
x4	2 GBs	4 GBs
x8	4 GBs	8 GBs
x16	8 GBs	16 GBs
x32	16 GBs	32 GBs

PCIe adapters connect to the processors via the I/O hub. The purpose of the I/O hub is to combine the various data streams from each of the PCIe slots into a single aggregate link to the processors using a dedicated QPI link to each processor. The x3850 X5 has two I/O hubs. The data transfer rate of the QPI link is negotiated between the processor and the I/O hub.

Table 3-9 on page 74 shows the QPI link speeds based on the type of installed processors. The I/O hub supports the highest QPI link speed that is shown in the table, 6.4 gigatransfers per second (GT/s). You can also adjust the QPI link speed to conserve power by booting into F1-Setup, selecting **System Settings** → **Operating Modes**, and setting QPI Link Frequency to values other than the default Max Performance, as shown in Figure 6-15 on page 240.

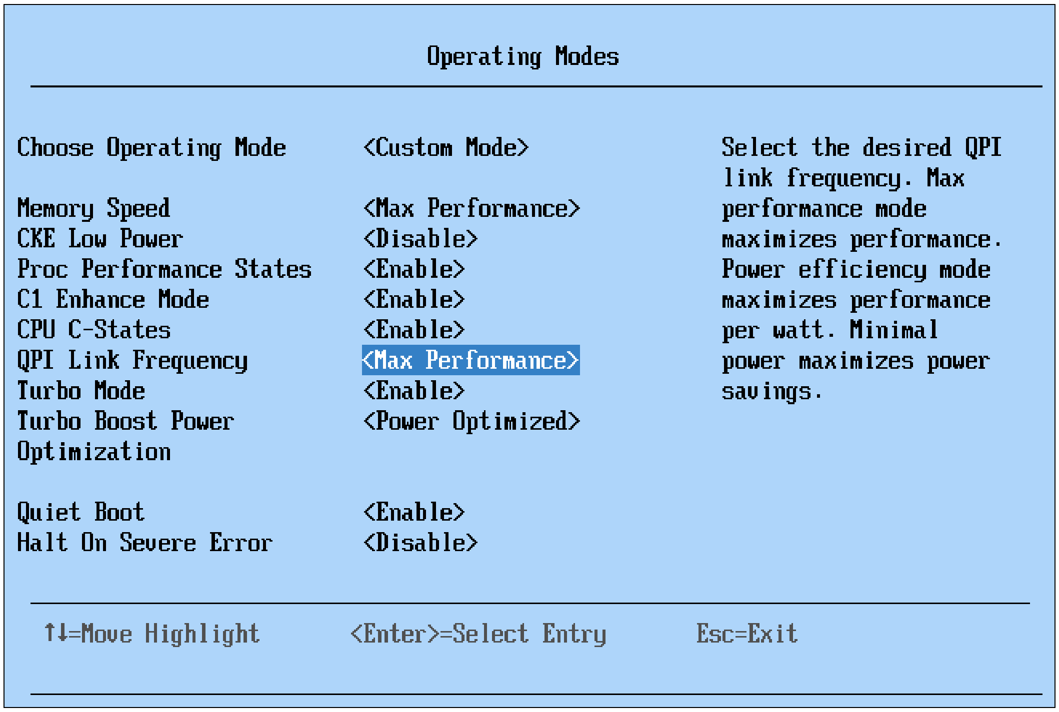


Figure 6-15 QPI Link Frequency setting

PCie adapter compatibility

IBM ServerProven tests IBM and non-IBM adapters that have been proven to function correctly in the System x3850 X5 server. See the following page for specifics:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/xseries/7148.html

Backward compatibility of Gen2 PCie slots to Gen1 adapters

All Gen2 PCIe slots are backward-compatible to Gen1 adapters; however, not all PCIe adapter vendors adopted the optional specification of Gen1 that allows a Gen2 PCIe slot to recognize a Gen1 adapter. When you install a Gen1 PCI adapter that is not recognized by the server, consider forcing the PCI slot in which the adapter is installed to a Gen1 slot in F1-Setup by selecting **System Settings** → **Devices and I/O Ports** → **PCIe Gen1/Gen2 Speed Selection**. Figure 6-16 on page 241 shows the resulting panel and the available selections. The change takes effect after a cold reboot of the server.

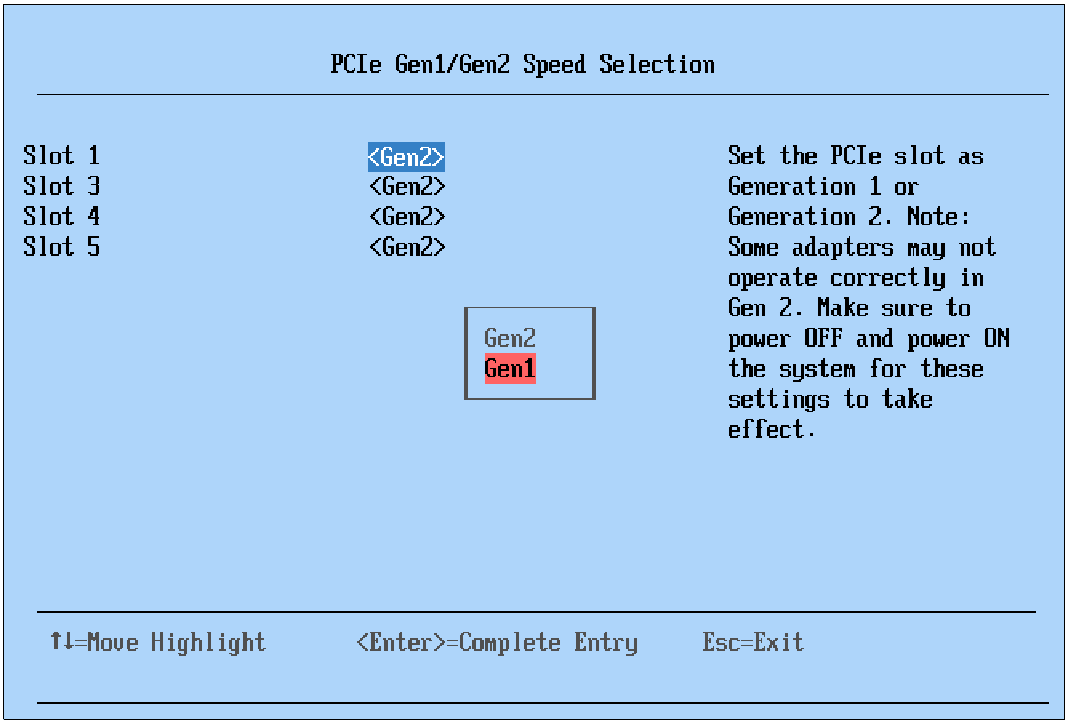


Figure 6-16 PCIe slot speed selection panel to force Gen1 compliance

Non-UEFI adapters in an UEFI environment

A number of Gen1 PCIe adapters were designed prior to the implementation of UEFI. As a result, these adapters are not recognized or might not have UEFI drivers that allow the adapter to function in an UEFI environment.

The way that the server supports these non-UEFI adapters is via a setting, Legacy Thunk Support, which is enabled by default. Legacy Thunk Support mode places the non-UEFI-aware Gen1 adapter into a generic UEFI wrapper and driver, which allows you to update the firmware of the adapter to support UEFI.

The recommendation is that all installed adapters either support UEFI as standard or be updated to support UEFI, because Thunking only provides limited support for non-UEFI adapters in an UEFI environment. For example, a legacy adapter in a Thunk UEFI wrapper cannot be seen in **System Settings** → **UEFI Adapters and Device Drivers**, nor can it natively access memory locations above 4 GB.

If you have previously disabled Thunking, you can re-enable it by using F1-Setup and selecting **System Settings** → **Legacy Support**. Figure 6-17 on page 242 shows the Legacy Support panel.

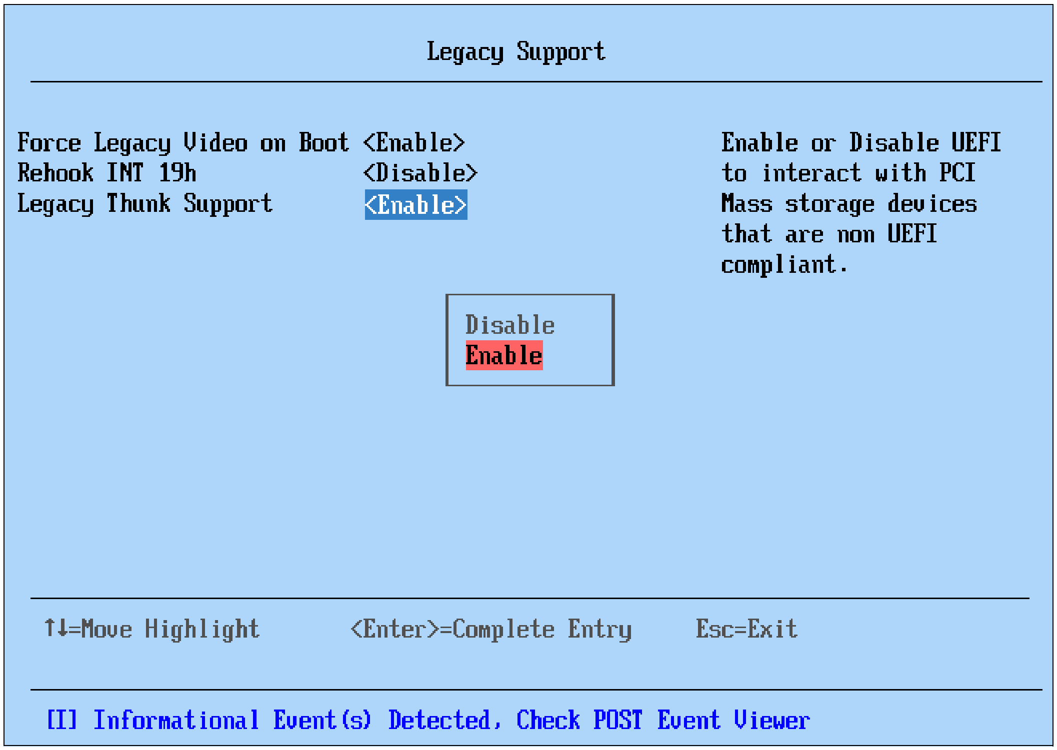


Figure 6-17 Legacy Support panel

Another possible solution to this problem is booting the server in *Legacy Only* mode. This mode allows both non-UEFI-aware operating systems and PCIe adapters to function as they function on a non-UEFI server. Many of the advanced memory addressing features of the UEFI environment are not available to the operating system in this mode.

Booting the server to an operating system in this mode allows you the ability to apply firmware updates to the Gen1 adapter that is not recognized in an UEFI environment. To enable this feature from within F1-Setup, select **Boot Manager** → **Add Boot Option**, scroll down, and select **Legacy Only** when it is displayed in the list of options. Figure 6-18 on page 243 shows where Legacy Only is located in the list of available boot options. If Legacy Only is not listed, it has already been added to the Boot Manager.

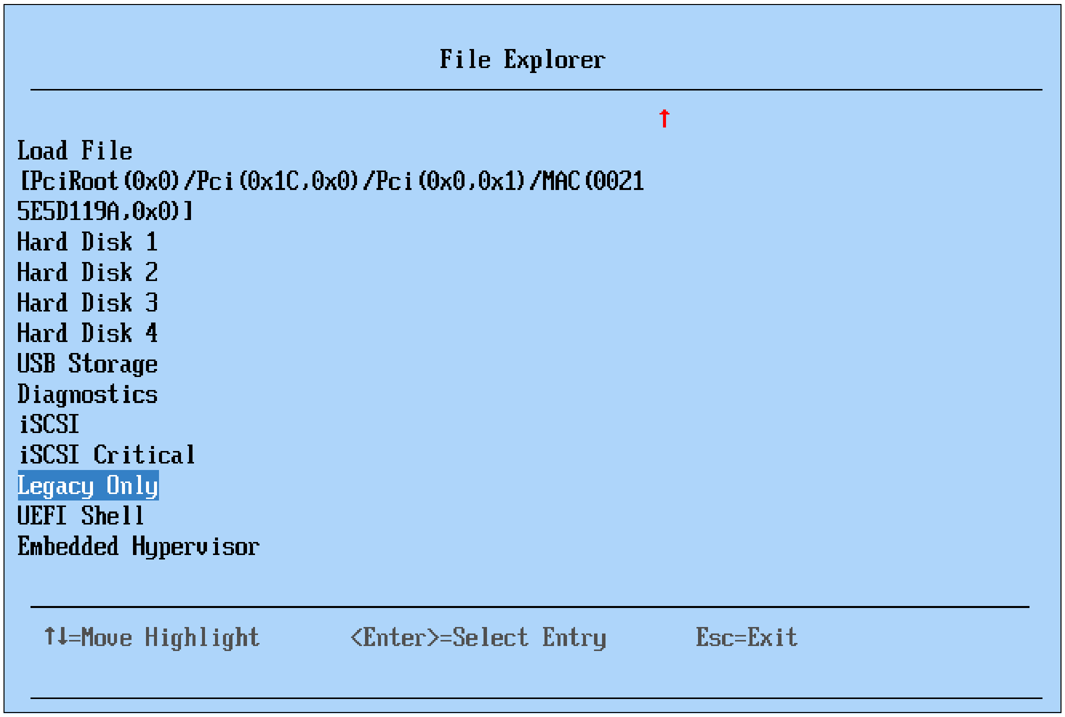


Figure 6-18 Selecting Legacy Only for a Boot option

Besides adding Legacy Only to the Boot Manager, you must also change the sequence to place Legacy Only at the top of the boot sequence. To change the boot order from within the Boot Manager panel, select **Change Boot Order.** Figure 6-19 shows how the boot order panel is activated by pressing Enter and then selecting the item to move by using the arrow keys. When Legacy Only is selected, use the plus (+) key to move it up the panel.

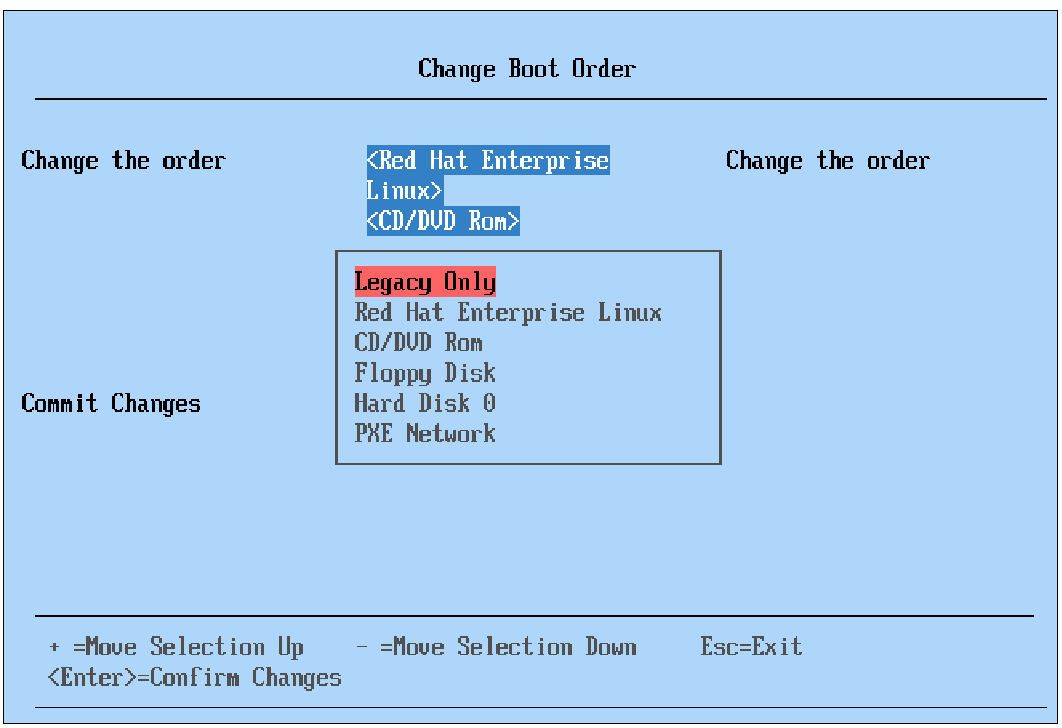


Figure 6-19 Change Boot Order panel

Microsoft Windows 2008 x64: When installed on an UEFI server, Microsoft Windows 2008 x64 will install Microsoft Boot Manager as part of the boot sequence. Regardless of how you change the boot sequence in the boot manager, Microsoft Boot Manager will always be at the top of the sequence.

When you install this same operating system with **Legacy Only** enabled, Microsoft Boot Manager is not installed as part of the boot manager. Removing the Legacy Only option from the boot manager will prevent the server from booting into the installed Windows 2008 x64.

6.6.2 PCIe adapters: Slot selection

A single-node x3850 X5 provides two independent zones of I/O interfaces. A 2-node x3850 X5 provides four independent zones of I/O interfaces. Each zone contains two processors and one I/O hub. Table 6-7 shows which I/O devices and processors communicate directly in each zone. Processors that need to initiate an I/O in a zone, of which they are not a member, must use their QPI links to the processors in the other zone to complete the I/O in that zone.

Table 6-7 I/O interface zones

Zone	Processors	I/O interfaces
1	1 and 2	► PCIe slots 5, 6, and 7 ► Two onboard 1GbE ports ► IMM, all USB ports, and SATA DVD ► 8x PCIe SAS port
2	3 and 4	PCIe slots 1, 2, 3, and 4

When installing teamed network adapters or multipath fibre host bus adapters (HBAs), each member of the team must be placed in separate zones to maximize the throughput potential of this server and minimize the loss of a single processor.

Consider these key points about these slots:

- PCIe slot 1 is an x16 slot to support possible future I/O adapters. It can be used to hold any PCIe adapter.
- PCIe slot 2 is actually an x4 with an x8 mechanical connection. Do not use this slot for an SSD RAID controller.
- PCIe slot 7 is an x8 slot and is specifically designed to support an Emulex 10GbE adapter that is designed for this server.
- The RAID SAS port is an x8 PCIe slot that supports a wide variety of ServeRAID and SAS HBAs. This slot is not suitable for adapters, such as SSD controllers, that generate a lot of heat; these adapters need to be installed in a slot at the rear of the server with better airflow.

While true performance on a given PCIe adapter can largely depend on the configuration of the environment on which it is used, there are general performance considerations with respect to the x3850 X5 server.

The I/O hub supports 36 lines of PCIe traffic with a combined bandwidth of 36 GBs. Each processor's QPI link to the I/O hub is capable of a maximum throughput of 26 GBs, depending on the processors installed on the server. With only one processor installed, the maximum combined bandwidth of all the PCIe lanes is reduced to the maximum bandwidth of a single

QPI link. If two matched processors of any QPI link speed are installed, this limit is no longer an issue.

Of all of the I/O adapters that can be installed on the server, the ServeRAID and 6 Gb SAS controllers managing SSD drives are the only adapters that can approach the theoretical limits of the x8 PCIe slot. Therefore, when you use SSD drives, connect no more than eight SSD drives to a single controller for the best performance of the SSD drives. Also, only use x8 slots to host the controllers that manage your SSD drives.

A single ServeRAID controller managing a single 4-drive SAS HDD array will function within the theoretical limits of an x4 PCIe slot. In this case, the mechanical nature of the HDD drives will limit the maximum throughput of data that passes through the PCIe slot.

The dual-port 8 Gbs Fibre Channel, 10Gbs Ethernet, and 10 Gbs Converge Network adapters (CNA) are all capable of approaching the theoretical limits of an x4 Gen2 PCI slot and might perform better in an X8 Gen2 slot.

6.6.3 Cleaning up the boot sequence

One of the most overlooked steps in completing a hardware setup is deciding from what you are going to boot. The server might have one or more ServeRAID controllers for internal drives or perhaps another ServeRAID adapter for external drives. You might also be using one or more Fibre Channel HBAs to access a SAN and you might have Preboot eXecution Environment (PXE) or iSCSI defined to boot to an operating system over the network.

By default, your server and the installed options came with the ability to boot from any of these sources other than USB storage devices. On every boot, the server is going to recognize each of these boot choices, determine if the bootable media device is attached, and add the optional ROM support to the boot ROM to determine the correct device from which to boot. This process adds time to the boot process.

To minimize this loss of time, you can disable the boot options for adapters from which you know you are never going to boot. The following sections describe common methods of disabling boot options.

Legacy only mode

When the server is instructed to boot to Legacy Only mode, the best way to disable unwanted boot sequences is to disable them in F1-Setup by selecting **System Settings** → **Device and I/O Ports** → **Enable / Disable Legacy Option ROM(s)**. Figure 6-20 on page 246 shows the available options. You need to know the specific PCIe slots that were used for each adapter, so that you will know which slot to leave enabled.

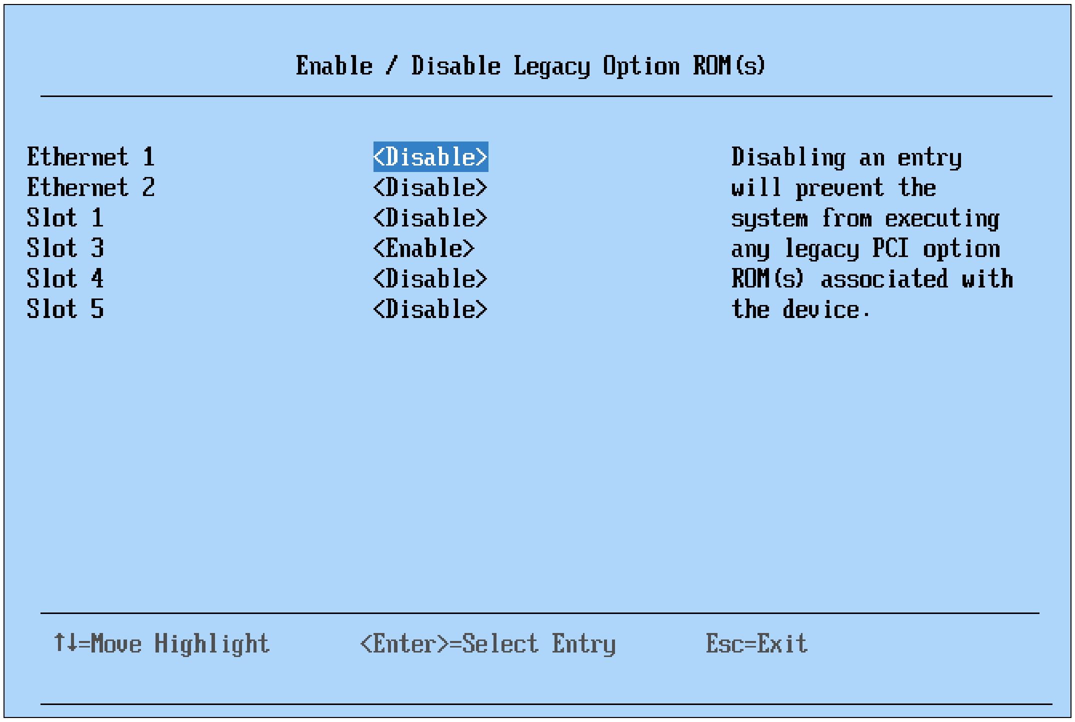


Figure 6-20 Legacy option ROM states

When booting from SAN with multiple paths for redundancy, you will need to enable the legacy option ROM for both HBAs.

The default UEFI mode

On the x3850 X5, you can sequence the order that the UEFI will search the various attached devices to locate a boot device. You can shorten the time that it takes to perform the search by moving the adapter that contains the boot device to the top of the list.

In UEFI mode, PXE boot can be disabled for the onboard network interface card (NIC) ports through F1-Setup by selecting **System Settings** → **Network** → **PXE Configuration** and then by selecting the port on which you want to disable PXE boot. Figure 6-21 on page 247 shows the panel that you will see to disable PXE boot on one of the two onboard network ports.

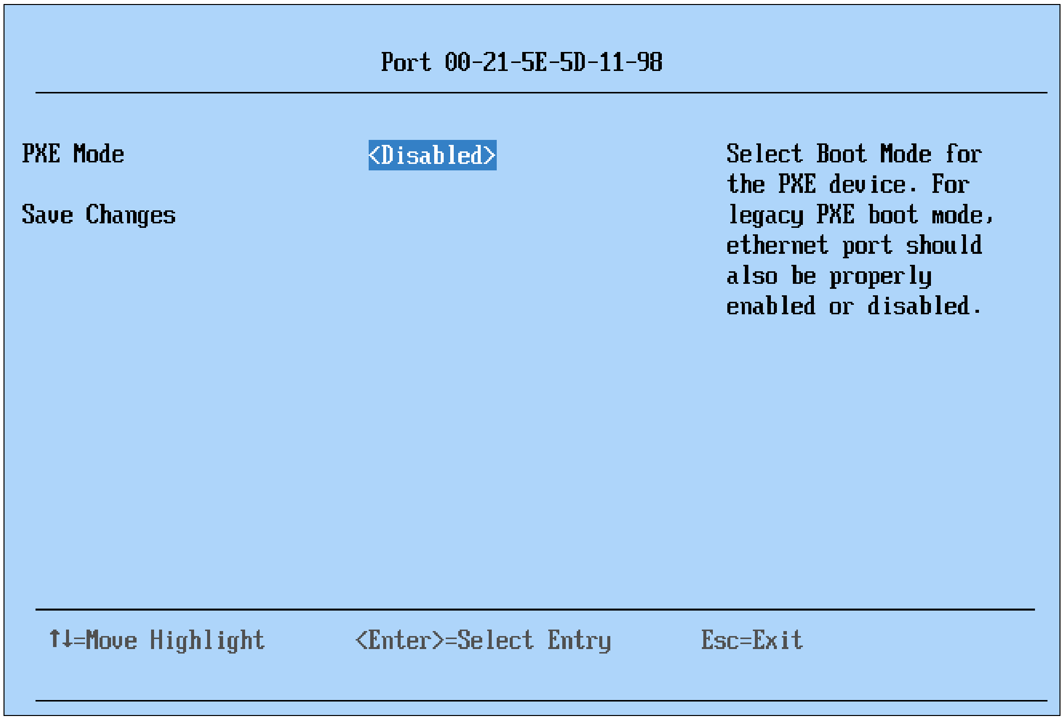


Figure 6-21 Disabling PXE boot of the onboard network ports

Other PCI adapters can have their boot option ROM disabled from within their configuration panels. To access individual adapter configuration panels from F1-Setup, select **System Settings** → **Adapters and UEFI Drivers** and press Enter. Figure 6-22 shows the selections that are presented by this panel when accessed.

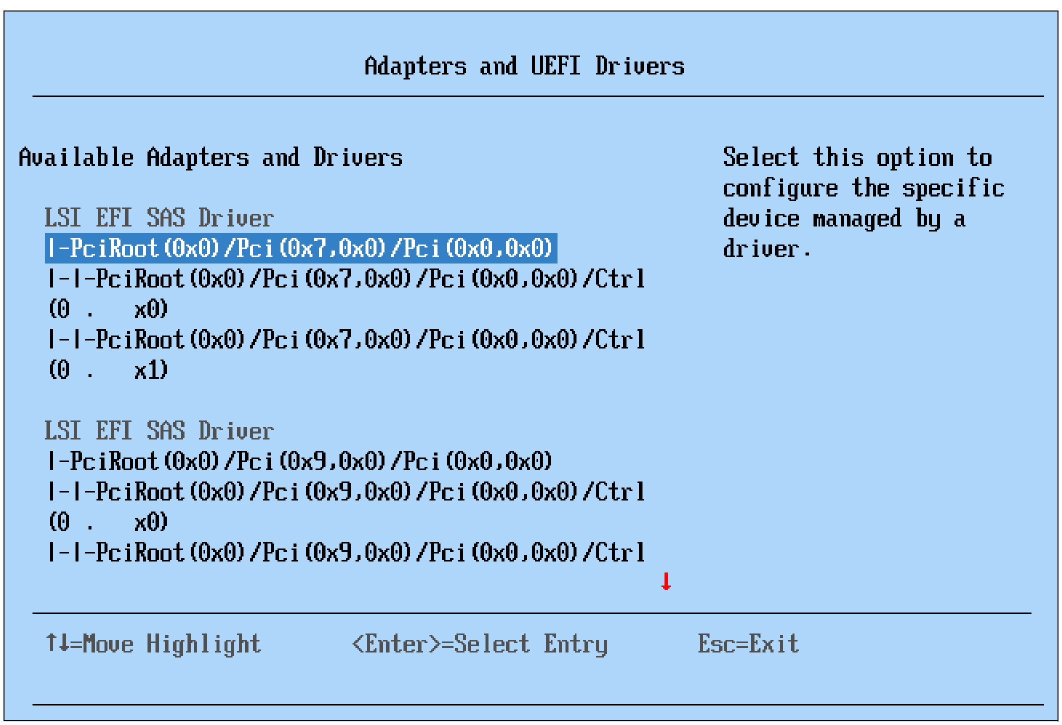


Figure 6-22 Accessing adapter-specific configuration information

To enter the configuration of a specific adapter, select the PciROOT directly under the adapter name. When you have multiple controllers of the same type, selecting the PciROOT of any of the same adapter types will select all of them and display a panel that allows you to select the specific adapter from within the configuration routine of the adapter type. Figure 6-23 demonstrates this process when two ServeRAID adapters are installed on the server.

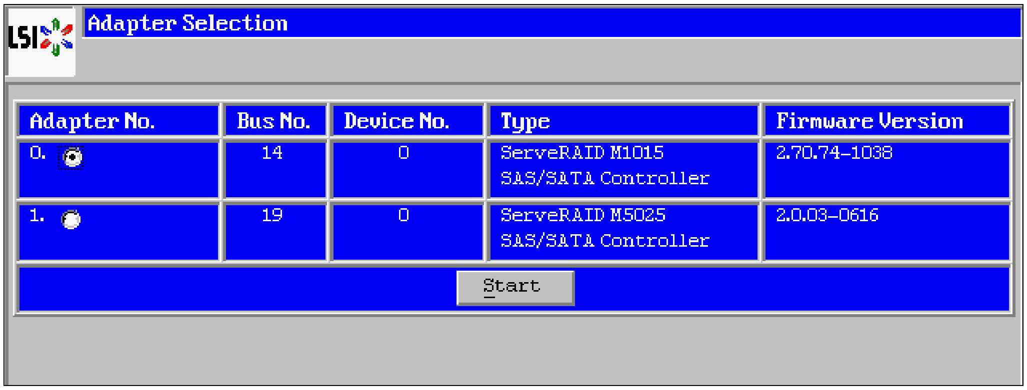


Figure 6-23 LSI adapter selection panel from within the LSI configuration panel

The controller configuration properties for adapters that are similar to the ServeRAID or Fibre Channel HBAs you are looking for are displayed, as shown in Figure 6-24.

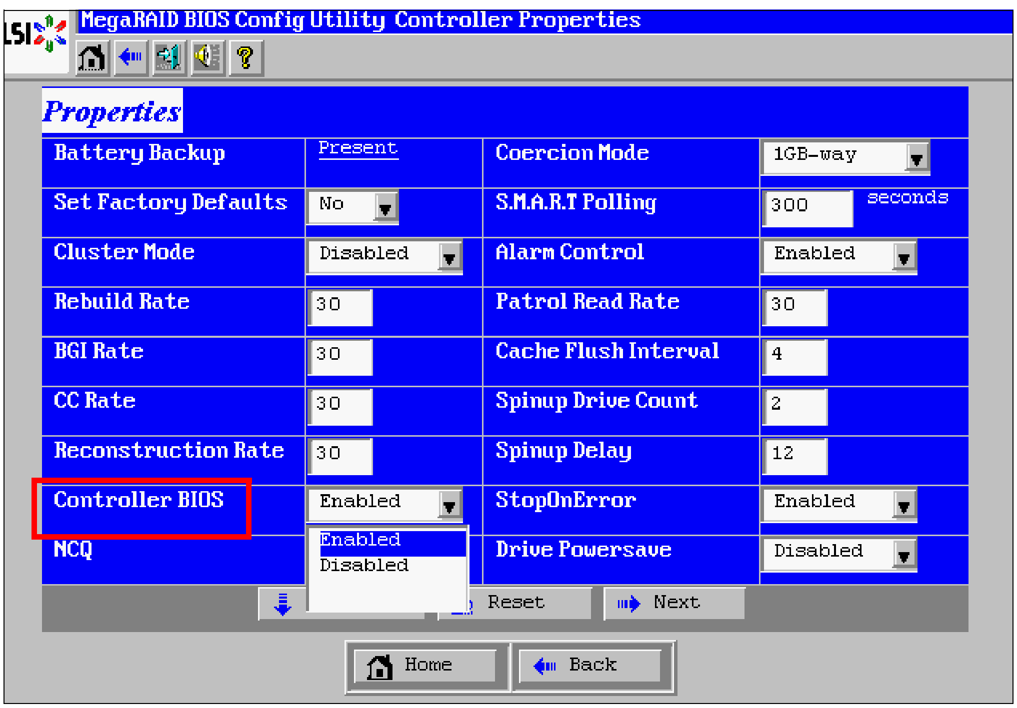


Figure 6-24 ServeRAID BIOS Config Utility Controller Properties: Disabling Boot ROM

Although it is not stated or obvious from the description, disabling the Controller BIOS only disables the Boot ROM execution during POST. All of the other operating characteristics of the adapter BIOS and firmware remain intact.

Option ROM execution order

Regardless of when you boot in Legacy Only mode or UEFI mode, you can control from which device you want to boot. This function is important when multiple storage adapters are installed in the server. To control the boot sequence, from within F1-Setup, select **System Settings → Device and I/O Protest Option ROM Execution Order**. Figure 6-25 shows what the panel looks like after pressing Enter on the list of possible choices. Use the up and down arrow keys to select a specific entry to move and use the plus or minus keys to move the item up or down the list.

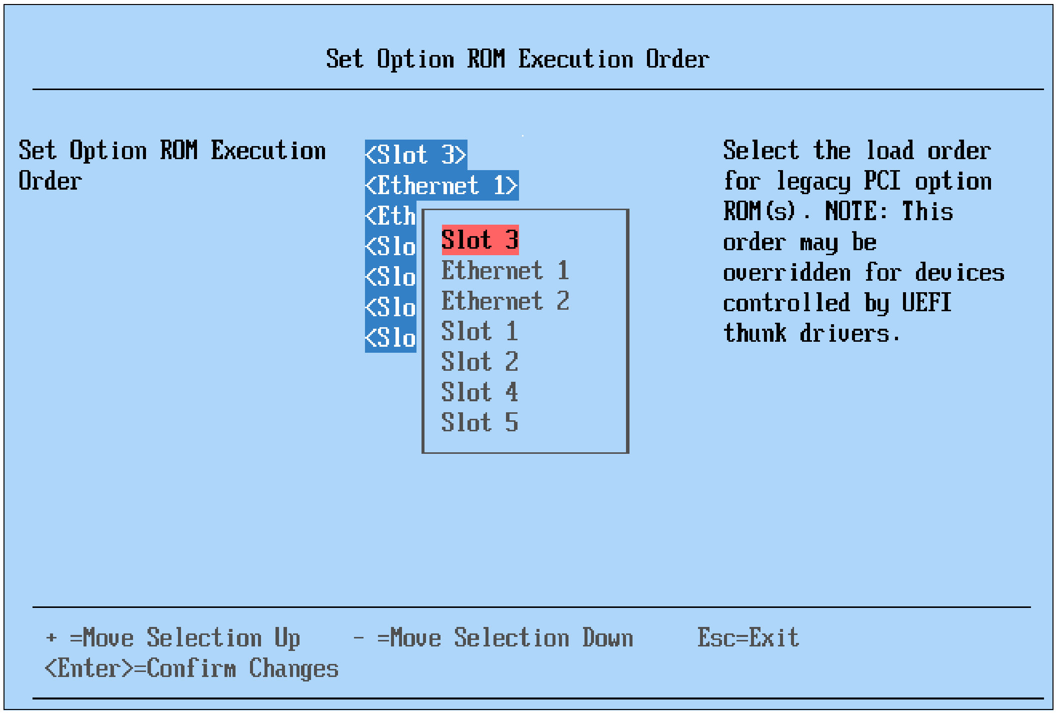


Figure 6-25 Set Option ROM Execution Order maintenance panel

6.7 Power supply considerations

Most models of the System x3850 X5 ship with two power supplies that will support the entire server with redundant power regardless of the configuration.

When the server loses one of the two power supplies, the server will report the following warning in the system event log:

“`Non-redundant: Sufficient Resources for Redundancy Degraded`”

The MAX5 also comes with two power supplies. If power fails in the MAX5, the server will power off. Powering the server back on will not be possible unless the MAX5 has power or until both QPI cables have been disconnected from the server, after removing ac power from the server. Be careful to ensure that both the MAX5 and the server to which it is attached are plugged into the same common power sources.

When two x3850 X5 servers are connected as a single complex, the loss of total ac power on either node will power both nodes down. The primary node will not power on unless ac power is supplied to the secondary node or the QPI cables are completely disconnected from the primary node.

You want to ensure that half of the power supplies of both units are plugged into one utility power source and that the remaining half are plugged into a separate utility power source.

This precaution will eliminate the possibility of a single breaker or circuit fault from taking down the entire server.

Think of the power supplies in your server as shock absorbers in your car. They are designed to absorb and overcome a wide variety of power conditions that can occur from an electric utility company, but like shock absorbers on a car, they will eventually begin to fail when fed a steady diet of unstable power. The time of their failure will most likely not coincide with a planned maintenance panel. For this reason, ensure that the two halves of the power supplies are plugged into two separate UPS sources to filter out all of the moderate to severe power fluctuations that occur.

6.8 Using the Integrated Management Module

For any successful server implementation, there must be provisions set aside to provide access to perform troubleshooting or routine maintenance. The x3850 X5 ships standard with the Integrated Management Module (IMM). The IMM is a separate, independent operating environment that starts to activate and remains active while the server is plugged into a good ac power source. The IMM monitors the hardware components of the server and the environment on which the server operates, looking for a potential hardware fault.

Part of the information that is stored in the IMM can be accessed with F1-Setup by selecting **System Settings** → **Integrated Management Module**. Figure 6-26 show the first panel of the IMM configuration panel.

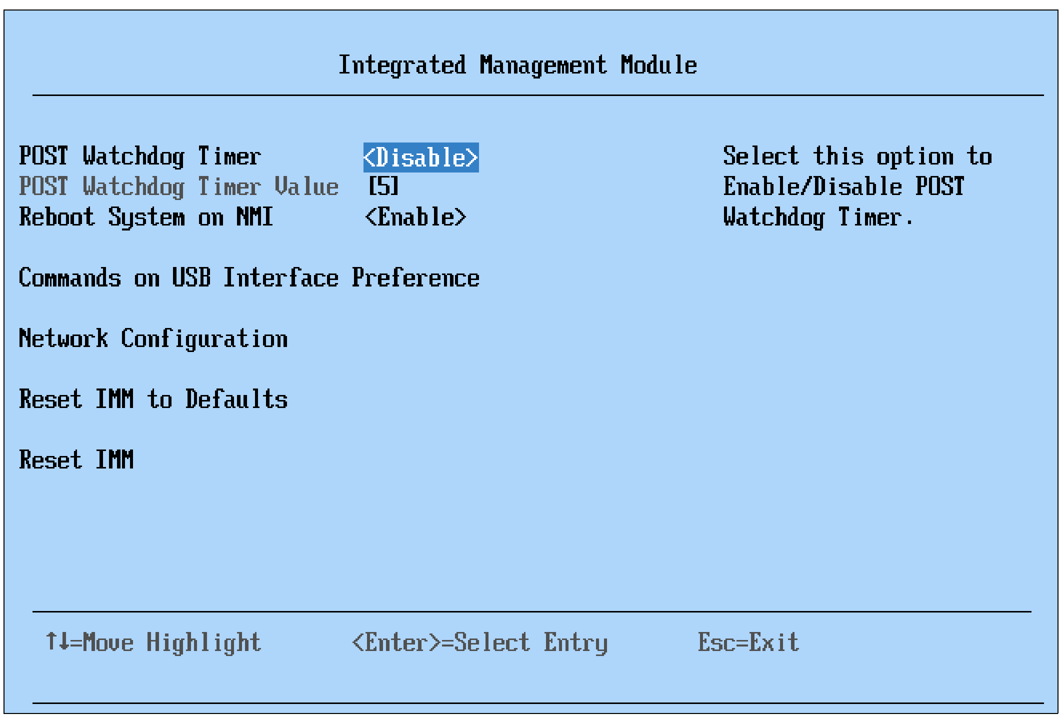


Figure 6-26 Integrated Management Module configuration panel

> **Tip:** If you have a server but you do not know the logon credentials, you can reset it by going to the panel that is shown in Figure 6-26, from F1-Setup, and then restore the IMM configuration to the factory defaults by selecting **Reset IMM to Defaults**.

6.8.1 IMM network access

The greatest strength of the IMM is the ability to completely monitor and manage the server from over the network. How much functionally you have through this remote access depends entirely on your configuration of the IMM.

IMM default configuration

The default network connection for the IMM on the x3850 X5 is through the System Management port on the back of the server. The following are the default settings of the IMM from the factory:

- Network IP: DHCP, if fails:
  - IP Address: `192.168.70.125`
  - Subnet Mask: `255:255:255:0`
  - Gateway: `0.0.0.0`
- Default user ID: `USERID`
- Default password: `PASSW0RD` where the `0` is a zero.

6.8.2 Configuring the IMM network interface

The IMM provides two paths to establish a network connection between you and the IMM by setting either Dedicated or Shared for the Network Interface Port in the Network Configuration panel of F1-Setup. In F1-Setup, you can access this panel by selecting **System Setting** → **Integrated Management Module** → **Network Configuration**, as shown in Figure 6-27 on page 251.

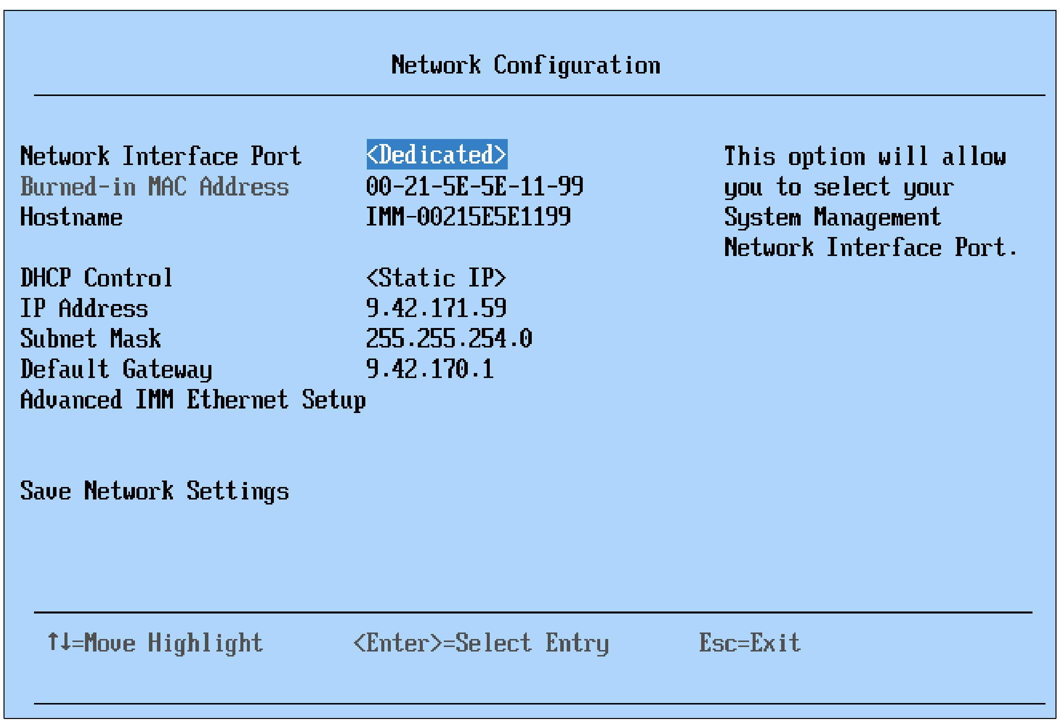


Figure 6-27 IMM Network Configuration panel

When configured as *Dedicated*, you are connecting to the network via the system management port. As shown in Figure 6-28, the port is located from the rear of the server on the left side of the video port. Using this port allows for easier separation of public and management network traffic. Separating the traffic is done when you connect your public network port to switch ports that belong to a public access virtual LAN (VLAN) and the management port is connected to a switch port defined by a separate management VLAN.

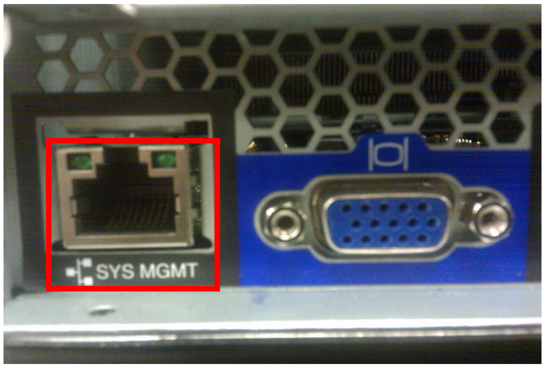


Figure 6-28 Dedicated 10/100 IMM system management port

When configured as *Shared*, you are sharing network traffic on the second onboard Ethernet port, the one closest to the power supply, as shown in Figure 6-29 on page 252. While this configuration eliminates a physical switch port and patch cable configuration, both the media access control (MAC) address for the second Ethernet port and the MAC address for the IMM will address through this single network port. This situation means at least two separate IP addresses for the same physical port, which prevents you from configuring the onboard Ethernet ports in a network team using 802.3ad load balancing.

Using this type of load balancing scheme will result in dropped packets for the IMM MAC address. Smart load balancing and failover are still available network teaming options. However, keeping the public traffic from the management traffic becomes more difficult.

To maintain separation between public and management traffic, network teaming software must be used to establish a VLAN to be used by the server to send public-tagged traffic to the network switch. The switch port must be configured as a trunk port to support both the public-tagged VLAN traffic, plus the untagged traffic for the management. The management VLAN must be defined as the native VLAN on the switch port, so that its untagged traffic from the switch will be accepted by the IMM MAC and dropped by the second Ethernet port's MAC.

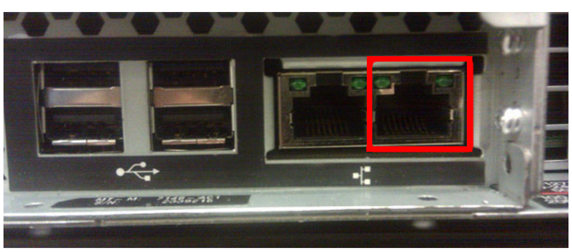


Figure 6-29 The onboard Ethernet port used when IMM Network interface is Shared

While the IMM uses a dedicated RISC processor, there are limitations as to the amount of network traffic that the IMM can be exposed to before complex functions, such as booting from a remote DVD, or USB storage will become unreliable because of timing issues. While the operating system has all of the necessary drivers in place to deal with these timing issues, the UEFI is not as tolerant. For this reason (maintaining secured access), the IMM must be kept on a separate management network.

6.8.3 IMM communications troubleshooting

The *Integrated Management Module User Guide* is an excellent guide to help you with every aspect of configuring and using the IMM. Download the guide from this website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5079770

Most communication errors are due to network switch configuration options, such as blocked ports or VLAN mismatches. The following procedure shows you how to determine this type of problem by connecting directly to the IMM port with a mobile computer and Ethernet patch cable and pinging and then starting a web session.

> **Crossover cable:** The management port is a 10/100 Ethernet port, so if your mobile computer does not have a 10/100/1000 Ethernet port on it, you will need to replace the patch cable with a 10/100 crossover cable. Only a 1Gb Ethernet port has the ability to auto-negotiate medium-dependent interface crossover (MDIX) when they auto-negotiate speed and duplex.

If you can ping the IMM, you have a good direct network link. If the web session fails, go through the following steps:

1. Try another web browser.
2. Directly access the Integrated Management Module configuration panel and reset the IMM in F1-Setup by selecting **System Settings** → **Integrated Management Module** → **Reset IMM.** You will have to wait about 5 minutes for the IMM to complete enough of its reboot to allow you to ping it. This IMM reset will have no impact on the operating system that is running on the server.
3. Try clearing the web browser cache.
4. Load the factory default settings back on the IMM through F1-Setup by selecting **System Settings** → **Integrated Management Module** → **Reset IMM to Defaults**. The IMM will need to be reset again after the defaults are loaded.
5. Contact IBM support.

6.8.4 IMM functions to help you perform problem determination

This section provides additional problem determination tips for the IMM. This section covers the following topics:

- "System Status"
- "Virtual light path diagnostics" on page 255
- "Hardware event log" on page 256
- "Remote control" on page 257

System Status

The first panel that you will see after completing the login and the session timeout limits panel is the System Status panel, as shown in Figure 6-30 on page 254. This panel provides a quick summary review of the hardware status of the server.

A green circle indicates that all is working from a strict hardware point of view. The IMM can check on the status of server components, the ServeRAID controllers, and PCIe interfaces to most PCIe adapters. It does not check on the functional status of most PCIe adapters with regard to their hardware connections to external devices. You will need to refer to the system event log from within the operating system or the switch logs of the network and fibre switches to which the server is connected to resolve connectivity issues.

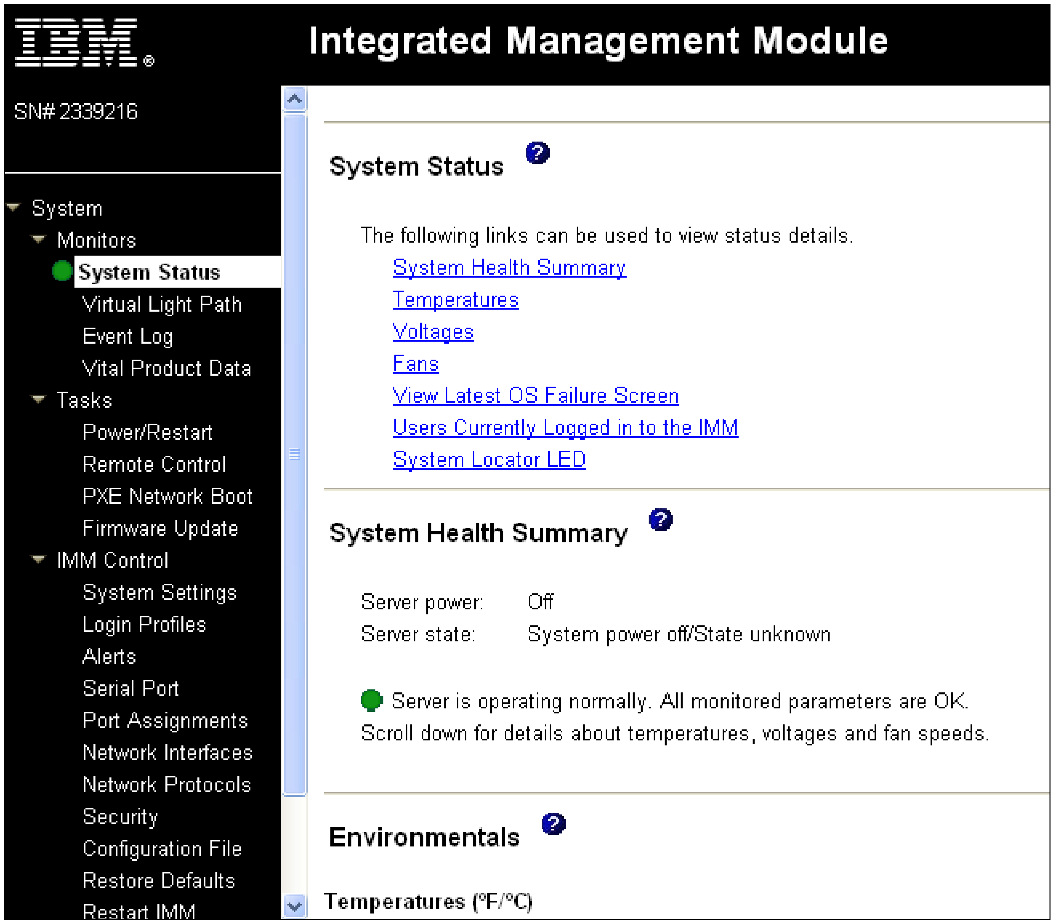


Figure 6-30 Integrated Management Module System Status

When an actual hardware error is detected in the server, the system status is represented by a red X. The System Health Summary will provide information about the errors presently unresolved in the server, as shown in Figure 6-31 on page 255.

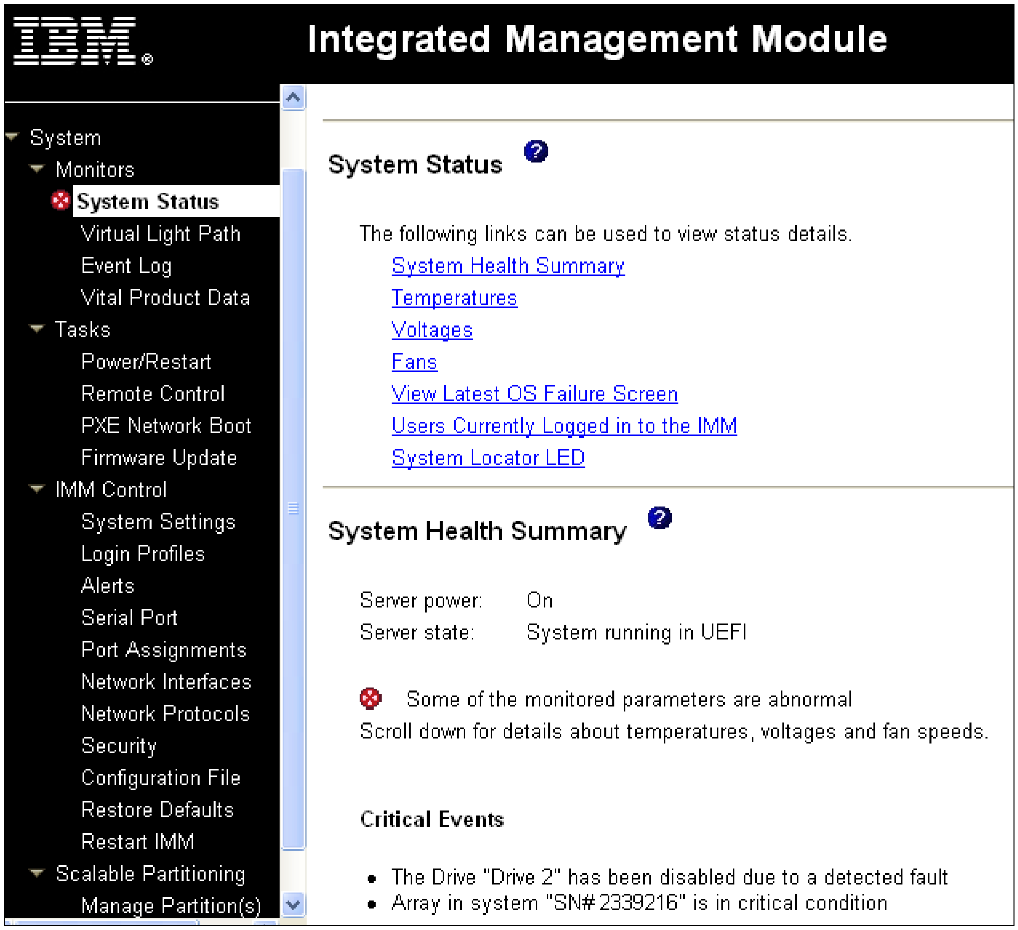


Figure 6-31 IMM System Status with a hard drive failure

Virtual light path diagnostics

If you are standing in front of the server, it is easy to track this problem by noticing the first tier of light path diagnostics, with the error light on the operator panel at the front of the server and on the rear of the server.

Pulling out the front operator panel reveals the second tier of light path diagnostics (as shown in Figure 6-32) that indicates the hardware subsystem that is experiencing the error.

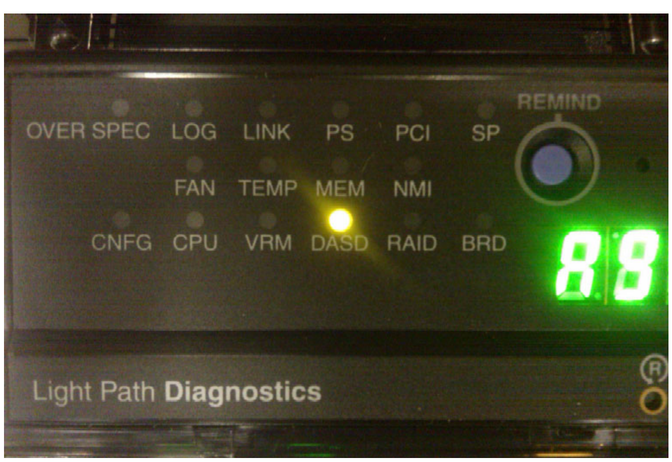


Figure 6-32 Tier 2 of light path diagnostics

Most servers are not physically located near the people who manage them. To help you see the event from a remote location, the IMM provides the capability of looking at all tiers of light path diagnostics, as shown in Figure 6-33.

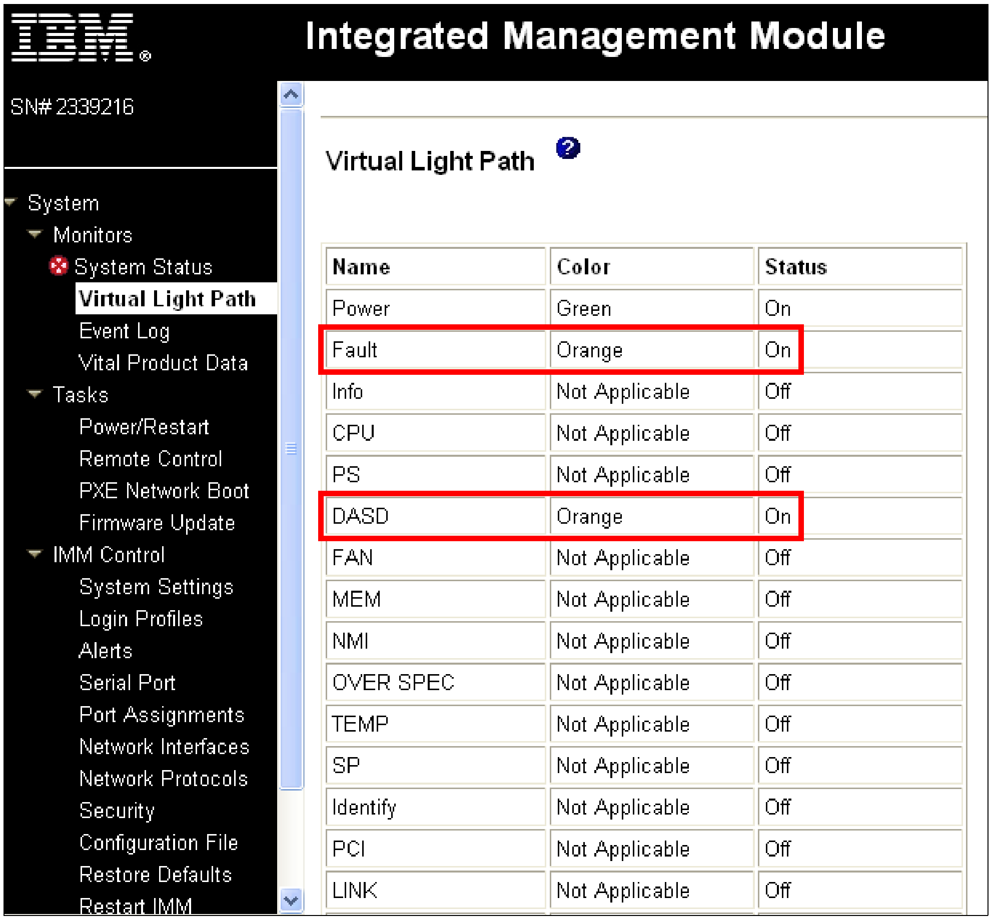


Figure 6-33 Integrated Management Module Virtual light path diagnostics

Hardware event log

For more detailed information, including the events that led up to a failure, you can access the hardware event log. Although not every event in the hardware event log is an event needing attention, the event log can provide insight to the cause or conditions that led up to the failure. The event log can be saved to a text file to be sent to IBM support.

Figure 6-34 on page 257 shows the IMM Event Log for the hard drive failure.

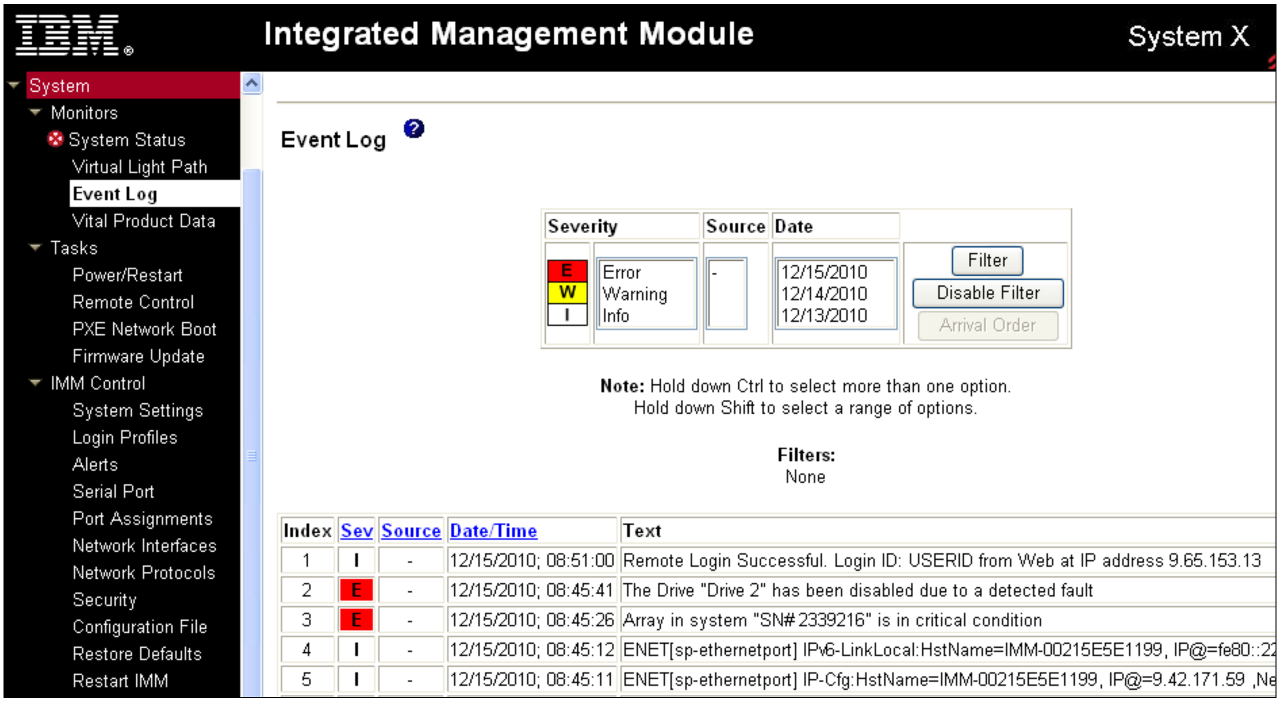


Figure 6-34 Integrated Management Module hardware Event Log

Remote control

Certain problems require that you get into the operating system or F1-Setup to detect them or fix them. For remotely managed servers, you can use the Remote Control feature of the Integrated Management Module. Figure 6-35 shows the options available for starting a remote control session.

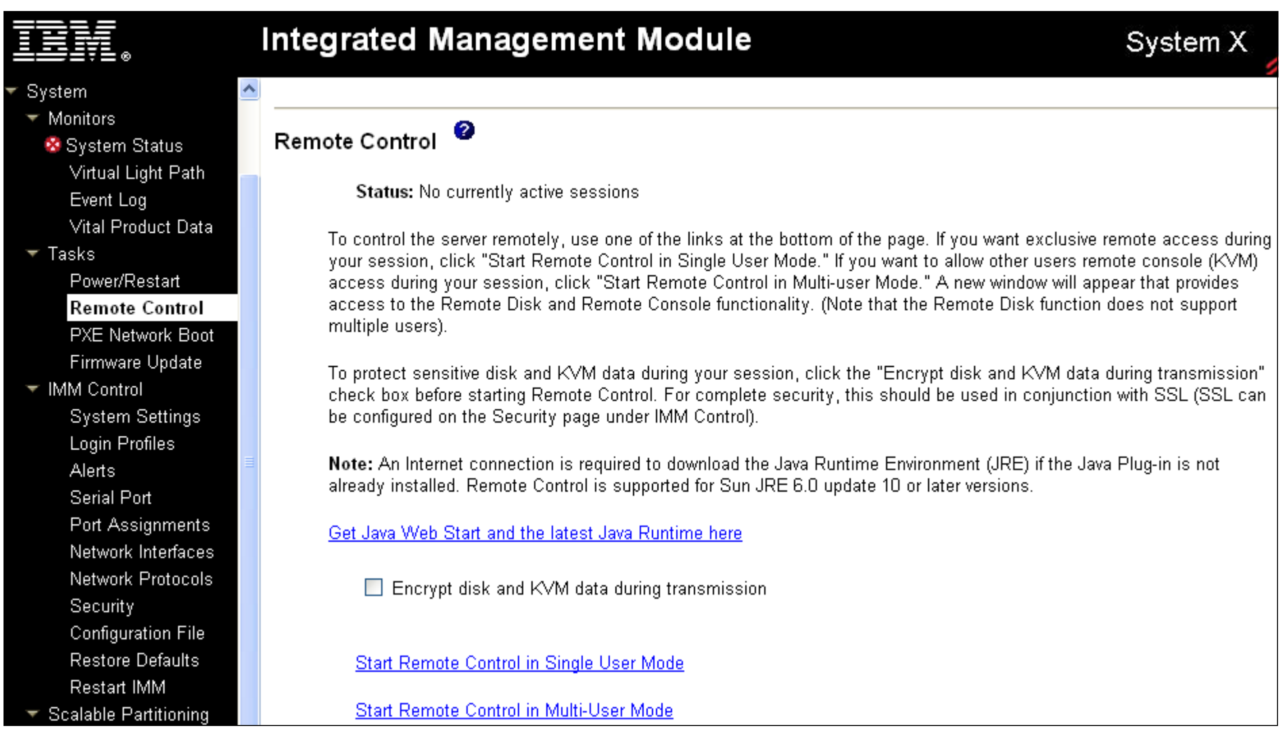


Figure 6-35 Integrated Management Module Remote Control session start-up panel

IMM Remote Control provides the following features:

- The remote control provides you with the same capability that you have with a keyboard, mouse, and video panel directly connected to the server.
- You have the ability to encrypt the session when used over public networks.
- You have the ability to use local storage or ISO files as mounted storage resources on the remote server you are using. These storage resources can be unmounted, changed, and remounted throughout the session, as needed.
- When combined with the Power/Restart functions of the IMM, you can power down, reboot, or power on the server while maintaining the same remote control session.

Depending on the application that you are accessing through the IMM Remote Control, you might notice that the mouse pointer is difficult to control. Fix this problem in the Video Viewer by selecting **Tools** → **Single Cursor**, as shown in Figure 6-36.

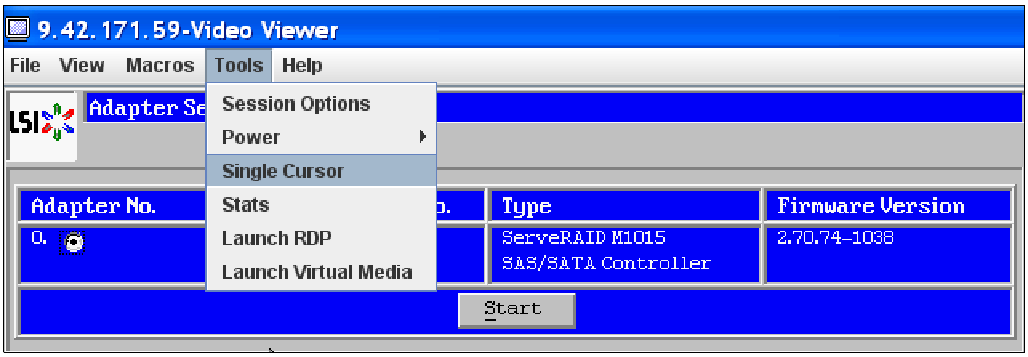


Figure 6-36 Fixing the mouse pointer in the Remote Control Video Viewer

6.9 UEFI settings

The Unified Extensible Firmware Interface (UEFI) is the interface between the operating system (OS) and platform firmware. UEFI provides a modern, well-defined environment for booting an OS and running pre-boot applications.

UEFI is effectively the replacement for BIOS. BIOS has been around for many years but was not designed to handle the amount of hardware that can be added to a server today. New IBM System x models and BladeCenter Blades implement UEFI to take advantage of its advanced features.

The UEFI page is accessed by pressing F1 during the system initializing process, as shown on Figure 6-37.

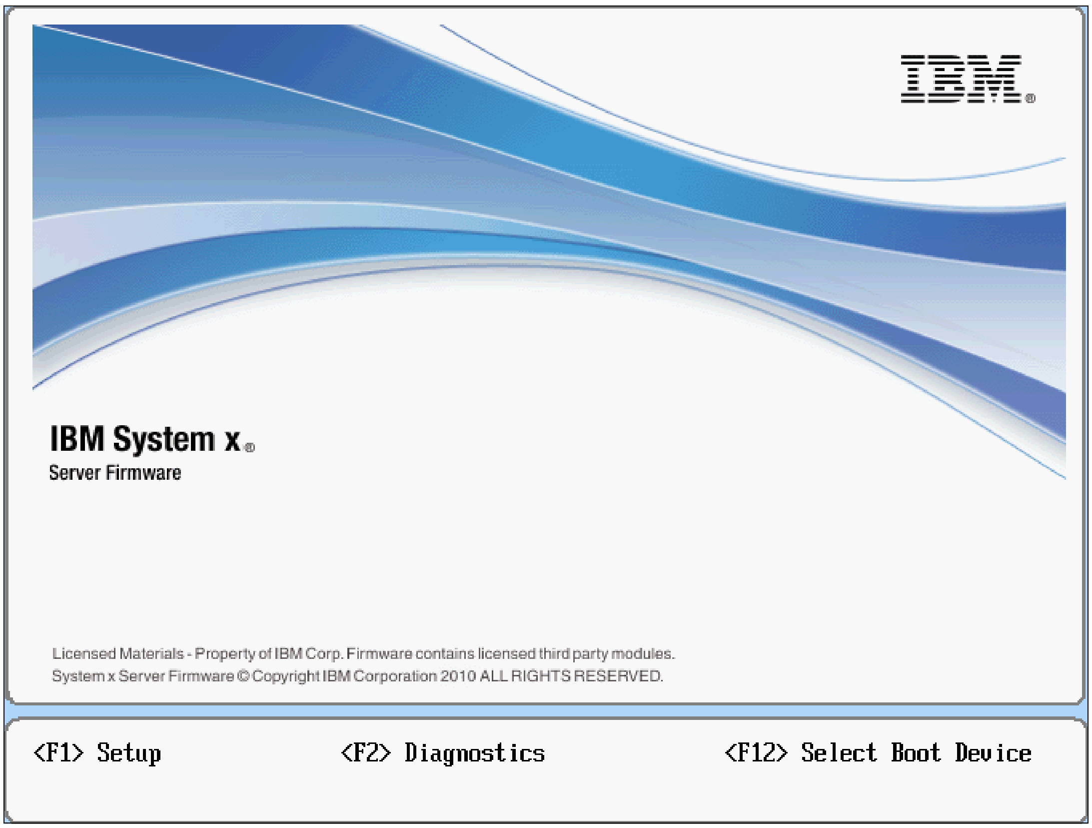


Figure 6-37 UEFI window on system start-up

If you use the factory defaults UEFI settings, the machine works in a either a 1-node, 2-node, or 1-node with MAX5 configuration. You can also change UEFI settings to meet your system requirements. In this section, we provide an overview of the UEFI settings for tuning your system for performance.

For an explanation of each setting, see 2.7, “UEFI system settings” on page 36.

You can use the Advanced Settings Utility (ASU) tool to change the UEFI settings values. ASU exposes more settings than the settings accessed using the F1-Setup panel. For more information about ASU, see 9.7, “Advanced Settings Utility (ASU)” on page 495.

Table 6-8 on page 260 provides an overview of the most common UEFI settings for optimizing system performance.

Table 6-8 UEFI settings, ASU values, and default settings

UEFI value	ASU value	ASU settings	Default
Processor settings			
TurboMode	uEFI.TurboModeEnable	Enable Disable	Enable
TurboBoost Power Optimization	uEFI.TurbBoost	Traditional PowerOptimized	PowerOptimized
Processor Performance States	uEFI.ProcessorEistEnable	Enable Disable	Enable
CPU C-States	uEFI.ProcessorCcxEnable	Enable Disable	Disable
C1 Enhanced Mode	uEFI.ProcessorC1eEnable	Enable Disable	Enable
Processor Data Prefetcher	uEFI.ProcessorDataPrefetch	Enable Disable	Enable
Hyper-Threading	uEFI.ProcessorHyperThreading	Enable Disable	Enable
Execute Disable Bit	uEFI.ExecuteDisableBit	Enable Disable	Enable
Intel Virtualization Technology	uEFI.ProcessorVmxEnable	Enable Disable	Enable
QPI Link Frequency	uEFI.QPISpeed	Max Performance Power Efficiency Minimal Power	Max Performance
Memory Settings			
N/A	IMM.ThermalModePolicy	Normal Performance	Normal
CKE Low Power	uEFI.CkeLowPolicy	Enable Disable	Disable
Memory Speed	uEFI.DdrSpeed	Max Performance Power Efficiency Minimal Power	Max Performance
Scheduler Policy (pagepolicy)	uEFI.SchedulerPolicy	Static Trade Off Static Read Primary Static Write Primary Adaptive	Adaptive
Mapper Policy	uEFI.MapperPolicy	Open Closed	Closed
Patrol Scrub	uEFI.PatrolScrub	Enable Disable	Disable
N/A	uEFI.DemandScrub	Enable Disable	Enable

6.9.1 Settings needed for 1-node, 2-node, and MAX5 configurations

Use these settings for each configuration:

- 1-node configuration

  If you use the factory default UEFI settings, the machine can work in a 1-node configuration. You can also change UEFI settings to meet your system requirements. See 2.7, “UEFI system settings” on page 36 for details.

- 2-node configuration

  If you use the factory default UEFI settings, the machine can work in a 2-node configuration. You can also change UEFI settings to meet your system requirements. See 2.7, “UEFI system settings” on page 36 for details.

- 1-node with MAX5

  Scaling an x3850 X5 system with a MAX5 makes a change in the UEFI settings. It adds the MAX5 Memory Scaling option in **System Settings** → **Memory**. This additional option is shown in Figure 6-38.

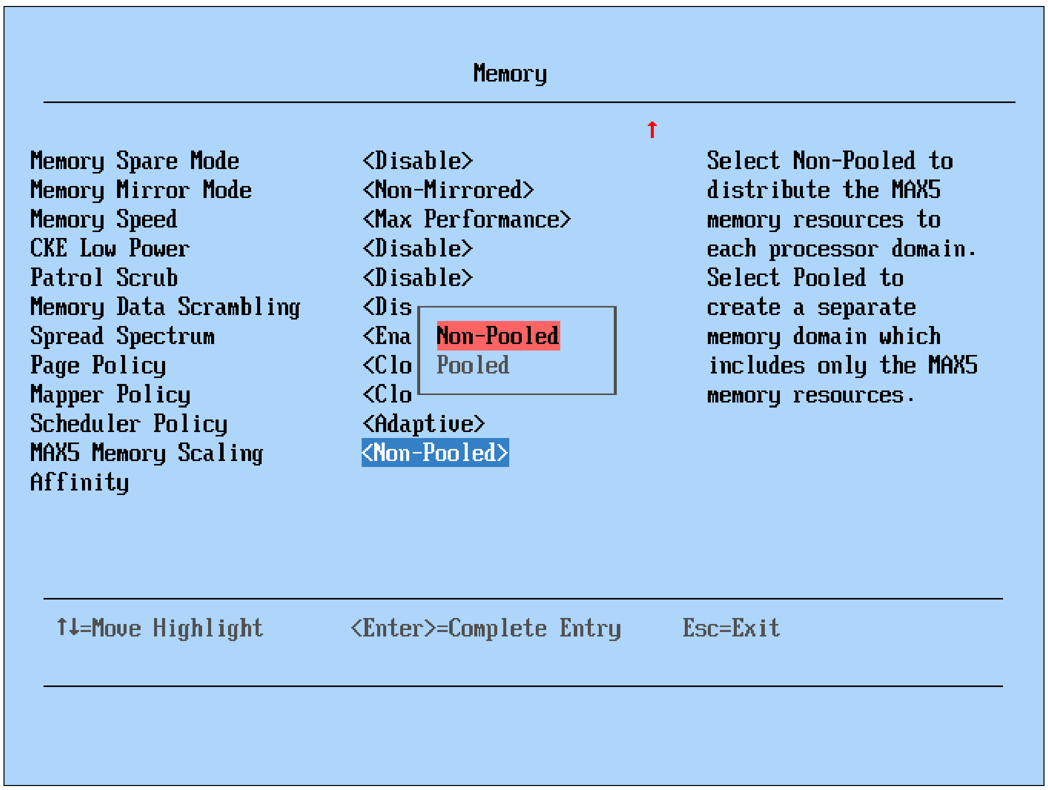


Figure 6-38 MAX5 Memory Scaling option in UEFI

The MAX5 Memory Scaling Affinity setting provides two options to determine how the system will present the memory capacity in the MAX5 unit to the OS:

- Non-Pooled

  The default option. Non-Pooled splits the memory in the MAX5 and assigns it to each of the installed processors. Configure VMware and Microsoft OSs to use the Non-Pooled setting.

- Pooled

  This option presents the additional memory in the MAX5 as a pool of memory without being assigned to any particular processor. Use this setting with Linux OSs.

6.9.2 UEFI performance tuning

Tuning the x3850 X5 for performance is a complicated topic, because it depends on which application you have installed or which workload this application generates. For example, a database server will generate a separate load on the hardware than a file and print server.

In this section, we provide general settings for the x3850 X5 that can be a good starting point for performance tuning. For more detailed information about the best settings for your specific environment and application needs, contact your IBM Business Partner or IBM representative.

Table 6-9 gives general recommendations for the most common UEFI settings.

Table 6-9 Overview of UEFI settings

Setting	MaxImum performance	Virtualization[b]	Low latency	Performance per watt	HPC
TurboMode[a]	Enabled	Enabled	Disabled	Disable	Disabled
TurboBoost	Traditional	Power Optimized	Automatically disabled	Automatically disabled	Automatically disabled
Processor Performance states	Enabled	Enabled	Disabled	Enabled	Disabled
C states	Disabled	Enabled	Disabled	Enabled	Enabled
C1E state	Disabled	Enabled	Disabled	Enabled	Enabled
Prefetcher	Enabled	Enabled	Enabled	Enabled	Enabled
Hyper-Threading	Enabled	Enabled	Disabled	Enabled	Enabled
Execute Disable	Disabled	Enabled	Disabled	Enabled	Disabled
Virtualization Extensions	Disabled	Enabled	Disabled	Enabled	Disabled
QPI Link Speed	Max Performance	Max Performance	Max Performance	Power Efficiency	Max Performance
IMM Thermal Mode	Performance	Performance	Performance	Normal	Performance
CKE Policy	Disabled	Disabled	Disabled	Disabled	Disabled
DDR Speed	Max Performance	Max Performance	Max Performance	Max Performance	Max Performance
Page Policy	Closed	Closed	Closed	Closed	Closed
Mapper Policy	Closed	Closed	Closed	Closed	Closed
Patrol Scrub	Disabled	Disabled	Disabled	Disabled	Disabled

Setting	MaxImum performance	Virtualization[b]	Low latency	Performance per watt	HPC
Demand Scrub	Enabled	Enabled	Disabled	Enabled	Disabled

[a] Depending on the processor workload, enabling TurboMode might also increase power consumption. The actual processing performance boost that you get from TurboMode depends on the environment that the server is in terms of temperature and humidity, because the processor will only boost performance up to the environmental limits set by the processor.
[b] These Virtualization settings are recommended for a stand-alone host only. For multiple virtualized hosts, in clustered workloads, use the Maximum performance settings instead.

6.10 Installing an OS

This section provides an overview of the options that you have when installing an OS on the x3850 X5. We provide instructions for VMware ESX/ESXi installation when a MAX5 is attached to the x3850 X5 and for installing an OS with a USB key are provided, as well as additional installation topics.

We recommend that you use ServerGuide for your OS install procedure. For more information about ServerGuide, see 9.8, “IBM ServerGuide” on page 501.

Topics in this section:

- 6.10.1, “Installing without a local optical drive” on page 263
- 6.10.2, “Use of embedded VMware ESXi” on page 271
- 6.10.3, “Installing the ESX 4.1 or ESXi 4.1 Installable onto x3850 X5” on page 275
- 6.10.4, “OS installation tips and instructions on the web” on page 288
- 6.10.5, “Downloads and fixes for x3850 X5 and MAX5” on page 293
- 6.10.6, “SAN storage reference and considerations” on page 294

6.10.1 Installing without a local optical drive

If you do not have a local optical drive, you can install an OS using any of the following methods:

- “IMM”
- “Local USB port” on page 265
- “ServerGuide Scripting Toolkit” on page 266
- “Preboot eXecution Environment (PXE)” on page 271
- “Tivoli Provisioning Manager for OS Deployment” on page 271

The following sections provide details for each method.

IMM

A remote control feature is available through the IMM web interface. You must log in to the IMM with a user ID that has Supervisor access. You can also assign to the server a CD or DVD drive, diskette drive, USB flash drive, or disk image that is on your computer.

The following server OSs have USB support, which is required for the Remote Disk feature:

- Microsoft Windows Server 2008
- Microsoft Windows Server 2003
- Red Hat Linux versions 4.0 and 5.0
- SUSE Linux version 10.0

- Novell NetWare 6.5

For more information, see the *Integrated Management Module User's Guide* at the following web page:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5079770

Follow these steps to mount a drive through IMM:

1. Connect to the IMM with your web browser.
2. Click **Task** → **Remote Control**.
3. If you want to allow other users remote control access during your session, click **Start Remote Control in Multi-user Mode**. Otherwise, click **Start Remote Control in Single User Mode**.
4. Two Java application windows open, as shown in Figure 6-39 and Figure 6-40.

Figure 6-39 Video Viewer window

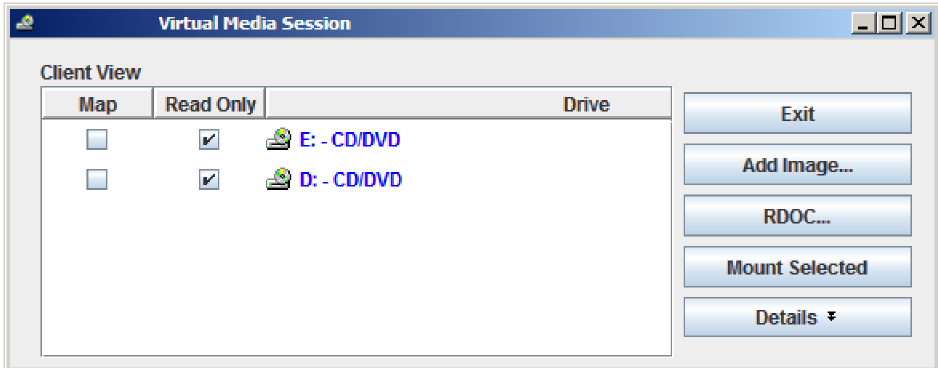


Figure 6-40 Virtual Media Session window

5. Select the **Virtual Media Session window.**
6. Click **Add Image** if you want to map an IMG or ISO image file.
7. Select the check box next to the drive that you want to map and click **Mount Selected**, as shown in Figure 6-41.

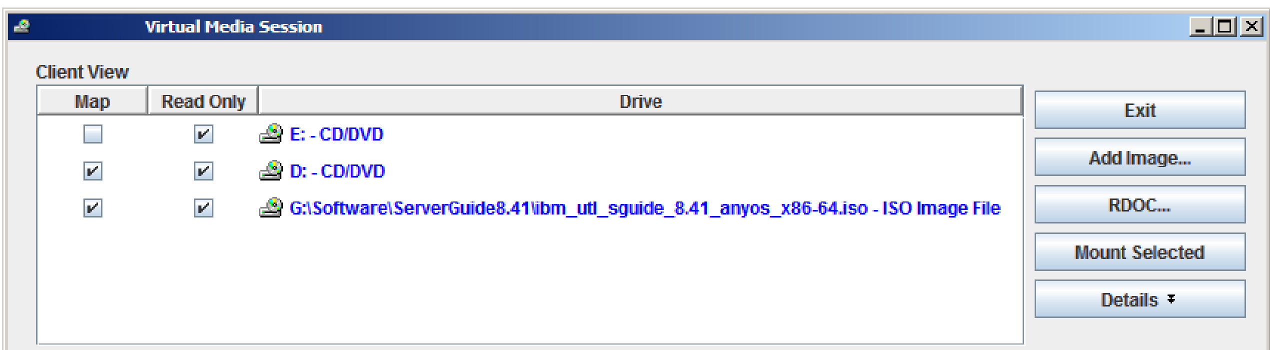


Figure 6-41 Overview of the selected drive

8. The image drive is now accessible by the system.

Closing the session: Closing the Virtual Media Session window when a remote disk is mapped to the machine causes the machine to lose access to the remote disk.

Internet Explorer: If you use Internet Explorer 7 or 8 and the remote control window does not open, see RETAIN tip H196657 for steps to solve the problem:

`http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5083262`

Local USB port

You can use the local USB port to attach a USB flash drive that contains the OS installation files. There are several methods to create a bootable flash drive.

For VMware, you can use the embedded hypervisor key, which is preinstalled with ESXi. You do not need to install VMware. For more information about the embedded hypervisor key, see 2.9.1, “VMware ESXi” on page 50.

For Linux, look on the vendor websites. They contain information about installation with a USB flash drive. For example, the following web pages provide details for using a USB key as an installation medium:

- Installing Red Hat Linux from a USB flash drive:

  `http://ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101131`

- How to create a bootable USB drive to install SLES:

  `http://www.novell.com/support/php/search.do?cmd=displayKC&docType=kc&externalId=3499891`

You can also use the ServerGuide Scripting Toolkit to create a bootable USB flash drive, as explained in the next section.

ServerGuide Scripting Toolkit

As described in 9.9, “IBM ServerGuide Scripting Toolkit” on page 507, you can use the ServerGuide Scripting Toolkit to customize your OS deployment. You can use the ServerGuide Scripting Toolkit for Windows, Linux, and VMware. This section contains information about deployment to allow you to begin using the Toolkit as quickly as possible.

For more information, see the *IBM ServerGuide Scripting Toolkit, Windows Edition User’s Reference* and *IBM ServerGuide Scripting Toolkit, Linux Edition User” Reference* at the following web page:

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

Windows installation

This section describes the process to install the ServerGuide Scripting Toolkit to create a deployment image for Windows 2008 R2 Enterprise Edition and then to copy this image to a USB key for deployment.

To configure a USB key for deployment, you need the following requirements:

- A system running Windows Vista, Windows Server 2008, Windows 7, Windows Server 2008 R2, Windows 2.1 Preinstallation Environment (PE), or a Windows 3.0 PE session
- A USB key with a storage capacity at least 64 MB larger than your Windows PE image, but not less than 4 GB.

We follow this procedure:

1. Install the ServerGuide Scripting Toolkit.
2. Create a deployment image.
3. Prepare USB key.

Installing the ServerGuide Scripting Toolkit

You must install the English language version of the Windows Automated Installation Kit (AIK) for Windows 7 family, Windows Server 2008 family, and Windows Server 2008 R2 family, which is available at the following website:

`http://www.microsoft.com/downloads/en/details.aspx?familyid=696DD665-9F76-4177-A811-39C26D3B3B34&displaylang=en`

Follow these steps to install the ServerGuide Scripting Toolkit, Windows Edition:

1. Download the latest version from the following website:

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

2. Create a directory, for example, `C:\sgshare`.

3. Decompress the *ibm_utl_sgtkwin_X.XX*`_windows_32-64.zip` file to the directory that you have created, for example, `C:\sgshare\sgdeploy`.

Creating a deployment image

Follow these steps to create a Windows installation image.

1. Start the Toolkit Configuration Utility in the `C:\sgshare\sgdeploy` directory.
2. Select **Add Operating System Installation Files**, as shown in Figure 6-42 on page 267.

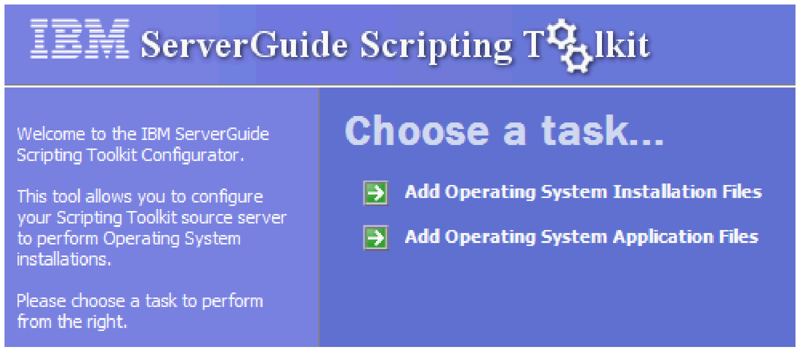


Figure 6-42 Main window

3. Choose the OS type that you want and click **Next**, as shown in Figure 6-43.

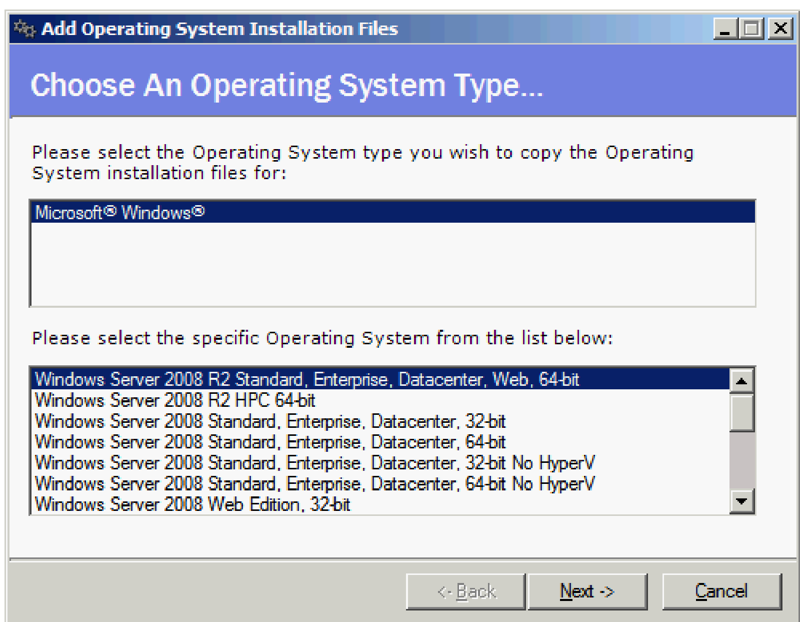


Figure 6-43 Select the type of operating system

4. Insert the correct OS installation media or select the folder that contains the installation files for the source, as shown in Figure 6-44 on page 268. If necessary, modify the target and click **Next**.

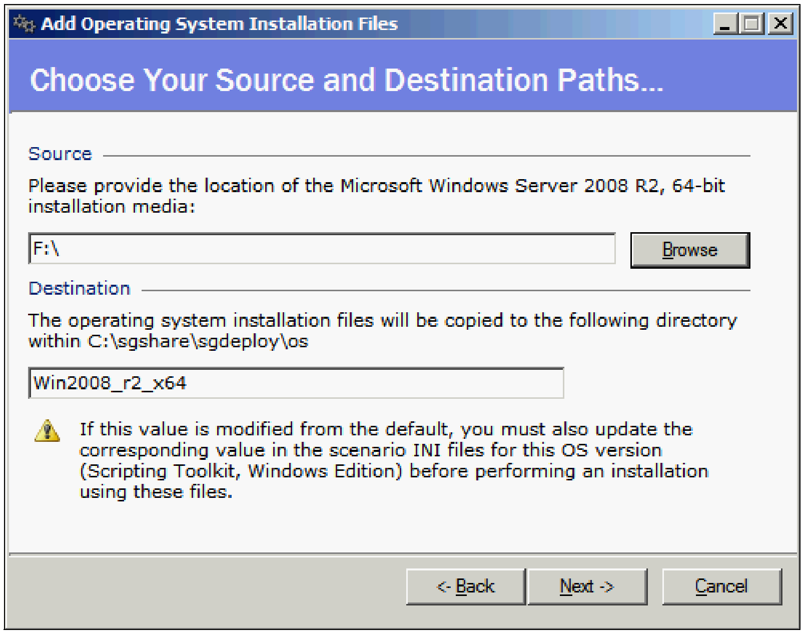


Figure 6-44 Define the source and target

5. When the copy process is finished, as shown in Figure 6-45, click **Finish**.

Figure 6-45 The copy process successfully completes

6. Open a command prompt and change to the `C:\sgshare\sgdeploy\SGTKWinPE` directory. Use the following command to create the Windows installation image:

```
SGTKWinPE.cmd ScenarioINIs\Local\Win2008_R2_x64_EE.ini
```

7. When the process is finished, as shown in Figure 6-46 on page 269, your media creation software is started to create bootable media from the image. Cancel this task.

```
18:26:21 - Creating the WinPE x64 ISO...

18:27:07 - The WinPE x64 ISO was created successfully.

*** WinPE x64 ISO:  c:\sgshare\sgdeploy\WinPE_ScenarioOutput\Local_Win2008_R2_x6
4_EE\WinPE_x64.iso

18:27:07 - Launching the registered software associated with ISO files...

*** Using ISO File:  c:\sgshare\sgdeploy\WinPE_ScenarioOutput\Local_Win2008_R2_x
64_EE\WinPE_x64.iso

18:27:08 - The WinPE x64 build process finished successfully.

SGTKWinPE complete.

c:\sgshare\sgdeploy\SGTKWinPE>
```

Figure 6-46   Build process is finished

Preparing the USB key

Follow these steps to create a bootable USB key with the Windows installation image that was created in “Creating a deployment image” on page 267:

1. Insert your USB key.

2. Enter **diskpart** to format the USB key using FAT32. All files on the USB key will be deleted. At the command prompt, type the commands that are listed in Figure 6-47.

```
C:\>diskpart
Microsoft DiskPart version 6.1.7600
Copyright (C) 1999-2008 Microsoft Corporation.
On computer:

DISKPART> list disk

  Disk ###  Status         Size     Free     Dyn  Gpt
  --------  -------------  -------  -------  ---  ---
  Disk 0    Online          271 GB      0 B        *
  Disk 1    Online          135 GB      0 B
  Disk 2    Online         7839 MB      0 B

DISKPART> select disk 2
Disk 2 is now the selected disk.

DISKPART> clean
DiskPart succeeded in cleaning the disk.

DISKPART> create partition primary
DiskPart succeeded in creating the specified partition.

DISKPART> select partition 1
Partition 1 is now the selected partition.

DISKPART> active
DiskPart marked the current partition as active.

DISKPART> format fs=fat32
  100 percent completed
DiskPart successfully formatted the volume.

DISKPART> assign
DiskPart successfully assigned the drive letter or mount point.

DISKPART> exit
```

Figure 6-47 Using diskpart to format the USB memory key

3. Copy the contents from `C:\sgshare\sgdeploy\WinPE_ScenarioOutput\Local_Win2008_R2_x64_EE\ISO` to the USB key. The USB key includes the folders and files that are shown in Figure 6-48.

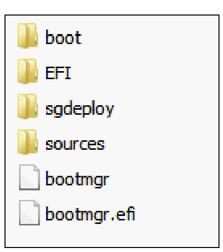


Figure 6-48 Contents of the USB key

4. Boot the target system from the USB key. The deployment executes automatically.

RAID controller: If the target system contains a RAID controller, RAID is configured as part of the installation.

Linux and VMware installation

The procedure for Linux and VMware is similar to the Windows procedure:

1. Install the ServerGuide Scripting Toolkit.
2. Create a deployment image.
3. Prepare a USB key.

For more information, see the *IBM ServerGuide Scripting Toolkit, Linux Edition User's Reference* at the following website:

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

Preboot eXecution Environment (PXE)

The Preboot eXecution Environment (PXE) is an environment to boot computers using a network interface for operating system deployment. All eX5 systems support PXE.

For example, you can use the ServerGuide Scripting Toolkit. For more information, see the *IBM ServerGuide Scripting Toolkit User's Reference* at the following web page:

http://www.ibm.com/support/docview.wss?uid=psg1SERV-TOOLKIT

Tivoli Provisioning Manager for OS Deployment

IBM Software has an offering for users needing advanced features for automating and managing the remote deployment of OSs and virtual images, in the form of Tivoli® Provisioning Manager for OS Deployment. It is available in a stand-alone package and as an extension to IBM Systems Director.

You can obtain more information about these offerings at the following web pages:

- Tivoli Provisioning Manager for OS Deployment:

  http://ibm.com/software/tivoli/products/prov-mgr-os-deploy/

- Tivoli Provisioning Manager for OS Deployment IBM Systems Director Edition:

  http://ibm.com/software/tivoli/products/prov-mgr-osd-isd/

6.10.2 Use of embedded VMware ESXi

ESXi is an embedded version of VMware ESX. The footprint of ESXi is small (approximately 32 MB) because it does not use the Linux-based Service Console. Instead, it uses management tools, such VirtualCenter, the Remote Command-Line interface, and Common Information Model (CIM) for standards-based and agentless hardware monitoring. VMware ESXi includes full VMware File System (VMFS) support across Fibre Channel and iSCSI SAN, and network attached storage (NAS). It supports 4-way Virtual SMP (VSMP). ESXi 4.0 supports 64 CPU threads (for example, eight x 8-core CPUs) and can address 1 TB of RAM.

The VMware ESXi 4.0 and 4.1 embedded virtualization keys are orderable. See 2.9.1, "VMware ESXi" on page 50 for the part numbers.

Setting the boot order

To ensure that you can boot ESXi successfully, you must change the boot order. The first boot entry must be Legacy Only and the second boot entry must be Embedded Hypervisor. Follow these steps:

1. Press F1 for the UEFI Setup.
2. Select **Boot Manager** → **Add Boot Option**.
3. Select **Legacy Only** and **Embedded Hypervisor**, as shown in Figure 6-49. If either option is not listed, the option is already in the boot list. When you have finished, press Esc to go one panel back.

Figure 6-49 Add boot options

4. Select **Change Boot Order.**
5. Change the boot order to Legacy Only followed by Embedded Hypervisor, as shown in Figure 6-50.

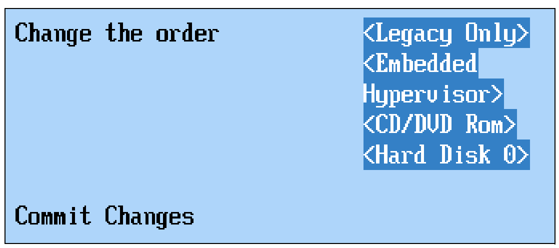


Figure 6-50 Example of a boot order

6. Select **Commit Changes** and press Enter to save the changes.

Installing system memory in a balanced configuration

When installing the ESXi Server OS on the x3850 X5, the memory must be balanced across all processors in the system. This rule applies to 1-node, 2-node, and x3850 X5 with MAX5 configurations. Failure to follow this rule will prevent the OS from installing correctly.

See 2.3.4, “Nonuniform memory architecture (NUMA)” on page 26 for more information about NUMA.

Configuring UEFI for embedded ESXi 4.1 if MAX5 is attached

Systems running VMware ESXi Server must use Non-Pooled mode in the MAX5 Memory Scaling option with the UEFI.

See 6.9.1, “Settings needed for 1-node, 2-node, and MAX5 configurations” on page 261 for instructions to configure the MAX5 Memory Scaling option.

Booting a new embedded ESXi 4.1 with MAX5 attached

To successfully boot the ESXi4.1 on an x3850 X5 with MAX5, the follow these instructions. For instructions to scale the x3850 X5 with MAX5, see 6.4, “Attaching the MAX5 memory expansion unit” on page 230.

1. Boot the host.

2. In the Loading VMware Hypervisor panel, press Shift+O when the progress bar is displayed.
3. Enter the following command at the prompt:

```
esxcfg-advcfg -k TRUE allowInterleavedNUMAnodes
```

4. After the system boots, connect to the system using the vSphere Client.
5. Select the **Configuration** tab of the host and click **Advanced Settings** under Software, as shown in Figure 6-51.

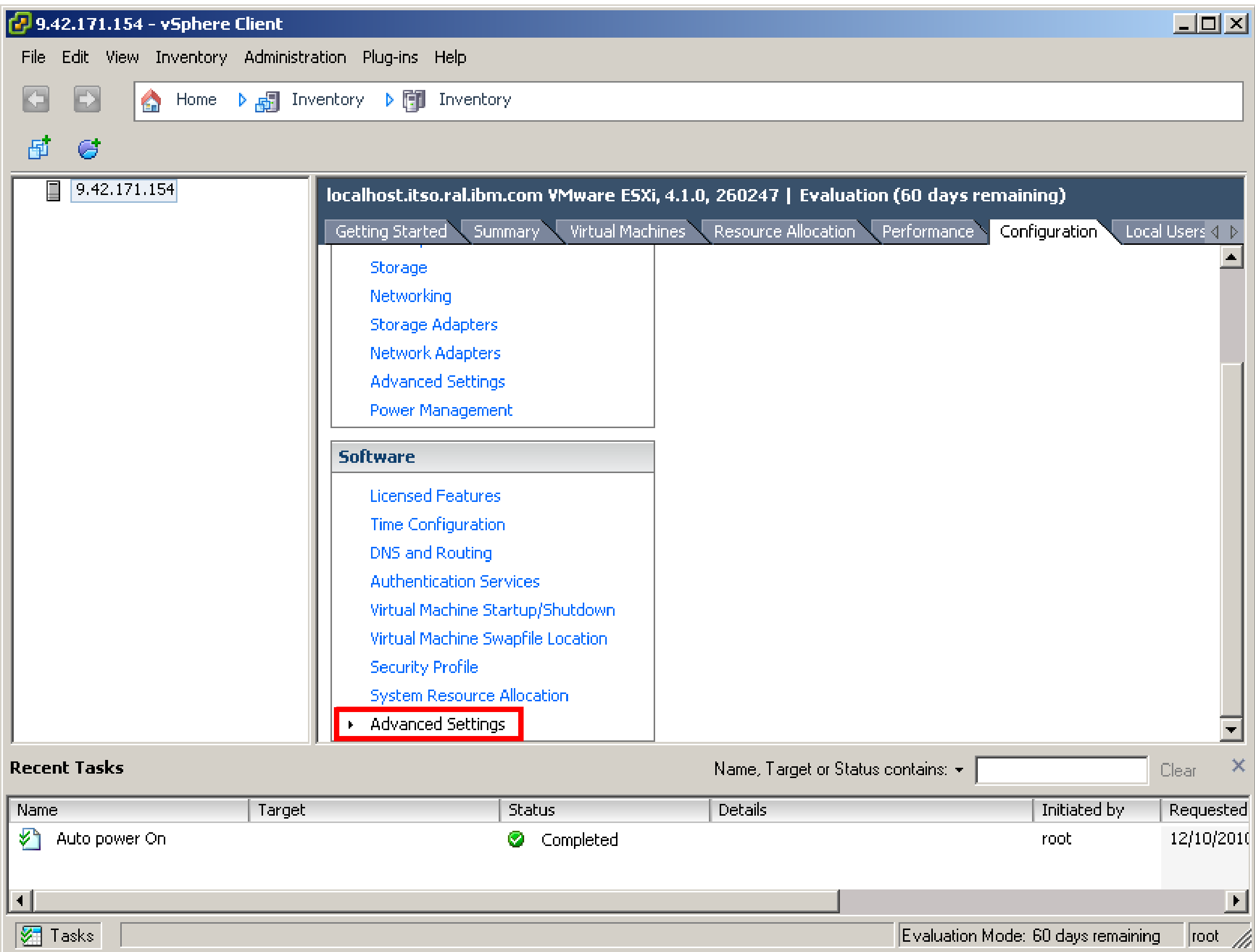


Figure 6-51 Advanced Settings in vSphere Client

6. Click **VMkernel** and select the check box next to **VMkernel Boot.allowInterleavedNUMAnodes**, as shown in Figure 6-52 on page 274.

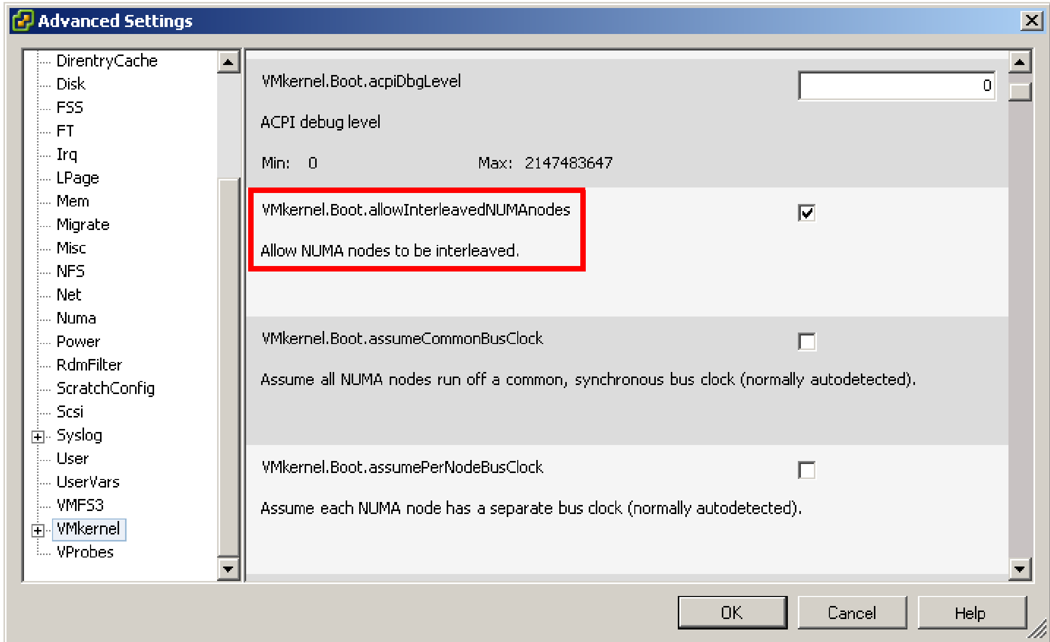


Figure 6-52 Editing the WMkernel settings in vSphere Client Advanced Settings

7. Click **OK** to save the settings.

Updating ESXi

You can install the latest version of ESXi 4 on IBM Hypervisor keys and it is supported by IBM. Use the following VMware upgrade mechanisms for the update:

- VMware Upgrade Manager
- Host Update Utility

For more information, see the VMware Documentation website:

http://www.vmware.com/support/pubs

ESXi recovery

You can use the IBM recovery CD to recover the IBM USB Memory Key to a factory-installed state. Table 6-10 provides the part numbers and versions for the CDs.

Table 6-10 VMware ESXi recovery CD

Part number	Description
68Y9634	VMware ESXi 4.0 U1
49Y8747	VMware ESXi 4
68Y9633	VMware ESX Server 3i v 3.5 Update 5
46M9238	VMware ESX Server 3i v 3.5 Update 4
46M9237	VMware ESX Server 3i v 3.5 Update 3
46M9236	VMware ESX Server 3i v 3.5 Update 2
46D0762	VMware ESX Server 3i version 3.5

To order a recovery CD, contact your local support center at the following website:

http://www.ibm.com/planetwide/region.html

6.10.3 Installing the ESX 4.1 or ESXi 4.1 Installable onto x3850 X5

Before installing any VMware OS, always refer to the latest OS support information that is contained on the IBM ServerProven site. You can obtain IBM ServerProven information at the following website:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/vmwaree.html

VMware supported versions based on x3850 X5 hardware configuration

IBM ServerProven provides general OS support information for the x3850 X5. Table 6-11 lists the versions of VMware ESX that are supported on the various x3850 X5 hardware configurations.

Table 6-11 WMware OS versions supported based on x3850 X5 hardware configuration

VMware OS	One-node	Two-node	x3850 X5 with MAX5
VMware ESX Server 4.0 Update 1	Yes	No	No
VMware ESXi Server 4.0 Update 1	Yes	No	No
VMware ESX Server 4.1	Yes	Yes	Yes
VMware ESXi Server 4.1	Yes	Yes	Yes

Tip: If your System has 1 Terabyte (TB) or greater of memory, see the following web page:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084935

Nonuniform memory access (NUMA)

When installing any supported version of the ESX Server OS onto the x3850 X5, the memory must be balanced across all processors in the system. This rule applies to 1-node, 2-node configurations, and x3850 X5 with MAX5 attached. Failure to follow this rule will prevent the OS from installing correctly. See 2.3.4, “Nonuniform memory architecture (NUMA)” on page 26 for details.

Installing the ESX/ESXi 4.1 Installable edition onto x3850 X5 with MAX5

To correctly configure and install ESX/ESXi 4.1 Installable editions of ESX/ESXi Server, follow these instructions:

1. Connect the MAX5, as described in 6.4, “Attaching the MAX5 memory expansion unit” on page 230.
2. Configure memory scaling. Systems running VMware ESX/ESXi Server must use Non-Pooled mode in the MAX5 Memory Scaling option. See “1-node with MAX5” on page 261 for instructions to configure the MAX5 Memory Scaling option.
3. Configure your RAID arrays.
4. Use the following procedures to install ESX/ESXi 4.1 Installable on an x3850 X5 with MAX5:
   - “Installing the ESX 4.1 OS” on page 276
   - “Installing the ESXi 4.1 OS” on page 278

Installing the ESX 4.1 OS

VMware ESX 4.1 is not UEFI aware, and therefore, you must use the Legacy Only option as the first boot option. To decrease the boot time, the optimal minimum configuration is Legacy Only, CD/DVD Rom, and Hard Disk 0.

Use the following steps to install ESX 4.1 Server onto x3850 X5 with MAX5 attached:

1. Power on the system and press F1 when the UEFI splash panel is shown.
2. Select **Boot Manager** → **Add Boot Option**.
3. Select **CD/DVD Rom**, **Legacy Only**, and **Hard Disk 0**. If either option is not listed, the option is already included in the boot list.
4. Press Esc when finished to go back one panel.
5. Select **Change Boot Order** and change the boot order to look like Figure 6-53.

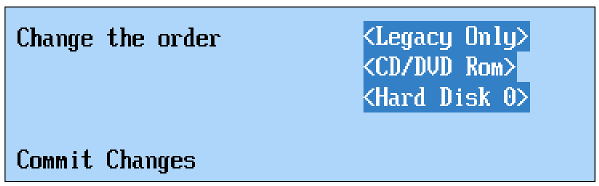


Figure 6-53 Example of a boot order

6. Select **Commit Changes** and press Enter to save the changes.
7. Boot the host from the ESX installation media.
8. Press F2 when you see the ESX 4.1 installation options panel, as shown in Figure 6-54.

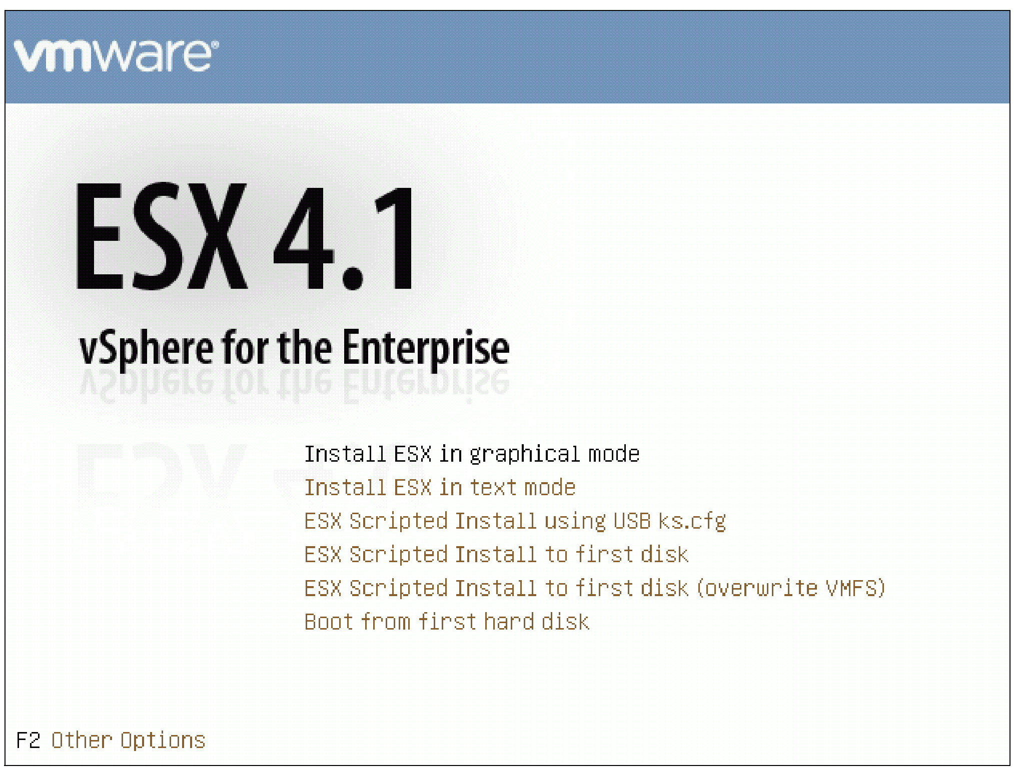


Figure 6-54 ESX installation options panel

9. The Boot Options line will appear on the panel. Type the following parameter at the end of the Boot Options line:

   `allowInterleavedNUMAnodes=TRUE`

   The edited result looks like Figure 6-55. Press Enter to proceed.

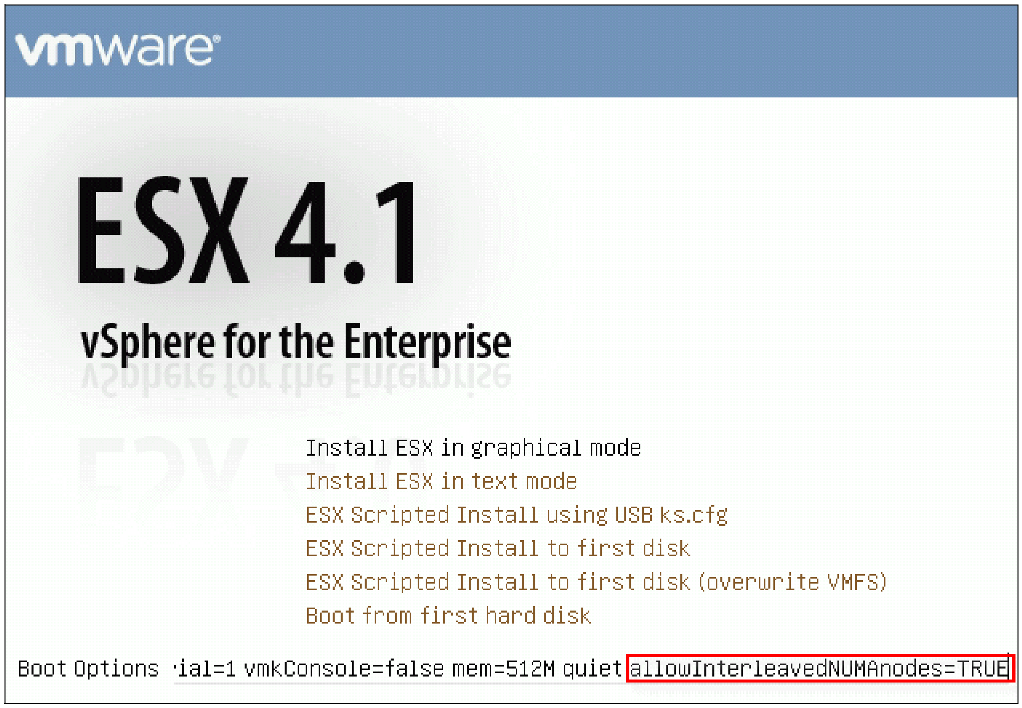


Figure 6-55 Editing the Boot Options

10. Proceed through the installer until you reach the Setup Type page. Click **Advanced setup** and clear **Configure bootloader automatically (leave checked if unsure)** as shown in Figure 6-56. Click **Next**.

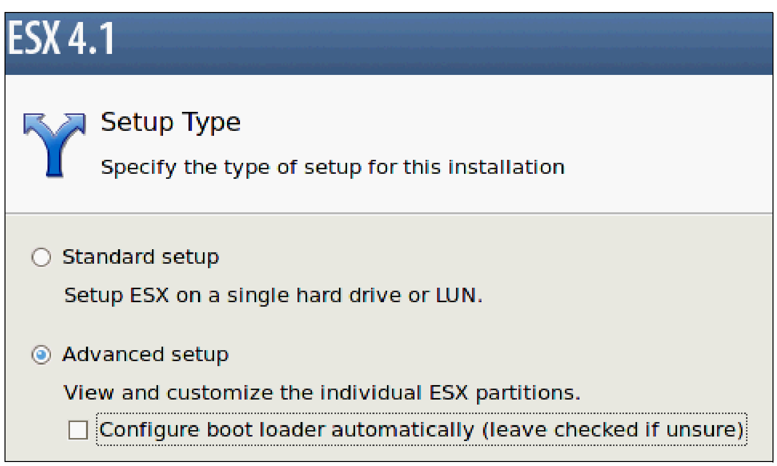


Figure 6-56 Modifying the ESX installation

11. Proceed through the installer to the Set Bootloader Options page and type the following parameter in the Kernel Arguments text box, as shown in Figure 6-57:

```
allowInterleavedNUMAnodes=TRUE
```

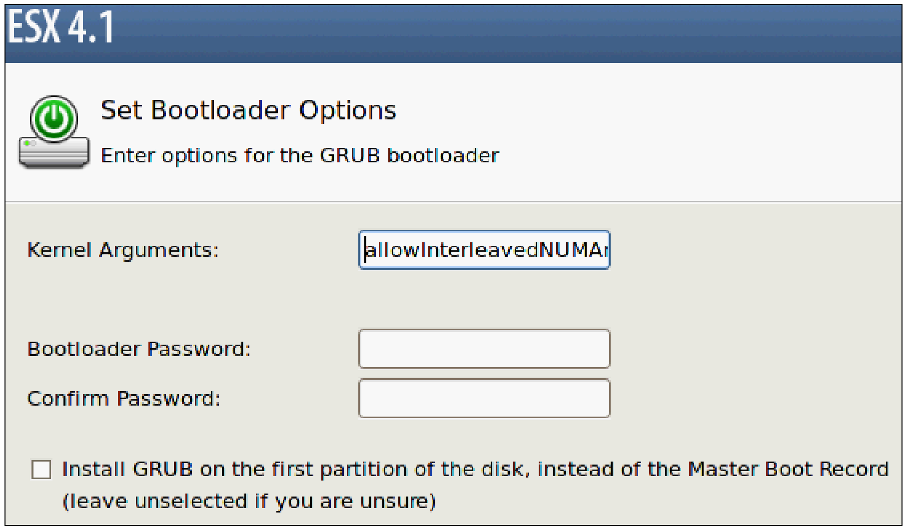


Figure 6-57 Editing the bootloader options

12. Complete the remainder of the ESX installation and reboot the host.

Installing the ESXi 4.1 OS

VMware ESXi 4.1 is not UEFI aware. Therefore, you must use the Legacy Only option as the first boot option. To decrease the boot time, the optimal minimum configuration is Legacy Only, CD/DVD Rom, and Hard Disk 0.

Use the following steps to install ESXi 4.1 Server onto x3850 X5 with MAX5 attached:

1. Power on the system and press F1 when the UEFI splash panel is shown.
2. Select **Boot Manager** → **Add Boot Option**.
3. Select **CD/DVD Rom**, **Legacy Only**, and **Hard Disk 0**. If either option is not listed, the option is already included in the boot list.
4. Press Esc when finished to go back one panel.
5. Select **Change Boot Order** and change the boot order to look like Figure 6-53 on page 276.

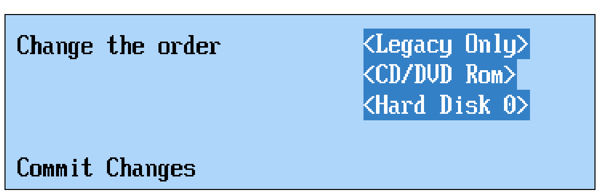


Figure 6-58 Example of a boot order

6. Select **Commit Changes** and press Enter to save the changes.
7. Boot from the ESXi Installable installation media.
8. Press the Tab key when the blue boot panel appears, as shown in Figure 6-59 on page 279.

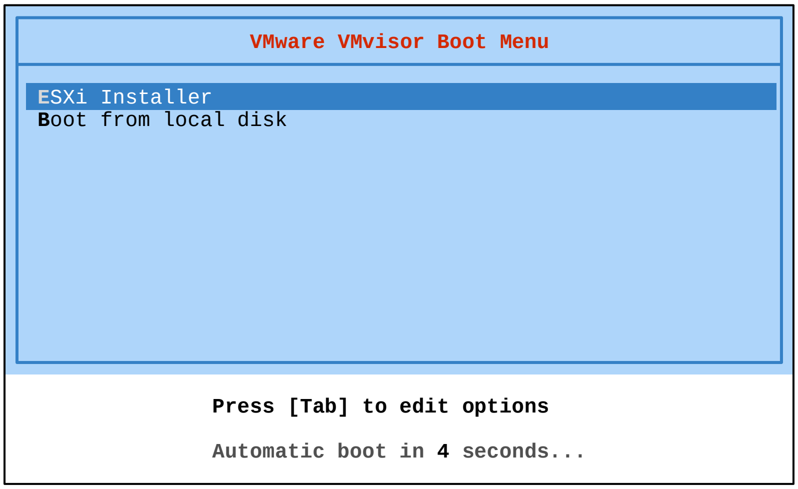


Figure 6-59 Installing ESXi 4.1 Installable edition

9. Add the following line after vmkboot.gz:

   `allowInterleavedNUMAnodes=TRUE`

   Ensure that you leave a space at the beginning and the end of the text that you enter (as shown in Figure 6-60); otherwise, the command will fail to execute at a later stage during the installation. Press Enter to proceed after the line has been edited correctly.

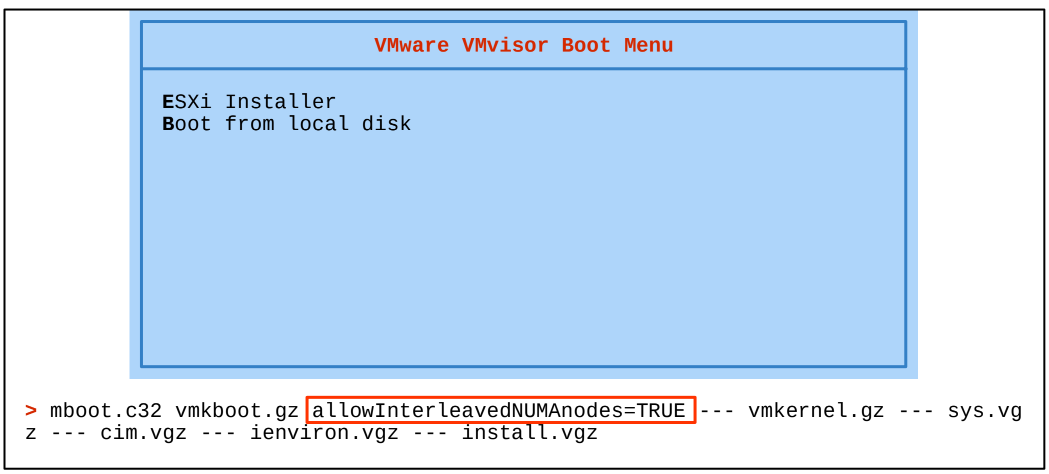


Figure 6-60 Editing the boot load command

10. Complete the ESXi installation and reboot when prompted. Ensure that you remove the media or unmount the installation image before the system restarts.
11. In the Loading VMware Hypervisor panel, press Shift+O when the progress bar is displayed, as shown in Figure 6-61 on page 280.

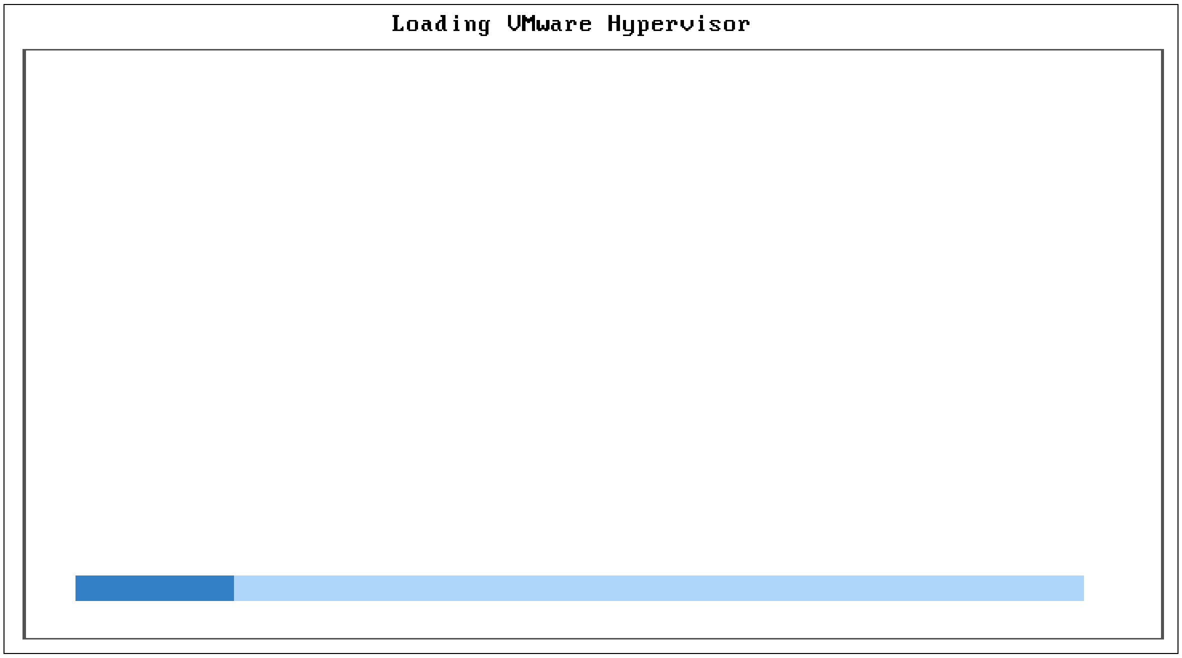


Figure 6-61 Loading VMware Hypervisor panel

> **IMPORTANT:** If you do not press Shift+O during the Loading VMware Hypervisor panel, you will receive the error that is shown in Figure 6-62:
>
> "`The system has found a problem on your machine and cannot continue. Interleaved NUMA nodes are not supported.`"

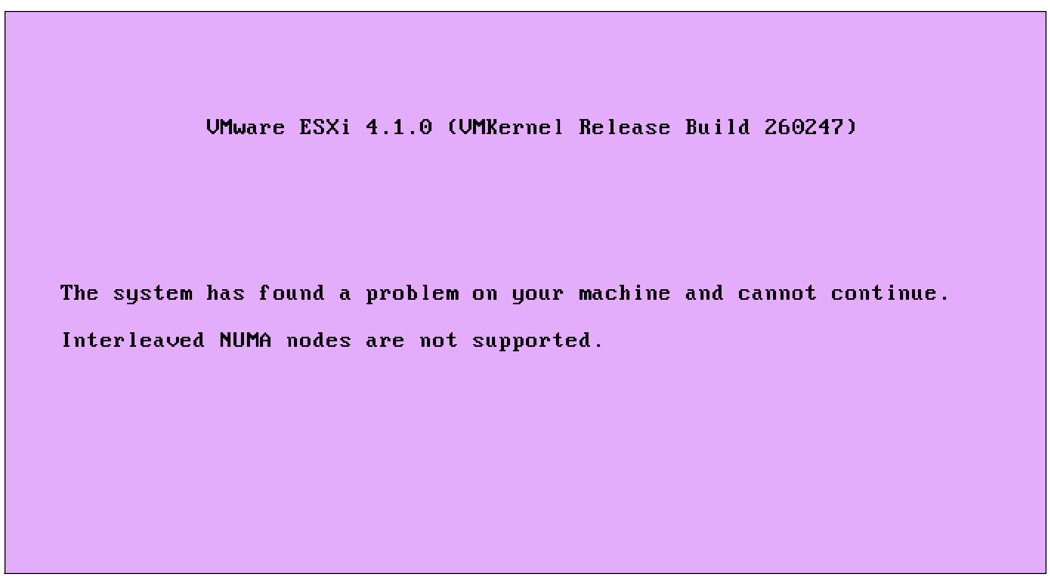


Figure 6-62 ESXi 4.1 Installable NUMA error

12. Enter the following command at the prompt after you have pressed Shift+O:

```
esxcfg-advcfg -k TRUE allowInterleavedNUMAnodes
```

Your output looks like Figure 6-63 on page 281.

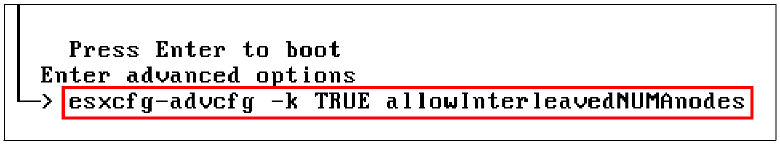


Figure 6-63 Loading VMware Hypervisor boot command

13. Press Enter after the command has been entered. Press Enter again to continue to boot.
14. After the system boots, connect to it using the vSphere Client.
15. Select the **Configuration** tab of the host and click **Advanced Settings** under the Software panel, as shown in Figure 6-64.

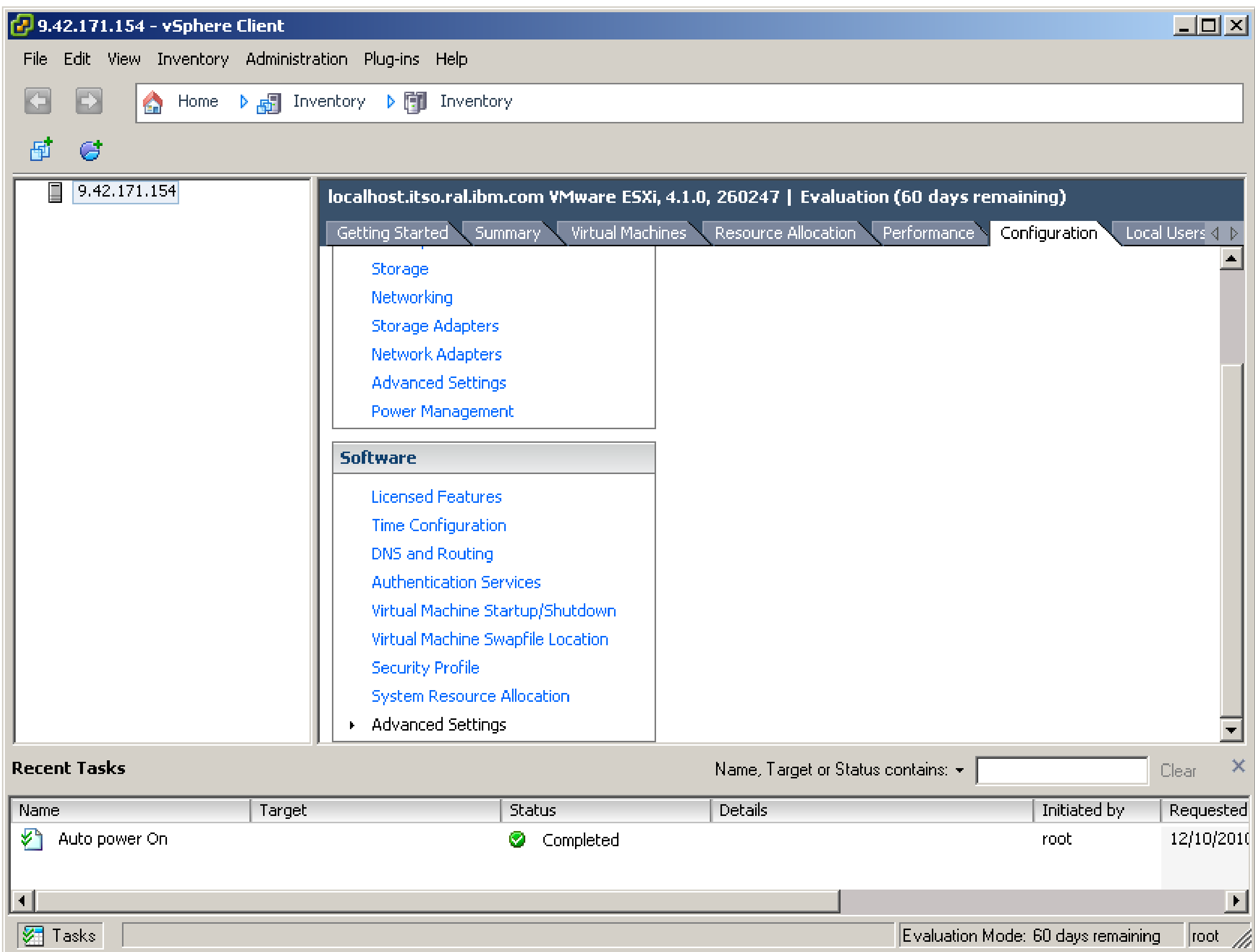


Figure 6-64 Advanced Settings in vSphere Client

16. Click **VMkernel** in the left pane and select **VMkernel.Boot.allowInterleavedNUMAnodes**, as shown in Figure 6-65 on page 282. Click **OK** when finished. This step concludes the installation process.

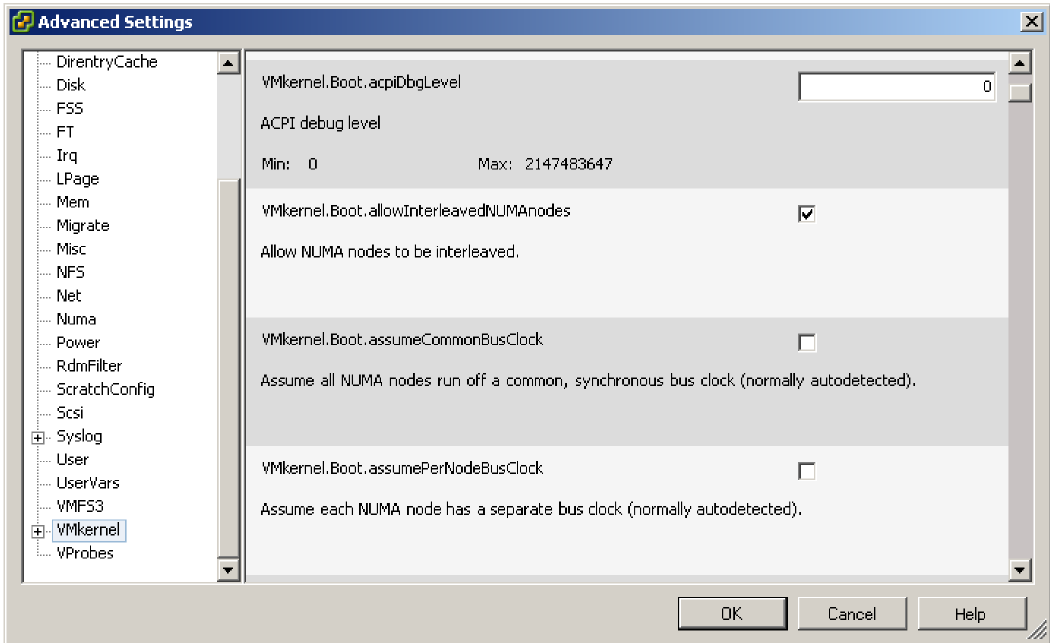


Figure 6-65 Editing the VMkernel settings in vSphere Client

Attaching MAX5 to an existing VMware ESX/ESXi 4.1 installation

Next, we add the MAX5 Memory option to an existing VMware ESX/ESXi installation on the x3850 X5. Perform these actions:

1. To successfully scale an x3850 X5 with MAX5, see 6.4, “Attaching the MAX5 memory expansion unit” on page 230.
2. Systems running VMware ESX/ESXi Server must use Non-Pooled mode in the MAX5 Memory Scaling option. See “1-node with MAX5” on page 261 for instructions to configure the MAX5 Memory Scaling option.
3. Use the following procedures to change your ESX/ESXi 4.1 installation:
   - “ESX installation”
   - “ESXi installation” on page 285

ESX installation

Use the following steps to change your ESX 4.1 installation:

1. Use the vSphere Client to connect to the system.
2. Select the **Configuration** tab of the host and click **Advanced Settings** under the Software panel.
3. Click **VMkernel** in the left pane and select **VMkernel.Boot.allowInterleavedNUMAnodes**, as shown in Figure 6-66 on page 283. Click **OK**.

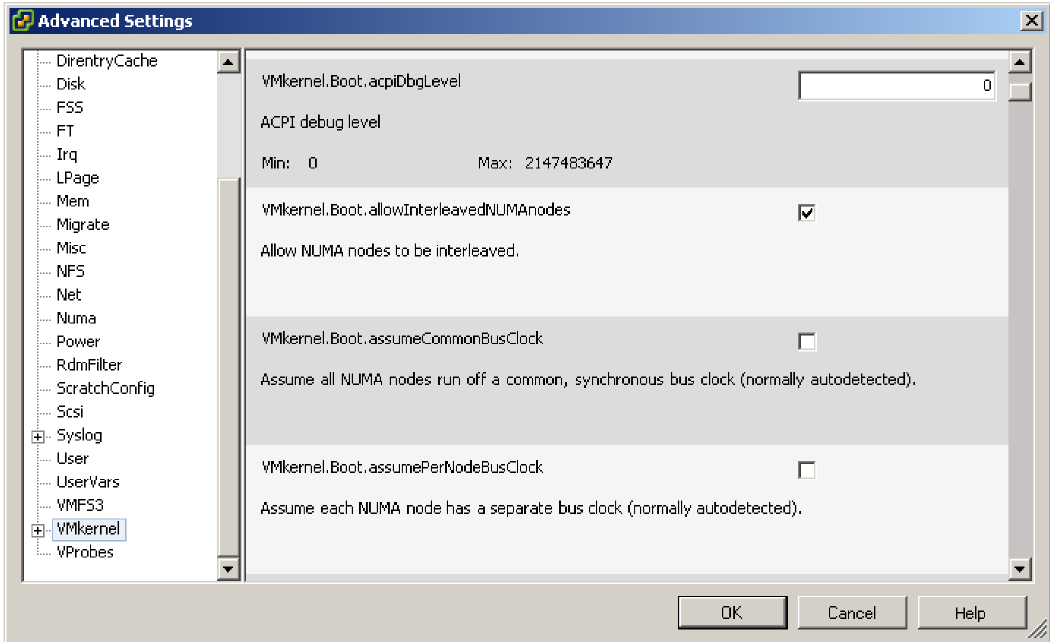


Figure 6-66 Editing the VMkernel settings in vSphere Client

4. Power off the x3850 X5 and unplug all power cords.
5. Attach the MAX5 to the x3850 X5. For more information, see 6.4, “Attaching the MAX5 memory expansion unit” on page 230.
6. Replug the power cords on x3850 X5 and MAX5 but *do not power on the complex configuration.*
7. Log in to the IMM web interface of the x3850 X5.
8. Click **Firmware Update** under Tasks, as shown in Figure 6-67.

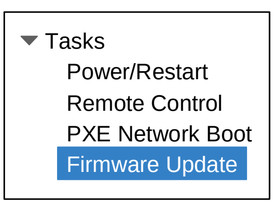


Figure 6-67 IMM menu bar

9. Click **Browse** and select the Field Programmable Gate Array (FPGA) update file.
10. Click **Update** to start the update process. A progress indicator opens as the file is transferred to the temporary storage of IMM.
11. When the transfer is complete, click **Continue** to complete the update process. A progress indicator opens as the firmware is flashed. A confirmation page opens to verify that the update was successful.
12. Unplug the power cords from the x3850 X5 and MAX5. Wait 1 minute and replug the power cords. This procedure activates the new FPGA code on the x3850 X5 and MAX5.
13. Wait approximately 10 minutes to power on the x3850 X5. There cannot be any LEDs lit on the Light Path panel. The System Health Status in the IMM web interface looks like Figure 6-68 on page 284.

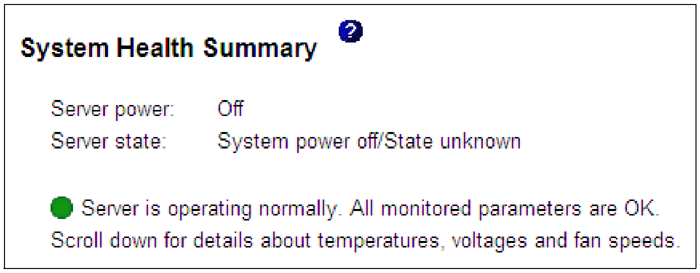


Figure 6-68 IMM System Health Summary

14. Power on the x3850 X5.
15. At the UEFI panel, press F1 for the UEFI menu.
16. Select **System Settings** → **Memory** and change the MAX5 Memory Scaling Affinity to **Non-Pooled**, as shown in Figure 6-69.

Figure 6-69 UEFI setting for MAX5 Memory Scaling Affinity

17. Press Esc two times to go back to the Main page. Select **Save Settings** and press Enter to save the UEFI settings.
18. Exit the UEFI and boot the system to ESX.

> **Important:** If you do not select **VMkernel.Boot.allowInterleavedNUMAnodes** under VMkernel, Figure 6-70 shows the error that will appear during boot.

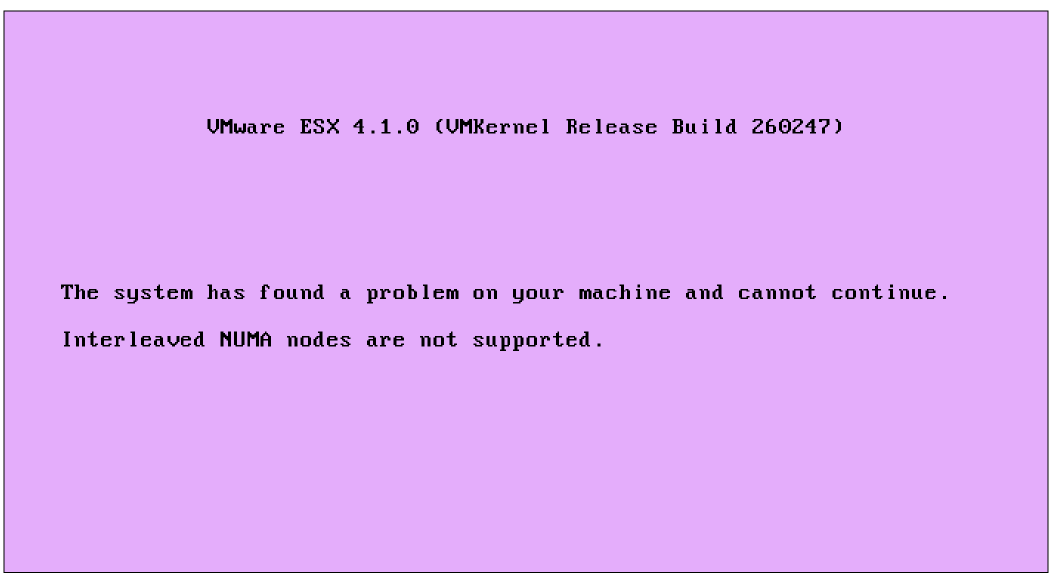


Figure 6-70 ESX 4.1 Installable NUMA error

If you see this error, follow these steps to change the boot code:

1. At the VMware bootloader panel, press a to modify the kernel arguments. Be sure that **VMware ESX 4.1** is highlighted, as shown in Figure 6-71 on page 285.

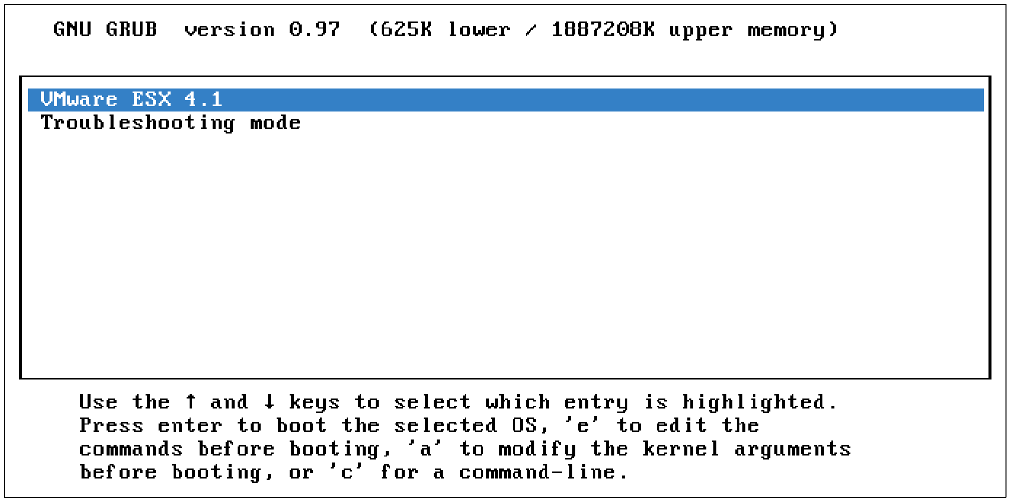


Figure 6-71 VMware ESX GRUB

2. Add the following line at the beginning:

   `allowInterleavedNUMAnodes=TRUE`

   Ensure that you leave a space at the beginning and the end of the text that you enter (as shown in Figure 6-72); otherwise, the command will fail to execute at a later stage during the boot process. Press Enter to proceed after the line has been edited correctly.

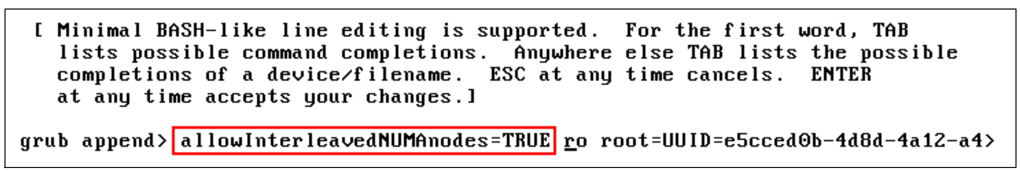


Figure 6-72 Editing the boot load command

3. Do not forget to select **VMkernel.Boot.allowInterleavedNUMAnodes** as active with the vSpere Client.

ESXi installation

Use the following steps to change your ESXi 4.1 installation:

1. Use the vSphere Client to connect to the system.
2. Select the **Configuration** tab of the host and click **Advanced Settings** under the Software panel.

3. Click **VMkernel** in the left pane and select **VMkernel.Boot.allowInterleavedNUMAnodes**. Click **OK**.

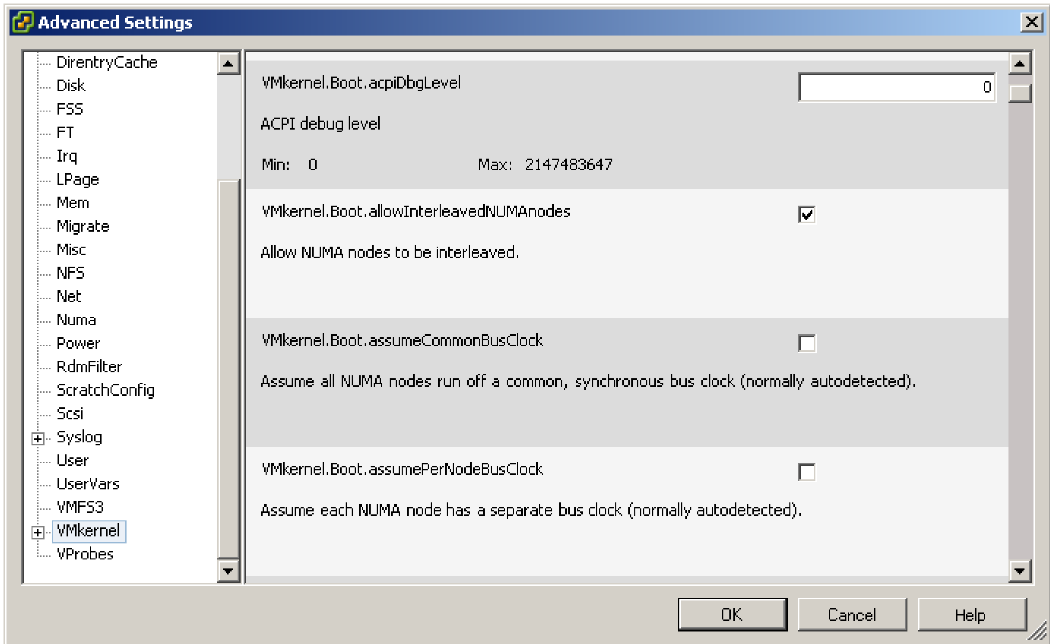


Figure 6-73 Editing the WMkernel settings in vSphere Client

4. Power off the x3850 X5 and unplug all power cords.
5. Attach the MAX5 to the x3850 X5. For more information, see 6.4, “Attaching the MAX5 memory expansion unit” on page 230.
6. Replug the power cords on x3850 X5 and MAX5 but *do not power on the complex configuration.*
7. Log in to the IMM web interface.
8. Click **Firmware Update** under Tasks, as shown in Figure 6-74.

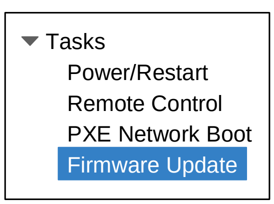


Figure 6-74 IMM menu bar

9. Click **Browse** and select the FPGA update file.
10. Click **Update** to start the update process. A progress indicator opens as the file is transferred to the temporary storage of IMM.
11. When the transfer is completed, click **Continue** to complete the update process. A progress indicator opens as the firmware is flashed. A confirmation page opens to verify that the update was successful.
12. Unplug the power cords from the x3850 X5 and MAX5. Wait 1 minute and replug the power cords. This procedure activates the new FPGA code on the x3850 X5 and MAX5.

13. Wait approximately 10 minutes to power on the x3850 X5. There must be no LEDs lit on the light path diagnostics panel. The System Health Status in the IMM web interface looks like Figure 6-75.

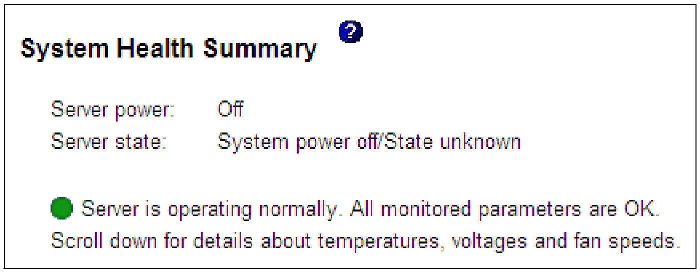


Figure 6-75 IMM System Health Summary

14. Power on the x3850 X5.
15. At the UEFI panel, press F1 for the UEFI menu.
16. Select **System Settings** → **Memory** and change the MAX5 Memory Scaling Affinity to **Non-Pooled**, as shown in Figure 6-76.

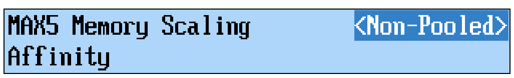


Figure 6-76 UEFI setting for MAX5 Memory Scaling Affinity

17. Press Esc two times to go back to the Main page. Select **Save Settings** and press Enter to save the UEFI settings.
18. Exit the UEFI and boot the system to ESXi.

> **Important:** If you forgot to select **VMkernel.Boot.allowInterleavedNUMAnodes** under VMkernel, Figure 6-77 shows the error that will appear during boot.

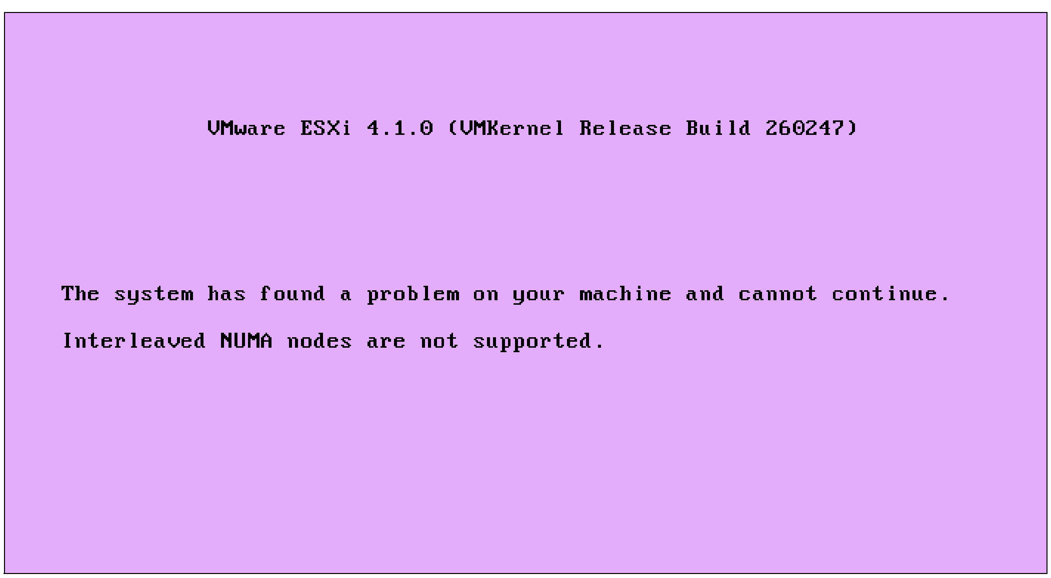


Figure 6-77 ESXi 4.1 Installable NUMA error

If this error occurs, follow these steps to change the boot code:

1. In the Loading VMware Hypervisor panel, press Shift+O when the progress bar is displayed, as shown in Figure 6-78.

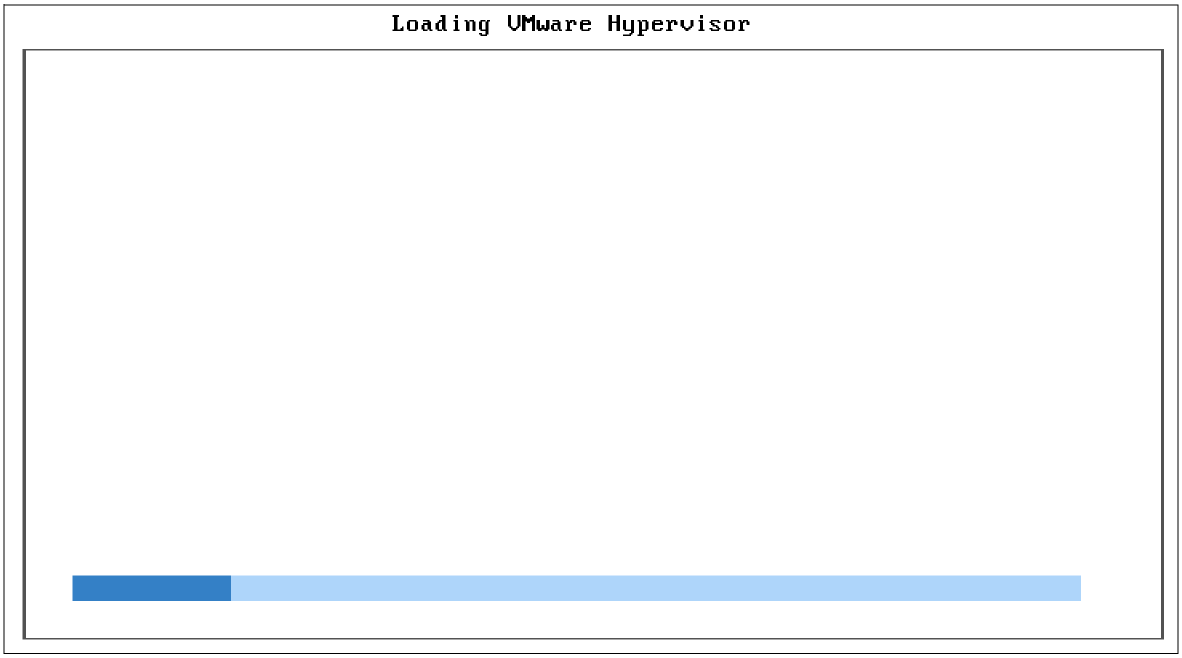


Figure 6-78 Loading WMware Hypervisor panel

2. Enter the following command at the prompt after you have pressed Shift+O:

   `esxcfg-advcfg -k TRUE allowInterleavedNUMAnodes`

   Your output looks like Figure 6-63 on page 281.

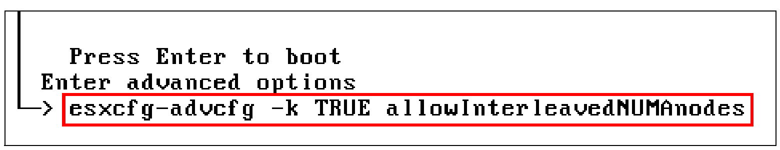


Figure 6-79 Loading WMware Hypervisor boot command

3. Press Enter after the command has been entered. Press Enter again to continue to boot.
4. Do not forget to select **VMkernel.Boot.allowInterleavedNUMAnodes** as active with the vSpere Client.

6.10.4 OS installation tips and instructions on the web

In this section, we provide information about the OS installation guides on the IBM website and describe certain (but not all) issues you might encounter.

Microsoft Windows Server 2008

This section provides useful information to assist with the installation of Windows Server 2008. The installation process has not been covered in this IBM Redbooks publication, because there are no particular deviations from a standard Windows Server installation. You can obtain a complete list of supported Windows Server OSs for x3850 X5 at this website:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/microsoft.html

Consider these useful tips:

- The following drivers are not included on the Windows Server 2008 DVD and must be downloaded separately (or use ServerGuide instead):
  - ServeRAID M5000 series RAID driver
  - Intel chip set driver
  - Broadcom NIC driver
  - Dual-port Emulex 10GbE NIC driver
- If installing Windows 2008 Enterprise Edition on an 8-socket system, Hyper-Threading must be turned off prior to installation or the system will stop with a blue window.
- The default installation installs the OS in UEFI mode. UEFI mode requires a separate imaging methodology, because UEFI mode requires the OS to reside on a GUID Partition Table (GPT) disk.
- The system can re-enumerate the boot drive when fibre devices are connected. This situation will cause the OS to not boot properly. Use Windows Boot Manager from the UEFI Boot Manager to fix this problem.
- You must set the MAX5 memory configuration in UEFI to Non-pooled mode.
- The recommendation is to review all RETAIN tips that are related to this system and the OS that you are installing. See the following website:

  http://ibm.com/support/entry/portal/Problem_resolution/Hardware/Systems/System_x/System_x3850_X5
- For further tuning information, see the Performance Tuning Guidelines for Windows Server 2008 at the following page:

  http://www.microsoft.com/whdc/system/sysperf/Perf_tun_srv.mspx

Microsoft Windows Server 2008 R2

This section provides useful information to assist with the installation of Windows 2008 R2. The installation process has not been covered in this IBM Redbooks publication, because there are no particular deviations from a standard Windows Server installation. You can obtain a complete list of supported Windows Server OSs for x3850 X5 at the following website:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/microsoft.html

The following web pages are the IBM installation guides for Windows Server 2008 R2:

- Installing Microsoft Windows Server 2008 R2 - IBM System x3850 X5 (7145, 7146), x3950 X5:

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5083420
- Installing Microsoft Windows Server 2008 R2 - IBM System x3850 X5 (7145, 7146) with MAX5:

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5085823

Consider these useful tips:

- The Dual-port Emulex 10GbE NIC driver is not included on the Windows Server 2008 DVD and must be downloaded separately (or use ServerGuide instead).
- You must ensure that the Flow Control is enabled under network adapter properties after upgrading the network device driver. When Flow Control is enabled, it allows the receipt or transmission of PAUSE frames. PAUSE frames enable the adapter and the switch to control the transmission rate.
- If installing Windows 2008 R2 Enterprise Edition on an 8-socket system, Hyper-Threading must be turned off prior to installation or the systems will show a blue screen.
- To correct an issue with the GPT disks, Windows 2008 R2 requires the Microsoft Hotfix that can be found at the following website:

  http://support.microsoft.com/kb/975535
- The default installation media installs the OS in UEFI mode unless Legacy Only is the first entry in the boot order.
- In 2-node configurations, edit the Windows Boot Configuration Data (BCD) Store to enable the High Precision Event Timer (HPET). For more information, see the following website:

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084072
- The system can re-enumerate the boot drive when fibre devices are hooked up. This action will cause the OS to fail to boot properly. Use Windows Boot Manager from the UEFI Boot Manager to fix this problem.
- If your system has one TB or more of memory installed, it is a requirement to apply the Microsoft Hotfix that is available at the following website:

  http://support.microsoft.com/kb/980598
- You must update with the latest BroadCom Ethernet device driver when your system has 64 or more processors (cores).
- You must set the MAX5 memory configuration in UEFI to Non-pooled mode.
- If your system has greater than 128 GB of memory and you plan to enable Hyper-V after installing Windows 2008 R2, you must first apply the Microsoft Hotfix that is available at the following website:

  http://support.microsoft.com/kb/979903/en-us
- The recommendation is to review all Retain Tips related to this system and the OS that you are installing. See the following website:

  http://ibm.com/support/entry/portal/Problem_resolution/Hardware/Systems/System_x/System_x3850_X5
- For further tuning information, see Performance Tuning Guidelines for Windows Server 2008 R2 at the following website:

  http://www.microsoft.com/whdc/system/sysperf/Perf_tun_srv-R2.mspx

VMware ESX and ESXi

This section provides useful information to assist with the installation of VMware ESX and VMware ESXi. The installation process has not be covered in this IBM Redbooks publication, because there are no particular deviations from a standard VMware installation. You can obtain a complete list of supported VMware Server OSs for x3850 X5 at this website:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/vmware.html

The following websites contain additional OS installation guides:

- Installing VMware ESX Server 4.1 - IBM System x3850 X5 (7145):

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5086213

- Installing VMware ESXi Server 4.1 Installable - IBM System x3850 X5 (7145):

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5086214

Consider these useful tips:

- VMware 4.0 U1 is the minimum supported OS level for a single-node x3850X5.
- The addition of a MAX5 or a second node requires at least VMware 4.1.
- Neither the included Emulex 10GbE NIC or the Emulex VFA, nor the Qlogic CNA driver, is included in VMWare ESX 4.0 U1. Download the drivers from VMWare's site:

  http://downloads.vmware.com/d/info/datacenter_downloads/vmware_vsphere_4/4#drivers_tools

- VMware 4.*x* does not include the driver for the Intel PRO/1000 Ethernet Adapter. Download the drivers from the VMware site:

  http://downloads.vmware.com/d/info/datacenter_downloads/vmware_vsphere_4/4#drivers_tools

- For a 2-node configuration or x3850 X5 with MAX5, you need to edit your `grub.conf` file with the following line:

  ```
  allowInterleavedNUMAnodes=TRUE
  ```

  For more information, see "Installing the ESX/ESXi 4.1 Installable edition onto x3850 X5 with MAX5" on page 275.

- ESX 4.1 generates a warning with 1 TB RAM, because 1 TB RAM is the maximum RAM for an ESX 4.1 host. For more information, see the Configuration Maximums for VMware vSphere 4.1 at the following website:

  http://www.vmware.com/pdf/vsphere4/r41/vsp_41_config_max.pdf

- ESX requires the amount of memory per CPU and per node to be balanced or it does not start. For more information, go to the following website:

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5069570

- ESX requires the MAX5 memory configuration in UEFI to be in Non-pooled mode.
- The recommendation is that you review all Retain Tips related to this system and the OS that you are installing. See the following website:

  http://ibm.com/support/entry/portal/Problem_resolution/Hardware/Systems/System_x/System_x3850_X5

- For further tuning information, see VMware vCenter Server Performance and Best Practices for vSphere 4.1 at the following website:

  http://www.vmware.com/resources/techresources/10145

Red Hat Enterprise Linux

This section provides useful information to assist with the installation of Red Hat Enterprise Linux (RHEL). The installation process has not been covered in this IBM Redbooks publication, because there are no particular deviations from a standard RHEL installation. You can obtain a complete list of supported RHEL versions for x3850 X5 at the IBM ServerProven page:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/redchat.html

The following websites contain additional OS installation guides:

- Installing Red Hat Enterprise Linux Version 6 - IBM System x3850 X5 (Type 7145)

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5086423

- Installing Red Hat Enterprise Linux Version 5 Update 4 - IBM System x3850 X5, x3950 X5

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5083917

Consider these useful tips:

- All system drivers are included in RHEL 5.4.
- With RHEL 5.4 and if the server has 1 TB of installed physical memory, see the Red Hat Knowledgebase article at the following website:

  http://kbase.redhat.com/faq/docs/DOC-25412

- RHEL 5.5 is required for 8-socket scaling.
- RHEL has a tendency to assign eth0 to the Emulex 10GbE NIC, which might be undesirable. The recommended action is to either blacklist the module during PXE installs or from within the operating system, hard-code the Ethernet device names based on their MAC addresses.
- RHEL requires the MAX5 memory configuration in UEFI to be in Pooled mode.
- We recommend that you review all RETAIN tips that relate to this system and the OS. See the following website:

  http://ibm.com/support/entry/portal/Problem_resolution/Hardware/Systems/System_x/System_x3850_X5

SUSE Linux Enterprise Server

This section provides useful information to assist with the installation of SUSE Linux Enterprise Server (SLES). The installation process has not been covered in this IBM Redbooks publication, because there are no particular deviations from a standard SLES installation. You can obtain a complete list of supported SLES versions for x3850 X5 at the ServerProven website:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/suseclinux.html

The following website contains additional OS installation guides.

- Installing SUSE Linux Enterprise Server 11 - IBM System x3850 X5, x3950 X5

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5083918

Consider these useful tips:

- All system drivers are included in SLES 11.
- SLES 11 SP1 can install in UEFI mode. It requires installation on a GPT disk.
- SLES 10 SP3 is supported, but it requires a mass storage driver for the M5015.
- SUSE Linux requires the MAX5 memory configuration in UEFI to be in Pooled mode.
- Because of an issue with SLES 11 and certain 4-core and 6-core processors that are supported on x3850 X5 servers, you must use an updated kernel during and after the installation. The updated CD is available at the following website:

  http://drivers.suse.com/ibm/x3850-X5/sle11/install-readme.html

- The recommendation is to review all Retain Tips related to this system and the OS you are installing. See the following website:

  http://ibm.com/support/entry/portal/Problem_resolution/Hardware/Systems/System_x/System_x3850_X5

6.10.5 Downloads and fixes for x3850 X5 and MAX5

Typically, updates are released to provide clients with enhanced capabilities, extended functions, and problem resolutions. Most of the updates are firmware, drivers, and OS patches.

We recommend performing a scheduled review of available updates to determine if they are applicable to the systems that are used in your environment.

Server firmware

Software that resides on flash memory and controls the lower-level function of server hardware is called *server firmware*. An IBM System, such as the x3850 X5, runs a number of firmware images that control separate components of the server.

This list shows the primary firmware for the x3850 X5:

- Unified Extensible Firmware Interface (UEFI)
- Integrated Management Module (IMM)
- Field-Programmable Gate Array (FPGA)
- Preboot Dynamic System Analysis (DSA)

Additional devices, such as network cards and RAID controllers, also contain their own firmware revisions.

Firmware updates are provided by IBM and can be downloaded from the IBM website, including proven firmware from other manufacturers to be applied on IBM systems. We describe several methods of performing firmware updates on IBM eX5 servers in Chapter 9, “Management” on page 447.

Tip: It is a recommended practice to update all System x firmware to the latest level prior to performing an OS or application installation.

Tip: IBM Bootable Media Creator (BoMC) is a tool that simplifies the IBM System x firmware update process without an OS running on the system. More information about this tool is available at the following website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-BOMC

Device drivers

Device drivers are software that controls hardware server components on the OS level. They are specific to the OS version, and therefore, critical device drivers are included with the installation media.

Device driver updates are provided by IBM, OS vendors, and component device vendors. They are mostly downloadable from each company’s support website.

Whenever possible, we recommend using the tested and approved driver updates from IBM.

Tip: IBM UpdateXpress is a tool that allows the IBM System x firmware and drivers to be updated via the OS. More information about this tool is available at the following website:

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

Operating system updates, fixes, and patches

The performance and reliability of an x3850 X5 tightly relate to the OS running on it. IBM supports an assortment of modern and widely used OSs capable of utilizing the system's potential. Each vendor supports their OS by releasing updates, fixes, and patches that provide enhanced functionality and fixes to known problems. Several of these updates, fixes, and patches only apply to certain configurations, while other updates, fixes, and patches apply to all configurations. The OS vendors support website has extensive information about these updates, fixes, and patches.

System update resources

Table 6-12 provides useful web address links to IBM tools, as well as vendor OS support links.

Table 6-12 Internet links to supports and downloads

Vendor	Product	Address
IBM	Systems Support	http://ibm.com/systems/support/
IBM	ServerGuide	http://ibm.com/support/entry/portal/docdisplay?lndocid=SERV-GUIDE
IBM	UpdateXpress	http://ibm.com/support/entry/portal/docdisplay?lndocid=SERV-XPRESS
IBM	Bootable Media Creator	http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-BOMC
Microsoft	Windows Server	http://support.microsoft.com/ph/14134
Red Hat	RHEL	https://www.redhat.com/support/
Novell	SLES	http://www.novell.com/support/
VMware	vSphere	http://downloads.vmware.com/d/

6.10.6 SAN storage reference and considerations

The System x3850 X5 with its MAX5 memory expansion capability is the high-end server solution in the IBM System x product line. Target workloads for this design include virtualization and database application. In both designations, it is typical for the user to attach storage area network (SAN) storage to the server for data storage.

SAN storage attachment

A storage area network (SAN) is a network whose primary purpose is the transfer of data between computer systems and storage elements. The following list describes the typically used SAN protocols, each with their own characteristics:

- Fibre Channel (FC)

  The Fibre Channel Protocol (FCP) is the interface protocol of SCSI on Fibre Channel. FCP is a transport protocol, which predominantly transports SCSI commands over Fibre Channel networks. Fibre Channel (FC) is the prevalent technology standard in the Storage Area Network (SAN) data center environment. Typical requirements for this configuration require an FC HBA and FC SAN infrastructure. Despite its name, Fibre Channel signaling can run on both twisted-pair copper wire and fiber optic cables.

- Fibre Channel over Ethernet (FCoE)

  FCoE is the transport, or mapping, of encapsulated FC frames over the Ethernet. The Ethernet provides the physical interface, and FC provides the transport protocol. System setup for FCoE requires Converged Network Adapter (CNA) to pass both network and storage data connected to 10Gb converged network infrastructure.

- Internet SCSI (iSCSI)

  iSCSI is an Internet Protocol (IP)-based storage networking standard for linking data storage facilities. The protocol allows clients (called *initiators*) to send SCSI commands to SCSI storage devices (targets). A hardware initiator might improve the performance of the server. iSCSI is often seen as a low-cost alternative to Fibre Channel.

- serial-attached SCSI (SAS)

  SAS uses point-to-point connection and the typical SAS throughput is 6 Gbps full duplex. Where a complex SAN configuration is not necessary, SAS is a good choice, although performance and distance will be limited compared to the other solutions.

Booting from SAN

This section provides useful guidelines for booting from SAN:

- See 6.1.1, “Verify that the components are securely installed” on page 220 to ensure that all PCI adapters are seated properly.
- Check if UEFI recognizes the adapter. Select **UEFI** → **System Settings** → **Adapters and UEFI Drivers**. The Adapters and UEFI Drivers panel displays, as shown in Figure 6-80. You need to see Card - HBA. If not, reflash the UEFI, IMM, and Firmware of the HBA and check again.

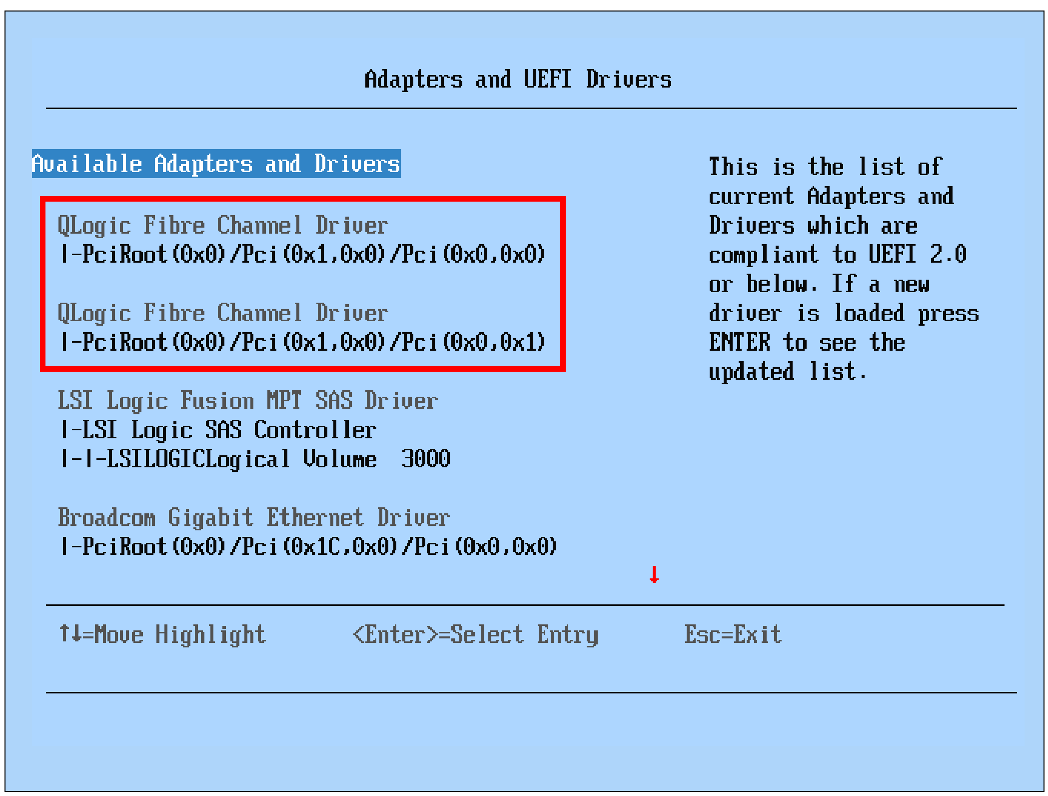


Figure 6-80 Adapters visible in UEFI

- If you do not have internal drives, disable the onboard SAS RAID Controller by selecting **System Settings** → **Devices and IO ports** → **Enable/Disable Onboard Devices** and disabling the SAS Controller or Planar SAS.
- Set the HBA as the first device in the Option ROM Execution Order by selecting **System Settings** → **Devices and IO Ports** → **Set Option ROM Execution Order**.
- For legacy OSs only (all OSs except Windows 2008 and SLES 11 SP1), set **Legacy Only** as the first boot device.
- Remove all devices, which might not host an OS, from the boot order. The optimal minimum configuration is CD/DVD and Hard Disk 0. For legacy OSs only, set **Legacy Only** as the first boot device.
- Enable the BIOS from your HBA.
- Verify that your HBA can see a logical unit number (LUN) from your storage.
- Make sure that the LUN is accessible through only one path, either Zoning or LUN masking.
- After installation, do not forget to install the multipath driver *before* you set more than one path, if you have more than one path to the LUN.

IBM Redbooks references for SAN-related information

A number of IBM Redbooks publications are available for reference. The following IBM Redbooks publications describe IBM System Storage products and their various implementations, including with IBM System x product lines:

- *IBM System Storage Solutions Handbook*, SG24-5250

  This book provides overviews and pointers for information about the current IBM System Storage products:

  http://www.redbooks.ibm.com/abstracts/sg245250.html

- *Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*, SG24-6116

  This book consolidates critical information while also covering procedures and tasks that you are likely to encounter on a daily basis when implementing an IBM/Brocade SAN:

  http://www.redbooks.ibm.com/abstracts/sg246116.html

- *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363

  This book represents a compilation of best practices for deploying and configuring IBM Midrange System Storage servers, which include the DS4000® and DS5000 family of products:

  http://www.redbooks.ibm.com/abstracts/sg246363.html

- *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065

  This book introduces the IBM System Storage DS3000, providing an overview of its design and specifications, and describing in detail how to set up, configure, and administer it:

  http://www.redbooks.ibm.com/abstracts/sg247065.html

- *Implementing an IBM/Cisco SAN*, SG24-7545

  This book consolidates critical information while discussing procedures and tasks that are likely to be encountered on a daily basis when implementing an IBM/Cisco SAN:

  http://www.redbooks.ibm.com/abstracts/sg247545.html

- *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659

  This book describes the concepts, architecture, and implementation of the IBM XIV® Storage System:

  http://www.redbooks.ibm.com/abstracts/sg247659.html

- *IBM Midrange System Storage Hardware Guide*, SG24-7676

  This book consolidates, in one document, detailed descriptions of the hardware configurations and options offered as part of the IBM Midrange System Storage servers, which include the IBM System Storage DS4000 and DS5000 families of products:

  http://www.redbooks.ibm.com/abstracts/sg247676.html

For more information regarding HBA storage-specific settings and zoning, contact your SAN vendor or storage vendor.

6.11 Failure detection and recovery

This section provides an overview of tools available to assist with problem resolution for the x3850 X5 in any given configuration. It also provides considerations for extended outages.

6.11.1 What happens when a node fails or the MAX5 fails

If you have power problems and one or both nodes fail or the MAX5 is no longer supplied with power, the complex configuration will shut down to avoid any damage (data loss, corrupt data, and so on). No OS can handle this sudden change to the system.

The MAX5 is turned off only if the connected server issues a power-off request and you have disconnected the MAX5 power cord from the power source. You cannot turn off the MAX5 expansion module manually.

For recovery options, see 6.11.4, “Recovery process” on page 299.

6.11.2 Reinserting the QPI wrap cards for extended outages

If one node becomes unavailable for any reason, you have the capability to boot your system in a single-node configuration. If you have QPI wrap cards, install the QPI cards for your system. The QPI wrap cards are not mandatory, but they provide a performance boost by ensuring that all CPUs are only one hop away from each other. For more information about QPI cards, see 3.4.2, “QPI Wrap Card” on page 66.

For recovery options, see 6.11.4, “Recovery process” on page 299.

6.11.3 Tools to aid hardware troubleshooting for x3850 X5

Use the following tools when troubleshooting problems on the x3850 x5 in any configuration.

Integrated Management Module

The first place to start troubleshooting the x3850 X5 is typically the IMM. Use the links under the Monitors heading to view the status of the server, as shown in Figure 6-81 on page 298.

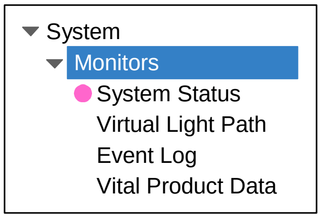


Figure 6-81 IMM web interface

From the System Status pages, you can perform these tasks:

- Monitor the power status of the server and view the state of the OS.
- View the server temperature readings, voltage thresholds, and fan speeds.
- View the latest server OS failure screen capture.
- View the list of users who are logged in to the IMM.

From the Virtual Light Path page, you can view the name, color, and status of any LEDs that are lit on a server.

From the Event Log page, you can perform these tasks:

- View certain events that are recorded in the event log of the IMM.
- View the severity of events.

For more information about the IMM, see 9.2, “Integrated Management Module (IMM)” on page 449.

Light path diagnostics panel

You can use the light path diagnostics to diagnose system errors quickly. When an LED is lit on the light path diagnostics panel, it helps you to isolate the error. The server is designed so that LEDs remain lit when the server is connected to an ac power source but is not turned on, provided that the power supply operates correctly. This feature helps you to isolate the problem when the OS is shut down.

For more information, see the *Problem Determination and Service Guide - IBM System x3850 X5, x3950 X5 (7145, 7146)* at the following website:

`http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084848`

System event log

This log contains POST and system management interrupt (SMI) events and all events that are generated by the Baseboard Management Controller (BMC) that is embedded in the IMM. You can view the system event log through the UEFI by pressing F1 at system start-up and selecting **System Event Logs** → **System Event Log**.

POST event log

This log contains the three most recent error codes and messages that were generated during POST. You can view the POST event log through the UEFI by pressing F1 at system start-up and selecting **System Event Logs** → **POST Event Viewer**.

IBM Electronic Service Agent

With an appropriate hardware maintenance and warranty contract, Electronic Service Agent™ enables your system to call home to submit diagnostic information and system statistics, report a problem, and, if a fix is available, download the solution immediately.

For more information and to download Electronic Service Agent, go to the following website:

https://www-304.ibm.com/support/electronic/portal/navpage.wss?category=5&locale=en_US

6.11.4 Recovery process

This section provides an overview of several recovery procedures. These procedures do *not* replace the *Problem Determination and Service Guide*. You can solve many problems without outside assistance by following the troubleshooting procedures in the *Problem Determination and Service Guide*, which is available at the following website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084848

The *Problem Determination and Service Guide* describes the diagnostic tests that you can perform, troubleshooting procedures, and explanations of error messages and error codes.

If you have completed the diagnostic procedure and the problem remains, and you have verified that all code is at the latest level, and all hardware and software configurations are valid, contact IBM or an approved warranty service provider for assistance.

Two-node configuration does not power on after a failure

The following list provides useful tips for bringing a 2-node configuration back online after a failure:

- Check the IMM Event Log.
- Check the LEDs on the light path diagnostic panel.
- Make sure that the power cord is connected to a functioning power source.
- Make sure that the power cord is fully seated in the power supply.
- Check the power-supply LEDs.
- Check the QPI link LEDs. When the QPI link LEDs are lit, they indicate that the QPI links are fully established.
- Remove all power cords from both nodes. Wait a few seconds and replug the cords in the following order: Top PS 2, Bottom PS 1, Top PS 1, and Bottom PS 2. You will need to have power to all nodes at relatively the same time to ensure that the IMMs can communicate across the QPI cables.
- Remove the QPI cables to separately debug the servers. Before you remove the QPI cables, the power *must* be removed.

One node fails in a 2-node configuration

In certain cases, resolving an issue in a 2-node configuration takes longer because it is a difficult task. When you need the system back online as quickly as possible, you can operate the system as a single-node configuration. Use the following steps to configure the node that is still working:

1. Remove all power cords from the working node.
2. Remove all QPI cables from the working node.
3. If available, install the QPI wrap cards.
4. If the working node is not fully populated with memory, add the memory of the second node. Remember the installation sequence for the memory.

5. If the working node does not have hard drives installed, add the hard drives with the operating system from the second node. At the boot sequence, you will get a prompt to import the RAID configuration.
6. Reconnect the power cords.
7. Power on the system.

Power failure on the MAX5

The MAX5 is turned off only if the connected server issues a power-off request and you have disconnected the MAX5 power cord from the power source. You cannot turn off the MAX5 expansion module manually. Use the following steps to boot your system without the MAX5:

1. Remove all power cords from the MAX5 and x3850 X5.
2. Remove all QPI cables.
3. If available, install the QPI wrap cards for the x3850 X5.
4. Reconnect the power cords to the x3850 X5.
5. Power on the system.

7

IBM System x3690 X5

The IBM System x3690 X5 offers flexibility in design to support a wide variety of uses. The x3690 X5 supports up to16 internal drive bays and offers up to 7.5 TB of redundant storage. This system is ideal for most general-purpose file servers.

Figure 7-1 System x3690 X5

This chapter provides assistance for making configuration, monitoring, and maintenance decisions for the x3690 X5 server. The information provided is meant to help you make informed decisions; however, the information provided must not be considered as an absolute implementation process.

This chapter covers the following topics:

7.1 Before you apply power for the first time after shipping

Before you begin populating your server with all of its processors, memory, and PCI adapters, and before you install an operating system (OS), perform the recommendations provided in this section.

7.1.1 Verify that the components are securely installed

Perform the following tasks to ensure that all of the electrical components of your server have proper connectivity:

- Inspect heat sinks to ensure that they are secure.
- Verify that dual inline memory modules (DIMMs) are mounted in the correct locations and are fully plugged in with their retain clips in the closed position.
- Inspect the PCIe adapters to ensure that they are securely plugged into their slots.
- Check all of the cable connections on the hard drive backplane, system board, and all internal disk controllers to ensure that they are properly snapped into place. A cable that can easily be unplugged from its connector must be plugged back in until it clicks or clips into place.

7.1.2 Clear CMOS memory

When a server is shipped from one location to another location, you have no idea what the server was exposed to. For all you know, it might have been parked next to a large magnet or electric motor, and everything in the server that stores information magnetically has been altered, including the CMOS memory. IBM does not indicate on the shipping carton that magnetic material is enclosed because the information is readily recoverable.

Booting the server to the F1 system configuration panel and selecting **Load Default Settings** will restore the default values for items that you can change in configuration. It will not change the setting of internal registers used between the Integrated Management Module (IMM) and the Unified Extensible Firmware Interface (UEFI). These registers define the system state of the server. If they become corrupt, the server can experience these problems:

- Fail to power on
- Fail to complete power-on self test (POST)
- Turn on amber light path diagnostic lights that describe conditions that do not exist
- Reboot unexpectedly
- Fail to detect all of the installed CPU, memory, PCIe adapters, or physical disks

These internal registers cannot be modified or restored to defaults by the F1 system configuration panel; however, they can be restored to defaults by clearing the CMOS memory.

With ac power removed from the server, CMOS memory can be cleared in one of two ways:

- Setting switch 1 on switch bock SW2 to the ON position for 30 seconds:
  a. Disconnect the ac power from the server and remove the cover.
  b. Locate the switch block SW2 in the back right of the server as you face the front of the server, as shown in Figure 7-2 on page 303. If the optional PCIe riser card is installed, it will need to be removed from the server first.

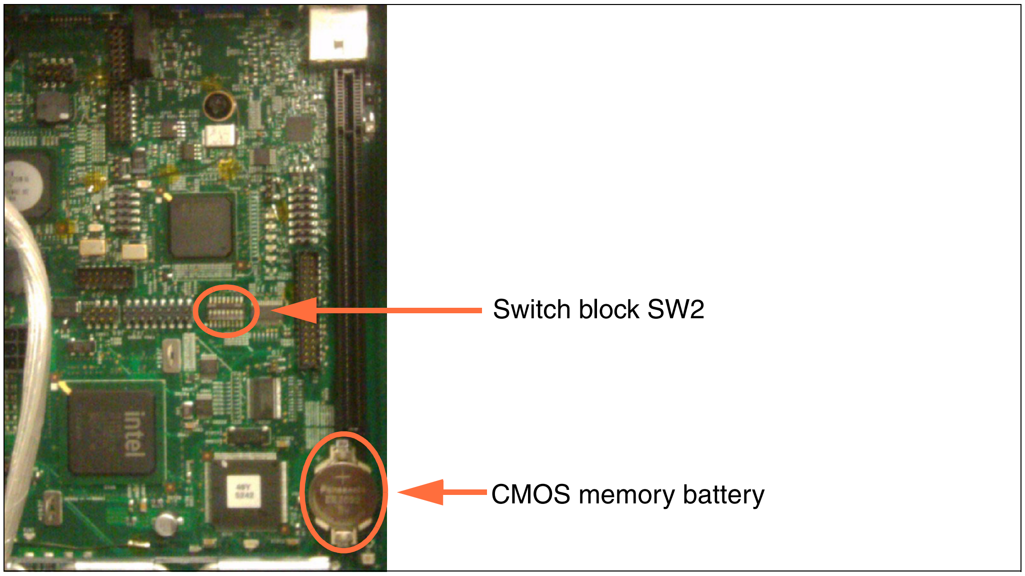


Figure 7-2 Location of switch block SW2 and the CMOS memory battery

c. The numbers on the switch block represent the OFF side of the switch. They are located on the side of the switch block that is closest to the front of the server. To clear CMOS, slide switch 1 (shown in Figure 7-3) to the ON position closest to the rear of the server.

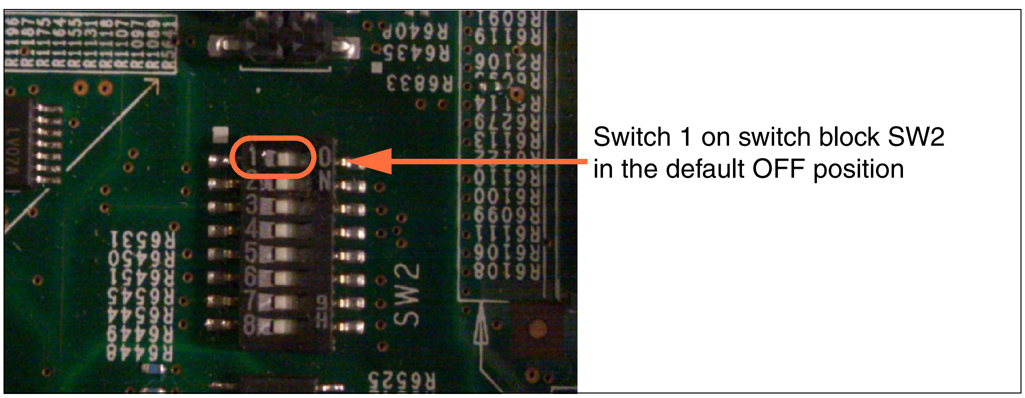


Figure 7-3 Location of switch 1 on switch block SW2

- Pull the CMOS memory battery for 30 seconds:

  a. Disconnect the ac power from the server and remove the cover.

  b. Locate the CMOS battery (shown in Figure 7-2).

  c. Use your finger to pry up the battery on the side closest to the neighboring IC chip. The battery will easily lift out of the socket.

  > **Note:** The light path diagnostic (LPD) lights are powered from a separate power source (capacitor) than the CMOS memory. LPD lights will remain lit for a period of time after ac power and the CMOS memory battery have been removed.

  d. After 30 seconds, insert one edge of the battery, with the positive side up, back into the holder.

e. Push the battery back into the socket with your finger and clip it back into place.

7.1.3 Verify that the server will complete POST before adding options

When you have ordered options for your server that have not yet been installed, it is a good idea to ensure that the server will complete POST properly before you start to add your options. Performing this task makes it easier to compartmentalize a potential problem with an installed option rather than having to look at the entire server to try to find a good starting point during problem determination.

7.2 Processor considerations

> **Tip:** To understand the information in this section, first read the information in 4.7, "Processor options" on page 130.

The x3690 X5 server supports up to two matched processors. The required match is in regard to the family of processors, the number of cores, the size of Level 2 cache, the core, and front side bus speeds.

As a matter of standard manufacturing, the producing vendor might alter the method of manufacturing the processor, which will result in various stepping levels, but this process will not affect the overall functionality of the processor in its ability to communicate with other processors in the same server. Any operational differences in stepping levels are handled by the microcode of the processor, the Integrated Management Module (IMM), and the UEFI.

7.2.1 Minimum processors required

The minimum number of processors required for the server to boot into any operational configuration is one. The processor must be installed in socket 1, as shown in Figure 7-4 on page 305.

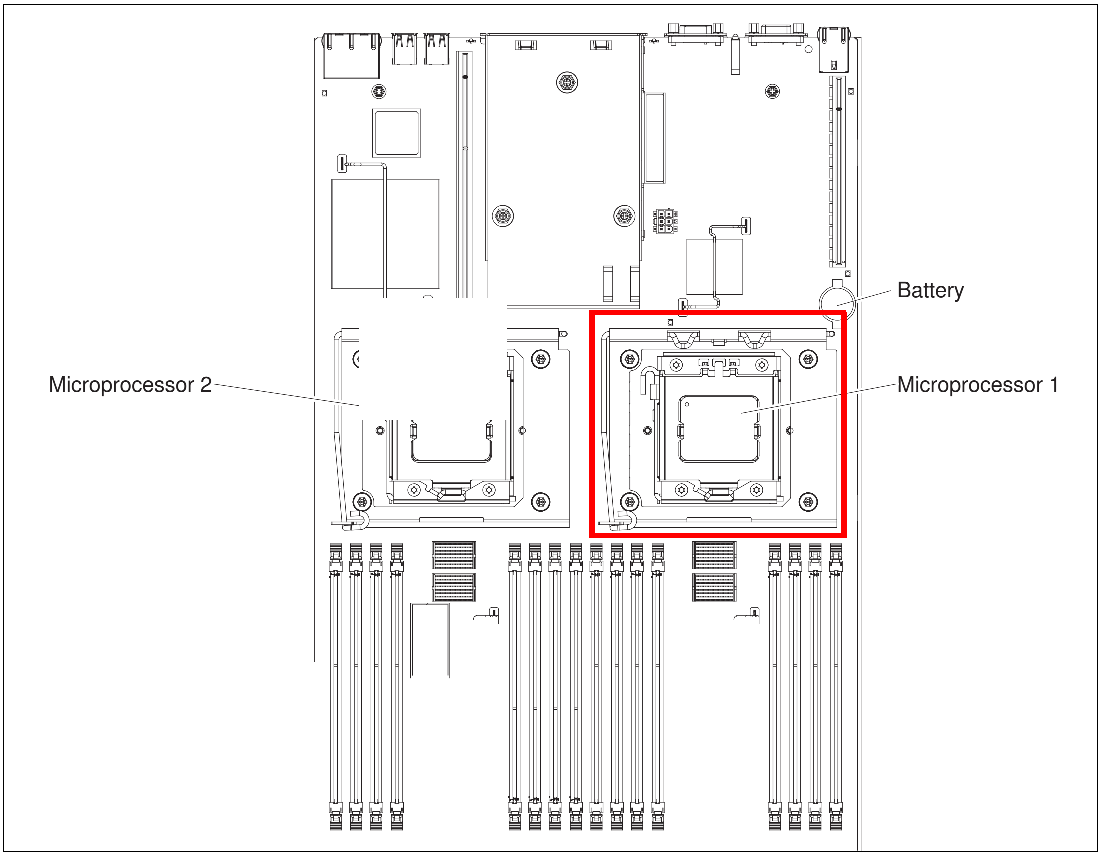


Figure 7-4 Processor locations in relation to surrounding components

7.2.2 Processor operating characteristics

Table 7-1 describes the operating characteristics of the server based on the number of processors and how the memory is installed. The table also describes how the server will react in the event of a failure in either processor 1 or 2.

Table 7-1 Operating characteristics of processor and memory installation options

Memory installed	Operational characteristics
Processor 1 is installed	
Memory installed only on the system board. Minimum of 2 DIMMs.	Performance improves as DIMMs are added to evenly populate all ranks on each memory controller.
Memory installed on both the system board and the mezzanine memory board.	The memory installed on the mezzanine memory board is not addressable by processor 1 and is ignored.
Memory installed on both the system board and the MAX5.	Performance is significantly improved when more active memory calls are using local memory on the system board.

Memory installed	Operational characteristics
Both processors 1 and 2 are installed	
Memory installed only on the system board (no memory mezzanine). Minimum of 2 DIMMs.	► Performance improves as DIMMs are added to evenly populate all ranks on each memory controller. ► Processing threads assigned to processor 2 will always have a significant drop in performance for memory-intensive tasks. ► Not an operational configuration for OSs, such as VMware.
Memory installed on both the system board and the mezzanine memory board.	► Performance improves as DIMMs are added to evenly populate all ranks on each memory controller. ► If processor 2 fails, memory on the memory mezzanine board is unusable. ► This configuration is only an operational configuration for VMware when both the system board and mezzanine memory board have the exact same total memory installed. For ease of maintenance, the best practice is to have an identical memory configuration on both the system and mezzanine memory board.
Memory installed on the system board, the mezzanine memory board, and the MAX5.	► Performance improves as DIMMs are added to evenly populate all ranks on each memory controller for local memory. ► Performance is significantly improved when more active memory calls are using local memory on the system board instead of the memory located in the MAX5. ► In the rare instance of a mezzanine failure, VMware operational requirements can be satisfied if all of the memory is installed in the MAX5. Memory access will be significantly slower if all of the memory is in the MAX5.

7.3 Memory considerations

Section 4.8, “Memory” on page 131 covers all of the various technical considerations regarding memory configuration for the System x3690 X5. There is a great deal of flexibility regarding the memory configuration of this server. As a result, there is a chance that you might configure a less than optimal memory environment.

Before you begin this section, understand how you are going to use the server. File servers that are used to provide access to disk storage for other servers or workstations are less affected by a less perfect memory latency speed than a database or mail server. Servers that are used as processing nodes in a high-performance cluster, database servers, or print servers for graphics printers require the best memory performance possible.

7.3.1 Local memory installation considerations

The following list describes considerations when installing memory inside the System x3690 X5:

- When installing memory for two processors and no mezzanine, only processing threads assigned to processor 2 will experience a 50% increase in memory latency. For a server with heavy I/O processing, this latency does not degrade the overall performance of the server.

- Any memory installed on the optional mezzanine board will not be seen by processor 1 when processor 2 is not installed or operational.
- For nonuniform memory access (NUMA)-aware OSs, when two processors are installed, you must install the same amount of memory on both the system board and the mezzanine. See Table 4-9 on page 136 for details.
- The best processor performance can be achieved when memory is installed to support *Hemisphere Mode*. When the mezzanine is not installed and both processors are installed, it is possible to have processor 1 in Hemisphere Mode and not processor 2. In this type of installation, having processor 1 in Hemisphere Mode will improve the memory access latency for both processors. To determine the DIMM population for Hemisphere Mode without the mezzanine board installed, see Table 4-9 on page 136.
- Consider installing all DIMMs of the same field-replaceable unit (FRU) to avoid conflicts that might prevent NUMA compliance or support Hemisphere Mode. The problems are most likely to occur when various-sized DIMMs are installed. The server supports various-sized DIMMs installed in each rank, but the configuration becomes complex and difficult to maintain. Although OSs that depend on NUMA compliance will inform you when the server is not NUMA compliant, there is nothing to inform you when the processors are not in Hemisphere Mode.
- It is better to install more smaller DIMMs than fewer larger DIMMs to ensure that all of the memory channels and buffers of each processor have access to the same amount of memory. This approach allows the processors to fully use their interleave algorithms to access memory faster by spreading it over multiple paths.

7.3.2 Testing the memory DIMMs

The best practice when installing memory is to run a memory quick test in diagnostics to ensure that all of the memory is functional. Memory might not be functional for the following reasons:

- Wrong DIMM for the type of server that you have. Ensure that only IBM-approved DIMMs are installed in your server.
- The DIMM is not fully installed. Ensure that the DIMM clips are in the locked position to prevent the DIMM from pulling out of its slot.
- The DIMM configuration is invalid. See the DIMM placement tables in “x3690 X5 memory population order” on page 133 for details.
- A possible bad DIMM, failed DIMM slot, bent processor pin, or resource conflict with a PCIe adapter. Swap the DIMM with a functional DIMM, reactivate the DIMM in F1-Setup, and retest:
  - When the problem follows the DIMM, replace the DIMM.
  - When the problem stays with the memory slot location, remove any non-IBM PCIe adapters, reactivate, and retest.
  - When the failed slot is on the mezzanine memory board, verify that the memory board is completely seated, reactivate, and retest.
  - Contact IBM support for parts replacement.

Use the built-in server diagnostics to test the installed memory. Use the following steps to perform a simple test of a new memory configuration before placing the server into production:

1. During POST, at the IBM System x splash panel, press F2-Diagnostics as shown in Figure 7-5 on page 308.

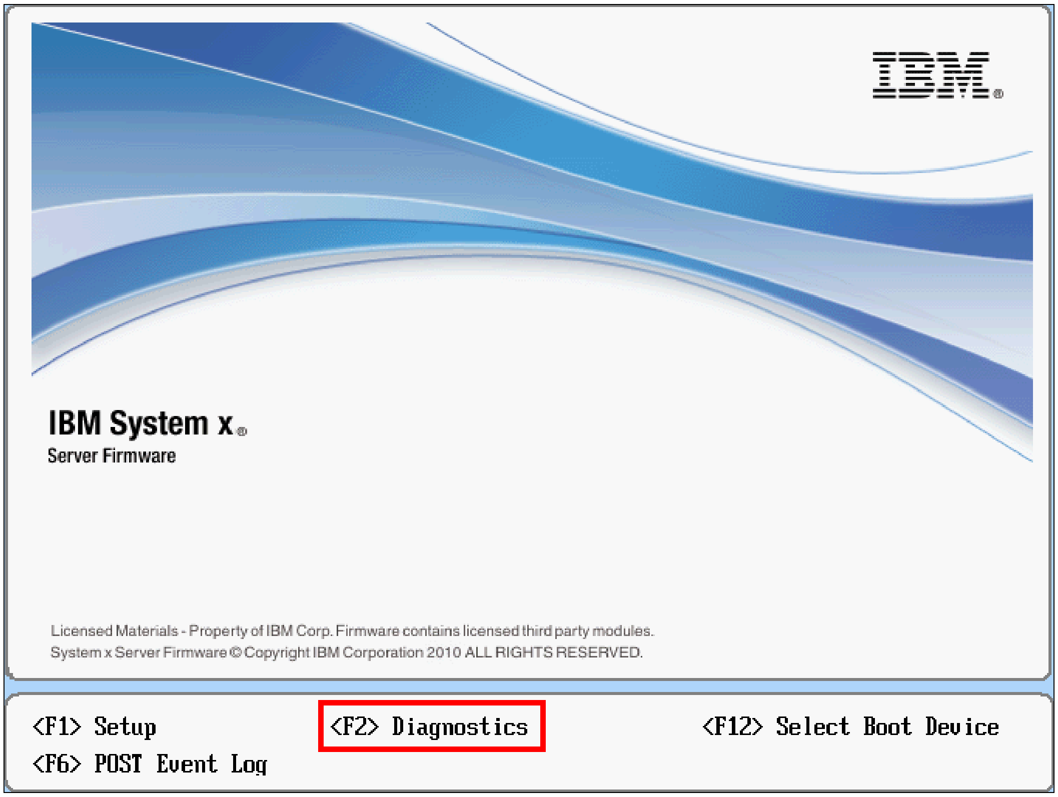


Figure 7-5 How to access diagnostics in POST

2. When the built-in diagnostics start, the diagnostics start the Quick Memory Test for all of the memory in the server, as shown in Figure 7-6 on page 309. You can stop the Quick Memory Test at any time and run a Full Memory Test, which runs the same test patterns multiple times and takes five times longer than the Quick Memory Test. The only time that you want to run the full memory test is if you have an intermittent memory problem that you are trying to isolate.

   Because the server identifies which specific DIMMs are experiencing excessive single-bit failures, it is far more efficient to swap reported DIMMs with similar DIMMs inside the server and see if the problem follows the DIMMs, stays with the memory slots, or simply goes away because the DIMMs were reseated.

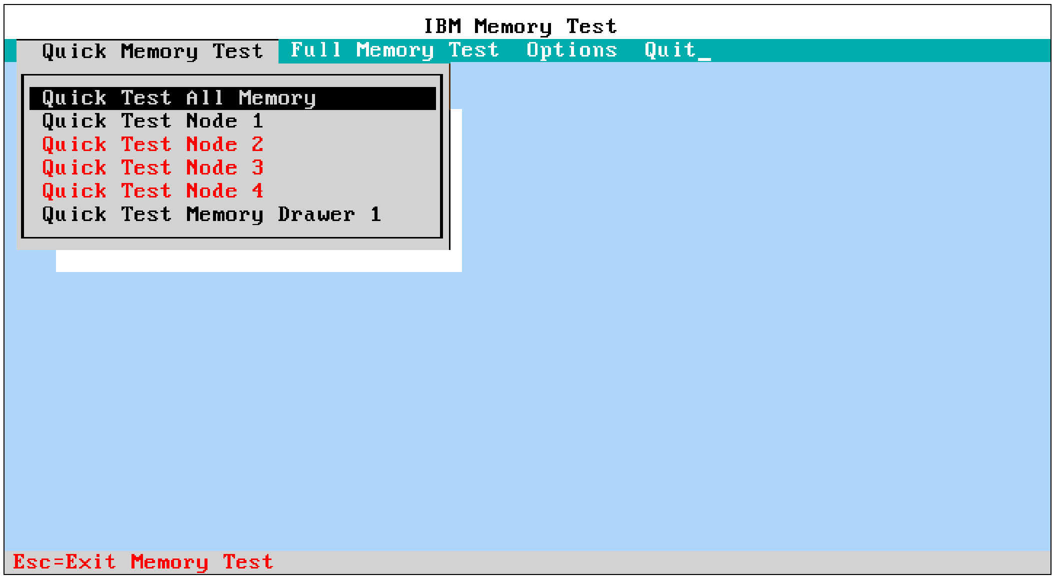


Figure 7-6 The start-up panel for the built-in diagnostics

3. The quick diagnostics continue to run, reporting each test that the quick diagnostics are currently performing and the length of time that it will take to complete that test. If an error occurs, the quick diagnostics stop and indicate the memory errors encountered before progressing into more advanced diagnostics, as shown in Figure 7-7.

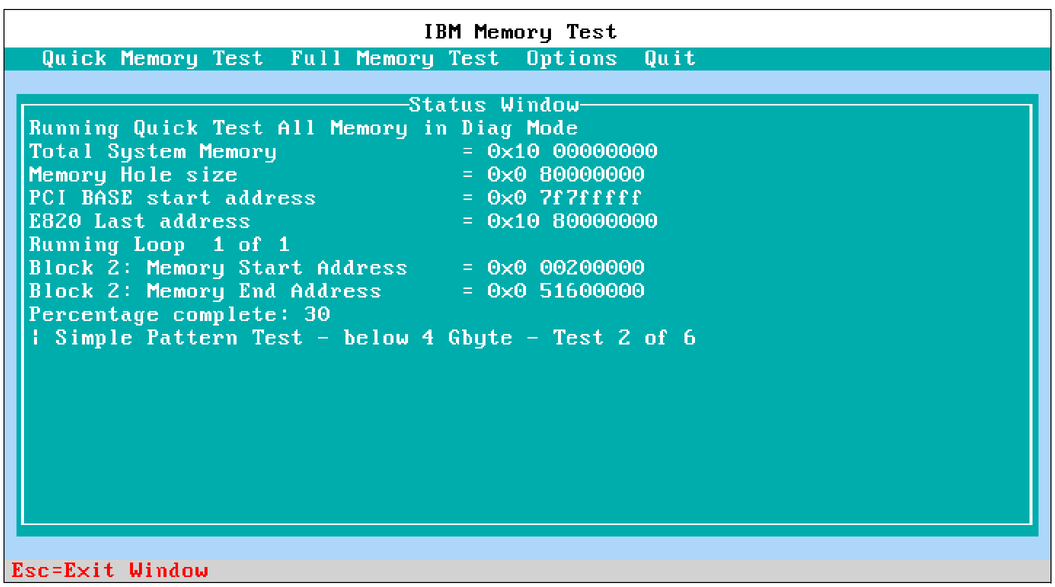


Figure 7-7 Quick Memory Test progress panel

You can terminate the diagnostics at any point by pressing Esc.

Important: Never warm-boot the server while running the built-in diagnostics. Several built-in functions that are used in the normal operation of the server are disabled during diagnostics to get direct results from the hardware. Only a normal exit from diagnostics or a cold boot of the server will re-enable those functions. Failure to perform this task correctly will cause the server to become unstable. To correct this problem, simply power off the server and power it back on.

7.3.3 Memory fault tolerance

For servers with high availability requirements, using the memory mirroring or memory sparing configuration allows the server to continue to function normally in the rare event of a memory failure. See 2.3.6, “Reliability, availability, and serviceability (RAS) features” on page 28 for an explanation of the memory mirroring and memory sparing functions.

Note, however, that considering that memory is like a solid-state device, it is unlikely that a DIMM failure will occur outside of the first 90 days of operation. Statistically, there is a higher risk for failure with power supplies, copper network adapters, storage devices, or processors. High availability is almost always a desired goal, but for true high availability, consider a cluster of host computers using common storage that allows virtual servers to be defined and moved from one host server to another host server.

If high availability is the most important aspect of your server (with cost and performance as secondary concerns), enable memory mirroring or memory sparing.

Use the following steps to establish memory mirroring or memory sparing:

1. Boot the server into F1-Setup.
2. From the System Configuration panel, select **System Settings** → **Memory** (Figure 7-8).

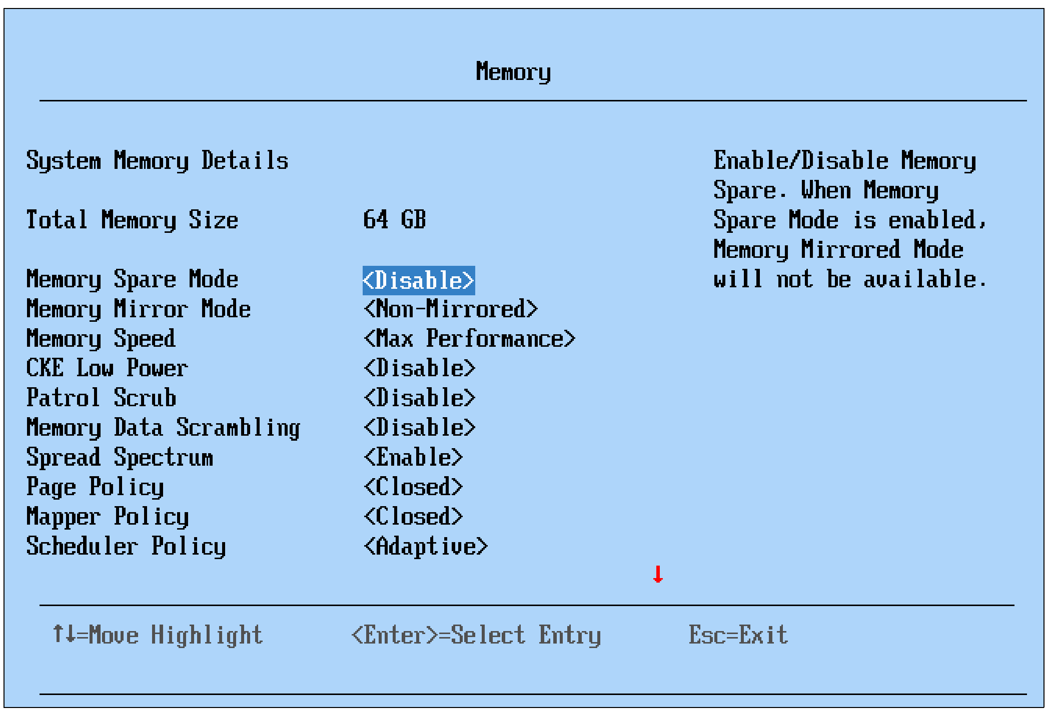


Figure 7-8 Memory configuration panel in F1-Setup

Figure 7-8 shows the Memory configuration panel with selections for performing memory sparing or memory mirroring, but not both at the same time. Select the desired option and reboot the server. If your memory population order does not support the requested option, the server will report a memory configuration error during the next reboot.

See 4.8.2, “x3690 X5 memory population order” on page 133 for the correct memory population order to support memory mirroring or sparing.

7.4 MAX5 considerations

On top of the half of a terabyte (TB) of memory that can be configured in the x3690 X5, an additional half of a TB of memory can also be configured in the MAX5 memory expansion unit and attached to the server to increase overall memory access performance and capacity.

MAX5 is best used with applications that can benefit from the increase in overall memory capacity. The most significant performance gains come in applications, which *require* the additional memory that MAX5 can provide. If an application does not need the extra memory, the potential for performance gains is reduced.

The best way to populate memory DIMMs in the server and the MAX5 depends greatly on the manner in which the applications that are being run address memory and the total amount of memory that is needed by the applications. However, you also need to consider the following rules:

- Always populate the DIMM sockets in the server first before installing DIMMs in the MAX5. Local memory has higher bandwidth and lower latency than MAX5 memory. MAX5 memory is limited by the speed of the QuickPath Interconnect (QPI) link.
- If possible, install DIMMs so that Hemisphere Mode is enabled. Without Hemisphere Mode, performance can suffer considerably.
- The MAX5 can be used on an x3690 X5 that has only one processor installed. However, you get the best performance by having both processors installed and installing memory to both the system planar and the memory mezzanine.
- MAX5 adds an additional path to memory through dedicated QPI ports, which results in potentially greater memory bandwidth. There might be instances where it is better to have reserve DIMMs for use in the MAX5.

7.4.1 Before you attach the MAX5

Before you can attach and use the MAX5, the x3690 X5 requires that the firmware, which is shown in Table 7-2, is in place. Perform firmware updates in the order listed.

Table 7-2 Minimum firmware levels to support the MAX5 memory expansion

Firmware type	Version	Build
1. IMM	1.20	YUOO75T
2. UEFI	1.21	MLE120B
3. FPGA	1.10	MLDU20EUS

Field Programmable Gate Array (FPGA) on the MAX5: The MAX5 also has FPGA firmware that resides on it. The FPGA firmware is updated automatically when the server to which it is attached has its FPGA firmware flashed. The default FPGA loaded onto the MAX5 from the factory is for the System x3850 X5. You are required to reflash the FPGA on your system x3690 X5 after the MAX5 is connected to it.

The server might fail to complete POST because of a significant difference in FPGA code between the MAX5 and the server. Correct this problem by flashing the FPGA through the IMM while the server is plugged into power but not powered on.

The recommended sequence of updates is shown in Table 7-2 on page 311. To update the server to this firmware level or a higher level, you can use one of the following choices:

- When you have a compatible OS installed, you can use the UpdateXpress System Packs Installer (UXSPI) tool that is described in 9.11, "UpdateXpress System Pack Installer" on page 511.
- Regardless of the version of OS or when an OS is not installed, you can use the Bootable Media Creator (BOMC) tool that is described in 9.12, "Bootable Media Creator" on page 514.
- After establishing a network connection and logging in to the IMM, use the firmware update function, as described in 9.13, "MegaRAID Storage Manager" on page 521.

> **UEFI update is a two-step process:** The UEFI flash is a two-step process. When the UEFI flash has completed, the server must be warm-booted to at least the F1-Setup panel to allow the second half of the UEFI flash to complete. Failure to perform the required warm boot will result in a corrupt version of the UEFI, which will force the server to boot into the recovery page of the UEFI until corrected by following the UEFI recovery process in the *Problem Determination and Service Guide* (PDSG).

For the latest firmware requirements using the MAX5, see RETAIN tip H197572:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5085756

7.4.2 Installing in a rack

The QPI cables that are used to connect the MAX5 to the x3690 X5 are extremely short, stiff cables that can be damaged easily. For this reason, that hardware ships with the MAX5 to allow the MAX5 to attach to the x3690 through a series of brackets and rail kits. Install the MAX5 and x3690 X5 in the rack before cabling them together.

The product publication *IBM eX5 MAX5 to x3690 X5 QPI cabling kit installation instructions* documents the process to connect the MAX5 to the server. This document is available at this website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5085207

7.4.3 MAX5 cables

The QPI cables that are used to cable the x3690 X5 and the MAX5 are extremely short and stiff. To plug the cables in, you must start the cable insertion on both sides of the cable into the correct receptors on both the x3690 X5 and MAX5 at the same time.

Use the following tips when plugging in the cables:

- QPI cables ship packaged with reusable plastic boots that protect the fragile outside edges of the cable connectors, as shown in Figure 7-9 on page 313. It is a good idea to keep a set of these plastic boots available for times when you want to remove the QPI cables when moving equipment from rack to rack or when servicing the unit.

Figure 7-9 Reusable QPI cable connector protective boot

- There are only two QPI cables connecting the two QPI ports of the x3690 X5 to two of the four QPI ports on the MAX5. Figure 7-10 shows how the cables must be connected between the QPI ports of the two units.

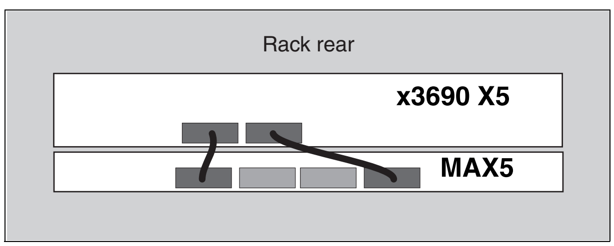


Figure 7-10 QPI cable installation

- The QPI cables are keyed to only be inserted one way. A quick visual check for cable orientation is to look for the 2U QPI or 1U QPI labels on the cable. The labels, along with the blue retainer release tab, are placed on what will become the visible top of the cables when they are installed correctly.
- The ends of the cables are labeled to indicate which end to insert into the correct equipment. The 2U QPI end of the cable plugs into the x3690 X5. The 1U QPI end of the cable plugs into the MAX5.
- The cable end must slide into the port until it clicks into place. You can disengage the retainer that holds the cable in place by pressing on the blue tab on the top of the QPI cable connection.
- Both cables must be installed, even when only one processor is installed, to allow the MAX5 to be controlled by the server. If one of the cables is detached, the server will not power on or complete POST.

Updating FPGA firmware: When attaching the MAX5 for the first time, you must reapply the FPGA firmware using the IMM firmware update tool after the server is plugged into ac power but prior to powering up the server.

Do not use Bootable Media Creator or USXPI to update the FPGA until the FPGA firmware is a match between the server and the MAX5. Mismatched FPGA will make the server unstable and it might power off during the flash, corrupting the FPGA. *This situation will result in requiring a hardware replacement, because there is no recovery for corrupt FPGA.*

Figure 7-11 shows the back of the x3690 X5 with the MAX5 attached.

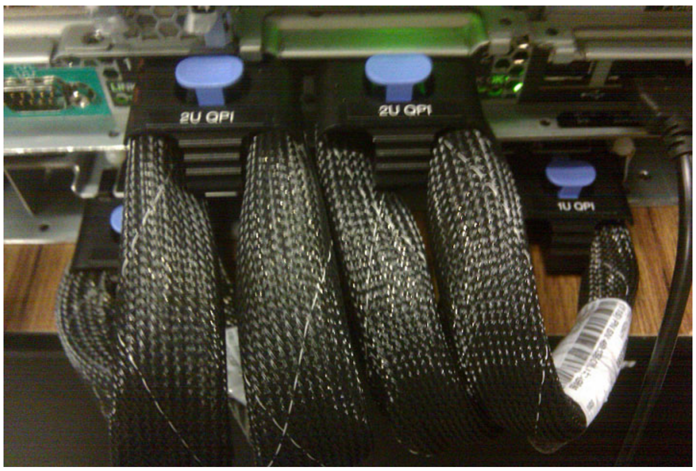


Figure 7-11 Lab photo of an x3690 X5 (top) attached to an eX5 MAX5 (bottom)

Important: With the server or MAX5 plugged into ac power and the scaled unit's power turned off, the QPI cables still have active dc power running through them. *You must only plug in or unplug the QPI cable when the MAX5 and the server are not plugged into ac power.* Failure to follow this rule will result in damaged circuits on either the MAX5 memory board or the system board of the server.

7.4.4 Accessing the DIMMs in the MAX5

To access the memory DIMMs in the MAX5, slide the memory board of the MAX5 forward out the chassis of the MAX5. Use the following steps to access the DIMMs:

1. Remove ac power from all of the server's power supplies and from the two MAX5 power supplies. Because the QPI cables are already held in alignment by the memory expansion chassis, it is not a requirement to remove the QPI cables before removing the memory board.
2. Remove the front bezel of the MAX5 by pressing in on the tab buttons on both sides of the bezel. The bezel then can be pulled away from the MAX5.
3. As shown in Figure 7-12 on page 315, there are two blue release tabs at the front of the MAX5. When pressed to the sides of the enclosure, they allow you to pull out the cam levers that are used to begin to pull out the memory tray.

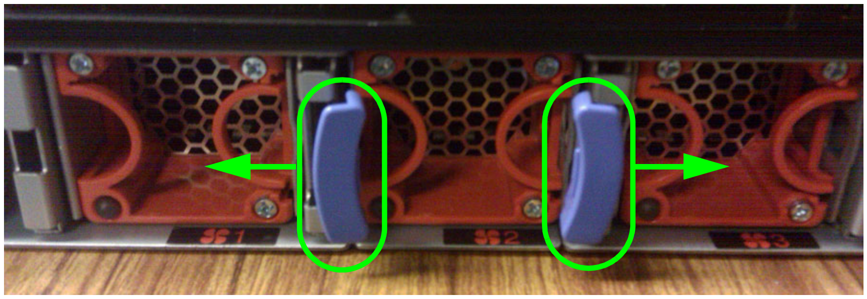


Figure 7-12   Blue release tabs for the memory tray cam levers

4. You can then pull the memory tray out about 30% before it stops. This design allows you to get a better grip on the tray on either side and then use your fingers to push in another set of blue release tabs on either side of the tray, as shown in Figure 7-13.

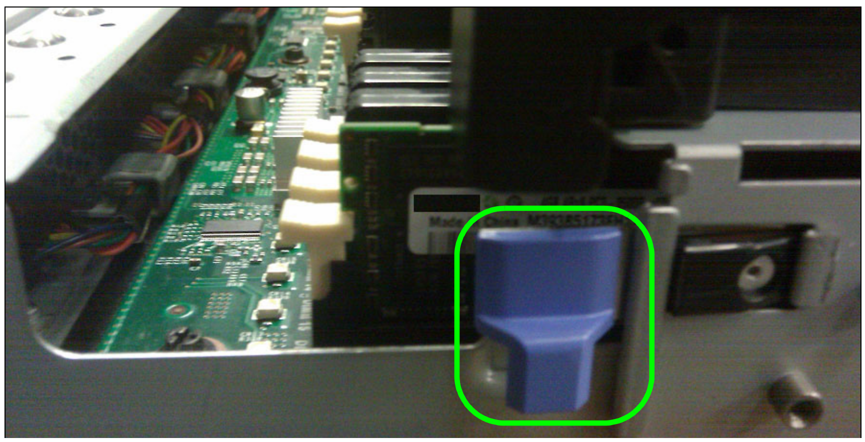


Figure 7-13   Final release tab for removing the MAX5 memory tray

5. Slide the memory tray completely out and place it on a flat work surface to work on the memory.

> **Power:** You must remove the ac power from both the server and the attached MAX5 before removing the memory tray. The FPGAs of both the server and the MAX5 are still active components when the server is powered down but not removed from utility power. FPGA components exist on both the server and the MAX5. Removing the memory board with ac power still active damages the FPGA components of both the server and the MAX5.

For the memory population order, see 4.8, “Memory” on page 131.

After the MAX5 is installed and configured properly, you can confirm a successful link between the server and the MAX5 when both units power on and off as one complete unit and the memory in the MAX5 can be seen by the server. You will also see the following message during POST:

“`System initializing memory with MAX5 memory scaling`”

7.5 PCIe adapters and riser card options

This section describes considerations for determining how to use your PCIe slots, depending on the types of PCIe riser cards that you have installed. The x3690 X5 is designed to function in a wide variety of tasks, starting from a high-performance graphics workstation up to a storage or processing node in a high-performance cluster.

7.5.1 Generation 2 and Generation 1 PCIe adapters

All of the PCIe slots in the x3690 X5 are at the Generation 2 (*Gen2*) specification. Gen2 offers additional error correction and addressing advancements, so that all of the slots of this server exchange data twice as fast as servers with Gen 1 PCIe slots.

Table 7-3 describes the theoretical limits of each of the common types of PCIe adapters.

Remember that theoretical limits are based on the mathematics of the frequency and data width of the bits that are transmitted over the interface. Theoretical limits do not take into account the communications needed to maintain the protocols required to interface between intelligent devices. Theoretical limits also do not take into account the inability to maintain a steady flow of data in full duplex.

Table 7-3 Theoretical data transfer limits of Gen1 PCIe slot types versus Gen2 PCIe slot types

PCIe slot type	Generation 1 limit	Generation 2 limit
x1	500 MBps	1 GBps
x4	2 GBps	4 GBps
x8	4 GBs	8 GBps
x16	8 GBps	16 GBps
x32	16 GBps	32 GBps

PCIe adapters connect to the processors via the I/O hub. The purpose of the I/O hub is to combine the various data streams from each of the PCIe slots into a single aggregate link to the processors by using a dedicated QPI link to each processor. The data transfer rate of the QPI link is negotiated between the processor and the I/O hub.

Table 4-5 on page 130 shows the QPI link speeds based on the types of installed processors. The I/O hub supports the highest QPI link speed on the table, which is 6.4 gigatransfers per second (GT/s). You can adjust the QPI link speed to conserve power by booting into F1-Setup, then selecting **System Settings** → **Operating Modes**, and setting QPI Link Frequency to values other than the default *Max Performance*, as shown in Figure 7-14 on page 317.

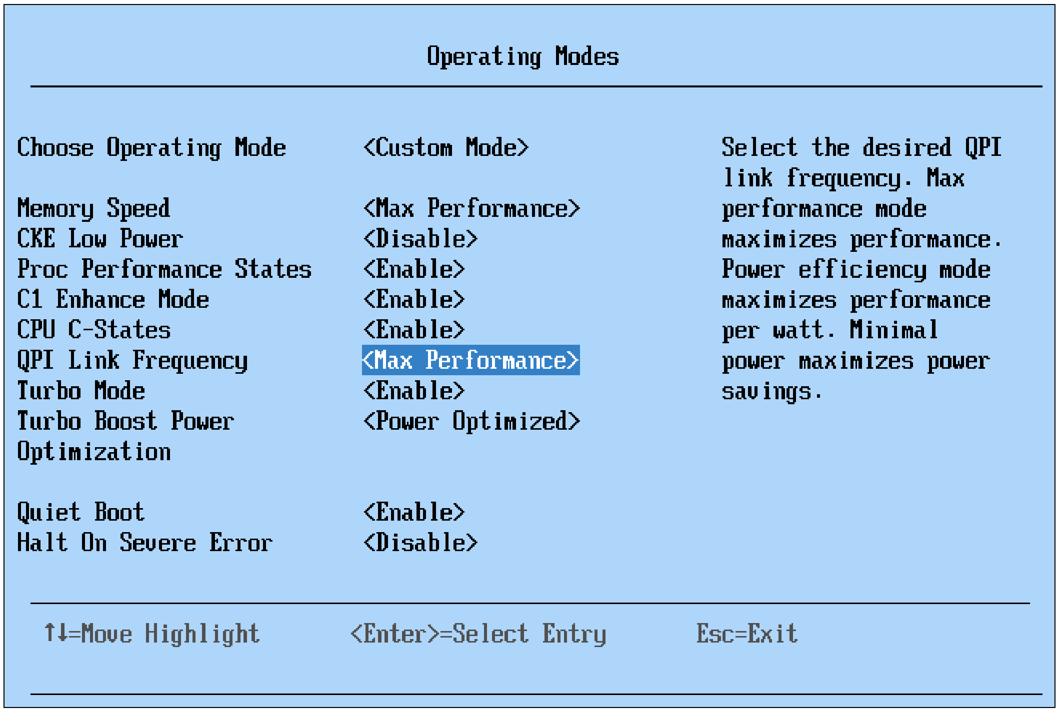


Figure 7-14 QPI Link Frequency setting

Backward compatibility of Gen2 PCIe slots to Gen1 adapters

Although all Gen2 PCIe slots are backward compatible to Gen1 adapters, not all PCIe adapter vendors adopted the optional specification of Gen1 that allows a Gen2 PCIe slot to recognize a Gen1 adapter. When you install a Gen1 PCI adapter that is not recognized by the server, consider forcing the PCI slot in which the adapter is installed to a Gen1 slot in F1-Setup by selecting **System Settings** → **Devices and I/O Ports** → **PCIe Gen1/Gen2 Speed Selection**. Figure 7-15 on page 318 shows the resulting panel and the available selections. The change takes effect after a cold reboot of the server.

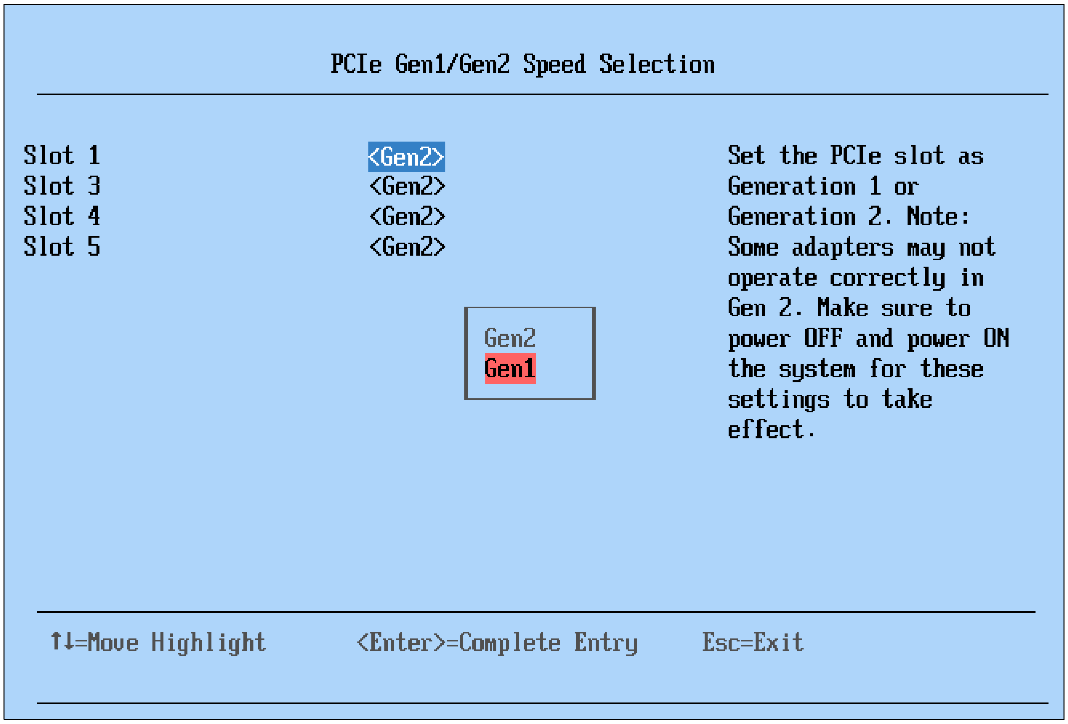


Figure 7-15 PCIe slot speed selection panel to force Gen1 compliance

Non-UEFI adapters in an UEFI environment

A number of Gen1 PCIe adapters were designed prior to the implementation of UEFI. As a result, these adapters are not recognized or might not have UEFI drivers that allow the adapter to function in an UEFI environment.

The server supports these non-UEFI adapters via a setting Legacy Thunk Support, which is enabled by default. Legacy Thunk Support mode places the non-UEFI-aware Gen1 adapter into a generic UEFI wrapper and driver, which allows you to update the firmware of the adapter to support UEFI.

We recommend that all installed adapters either support UEFI as standard or be updated to support UEFI, because Thunking only provides limited support for non-UEFI adapters in an UEFI environment. For example, a legacy adapter in a Thunk UEFI wrapper cannot be seen in **System Settings** → **UEFI Adapters and Device Drivers**, nor can it natively access memory locations above 4 GB.

If you have previously disabled Thunking, you can re-enable it by using F1-Setup and selecting **System Settings** → **Legacy Support**. Figure 7-16 on page 319 shows the Legacy Support panel.

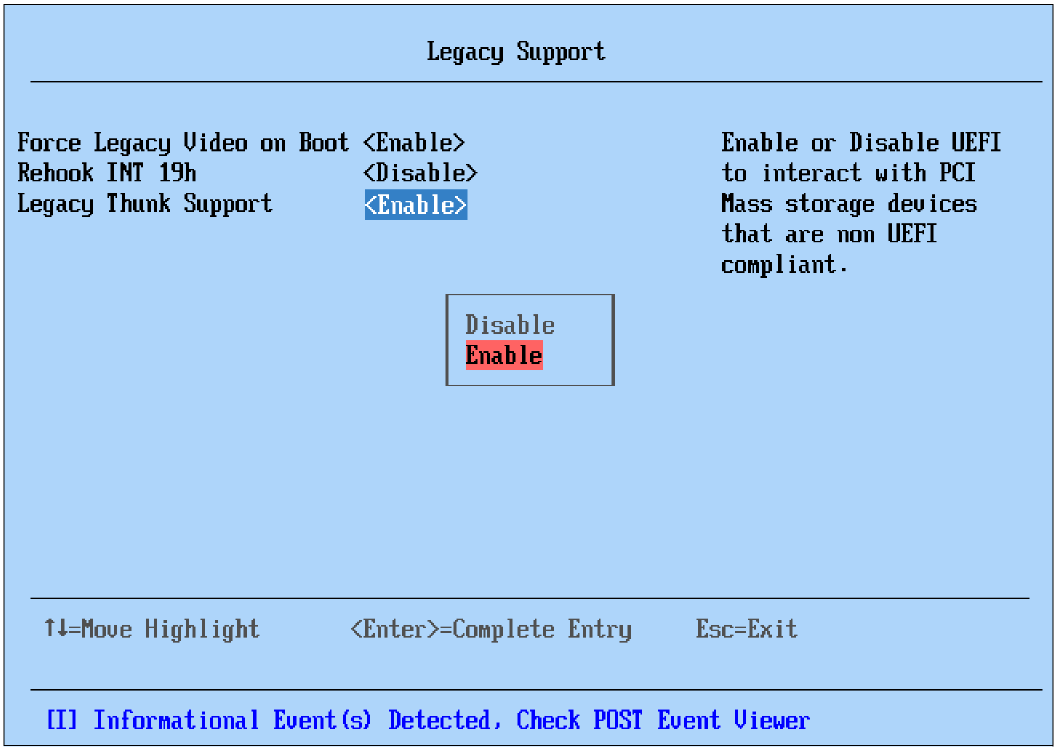


Figure 7-16 Legacy Thunk Support in UEFI

Another possible solution to this problem is booting the server in Legacy Only mode. This mode allows both non-UEFI-aware OSs and PCIe adapters to function as though they are on a non-UEFI server. Many of the advanced memory addressing features of the UEFI environment will not be available to the OS in this mode.

Booting the server to an OS in this mode allows you the ability to apply firmware updates to a Gen1 adapter that it not recognized in an UEFI environment.

To enable this feature from within F1-Setup, select **Boot Manager** → **Add Boot Option**, scroll down until Legacy Only is displayed in the list of options, and select it, as shown in Figure 7-17 on page 320. If Legacy Only is not listed, it has already been added to the boot manager.

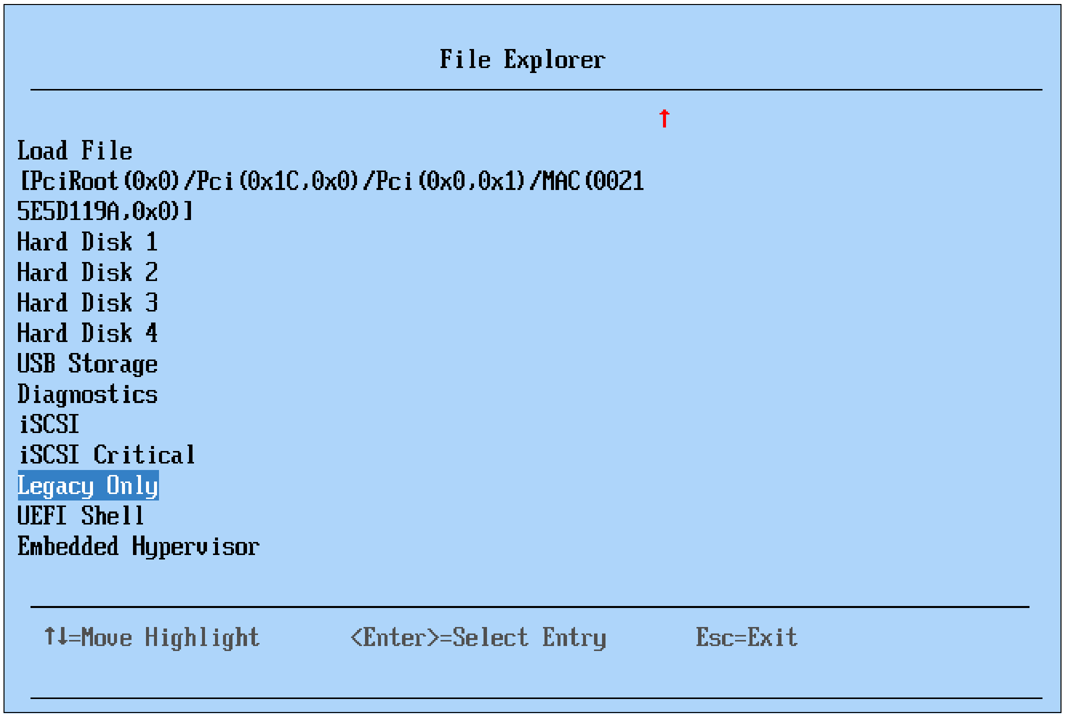


Figure 7-17 Selecting Legacy Only for a boot option

If you add Legacy Only to the boot manager, you must also change the boot sequence to place Legacy Only at the top of the boot sequence. To change the boot order from within the Boot Manager panel, select **Change Boot Order**. Figure 7-18 shows the Change Boot Order panel. Use the arrow keys to select **Legacy Only** and move it to the top of the list. Press Enter to confirm your changes.

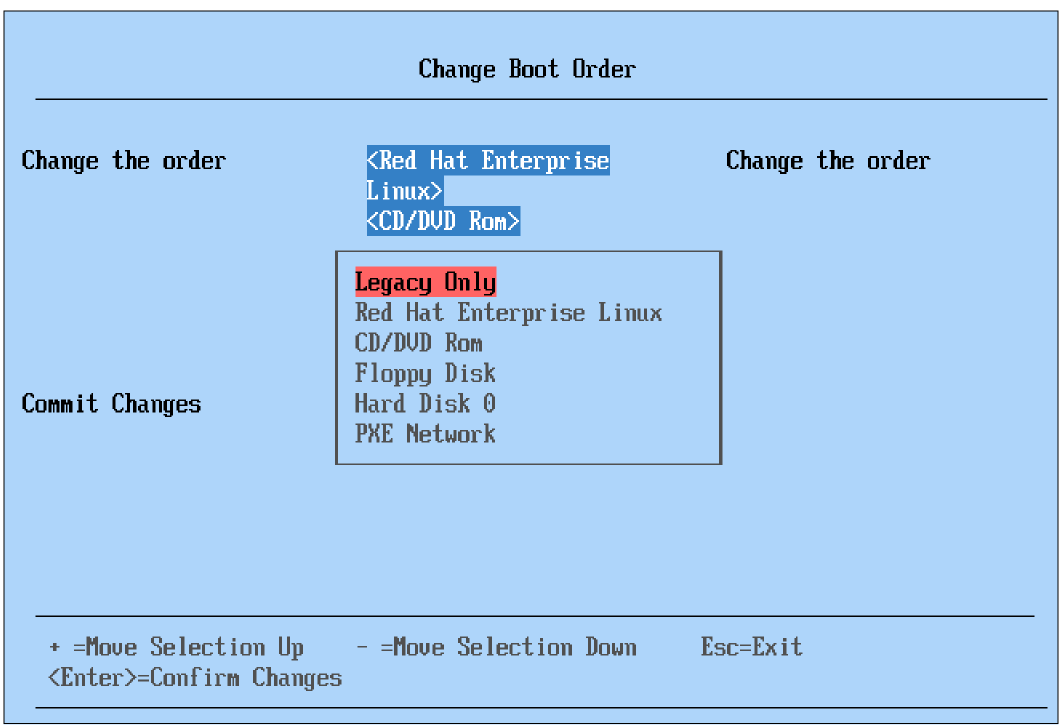


Figure 7-18 Change Boot Order panel

Microsoft Windows 2008 x64: When installed on an UEFI server, Microsoft Windows 2008 x64 will install Microsoft Boot Manager as part of the boot sequence. Regardless of how you change the boot sequence in the boot manager, Microsoft Boot Manager will always be at the top of the sequence.

When you install this same OS with Legacy Only enabled, Microsoft Boot Manager is not installed as part of the boot manager. Removing the Legacy Only option from the boot manager will prevent the server from booting into the installed Windows 2008 x64.

7.5.2 PCIe adapters: Slot selection

The x3690 X5 offers a wide variety of PCIe adapter configurations among possible riser card combinations, as described in 4.10, “PCIe slots” on page 164. The two riser slots have these riser card options:

- Riser slot 1 can contain one of the following options:
  - Two x8 slots (installed in most standard models)
  - One x16 slot for 3/4-length cards
  - One x16 slot for full-length cards (memory mezzanine cannot be used)
- Riser slot 2:
  - Three x8 slots (one slot is wired as an x4 slot) (installed in most standard models)

The server can be ordered without riser card 1 or with any of the three possible riser cards. Riser card 2 comes standard in most models of the server, as described in 4.10, “PCIe slots” on page 164.

Note these key points about the riser cards:

- Riser card 1 supports a single x16 adapter, such as a video card. If needed, an auxiliary power connector is located on the system board near the riser card.
- Two x16 riser cards are available for riser slot 1: one for full-length cards and one for 3/4-length cards. If you use the full-length card riser, you cannot also install the memory mezzanine, because the two options do not both physically fit in the server.
- As described in 4.9.5, “ServeRAID Expansion Adapter” on page 157, the ServeRAID Expansion Adapter is a SAS expander that allows you to create RAID arrays of up to 16 drives and across up to four backplanes. The Expansion Adapter must be installed in PCI Slot 1 (in riser card 1) and the ServeRAID adapter must be installed in PCI Slot 3 (in riser card 2). See Figure 4-44 on page 164 for the locations of these slots.
- Slot 4 of riser card 2 is physically an x8 slot, but electronically, it is an x4 slot.
- If the Emulex 10Gb Ethernet Adapter is installed (it is standard in certain models, as listed in 4.3, “Models” on page 124), the adapter is installed in slot 5 in riser card 2. See 4.10.3, “Emulex 10Gb Ethernet Adapter” on page 166 for details about the adapter.

Although true performance on a given PCIe adapter can largely depend on the configuration of the environment in which it is used, general performance considerations exist with respect to the x3690 X5 server.

The I/O hub supports 36 lanes of PCIe traffic with a combined bandwidth of 36 GBs. Each processor’s QPI link to the I/O hub is capable of a maximum throughput of 26 GBs, depending on the processors installed in this server. With only one processor installed in this server, the maximum combined bandwidth of all the PCIe lanes is reduced to the maximum bandwidth of a single QPI link.

Of all of the I/O adapters that can be installed on the server, the ServeRAID and 6 Gbps SAS controllers, when managing solid-state drives (SSDs), are the only adapters that can approach the theoretical limits of the x8 PCIe 2.0 slot. When you use SSD drives, connect no more than 8 SSD drives to a single controller for the best performance of the SSD drives. Only use x8 slots to host the controllers that manage your SSD drives (that is, not slot 4, because it is an x4 slot).

A single ServeRAID controller managing a single 4-drive SAS hard disk drive (HDD) array will function within the theoretical limits of an x4 PCIe slot. In this case, the mechanical nature of the HDD drives will limit the maximum throughput of data that passes through the PCIe slot.

The dual-port 8 Gbs Fibre Channel, 10Gbs Ethernet, and 10 Gbs Converge Network Adapters (CNAs) are all capable of approaching the theoretical limits of an x4 PCIe 2.0 slot and might perform better in an x8 PCIe 2.0 slot.

7.5.3 Cleaning up the boot sequence

One of the most overlooked steps in completing a hardware setup is deciding from what you are going to boot. The x3690 X5 server might have one or more ServeRAID controllers for internal drives or perhaps another ServeRAID adapter for external drives. You might also use one or more Fibre Channel host bus adapters (HBAs) to access a SAN and you might have Preboot eXecution Environment (PXE) or iSCSI defined to boot to an OS over the network.

By default, your server and the installed options came with the ability to boot from any of these sources besides USB storage devices. On every boot, the server recognizes each of these boot choices, determines if bootable media is attached, and adds the optional ROM support to the boot ROM to determine the correct device from which to boot. Checking all of these boot choices adds time to the boot process.

To minimize this loss of time, you can disable the boot options for adapters that you know you are never going to boot from. The following sections describe the common methods of disabling boot options.

Legacy only mode

When the server is instructed to boot to *Legacy Only* mode, the best way to disable unwanted boot sequences is to disable them in F1-Setup by selecting **System Settings → Device and I/O Ports → Enable / Disable Legacy Option ROM(s)**. Figure 7-19 on page 323 shows the available options. You need to know the specific PCIe slots that are used for each adapter so that you will know which slot to leave enabled.

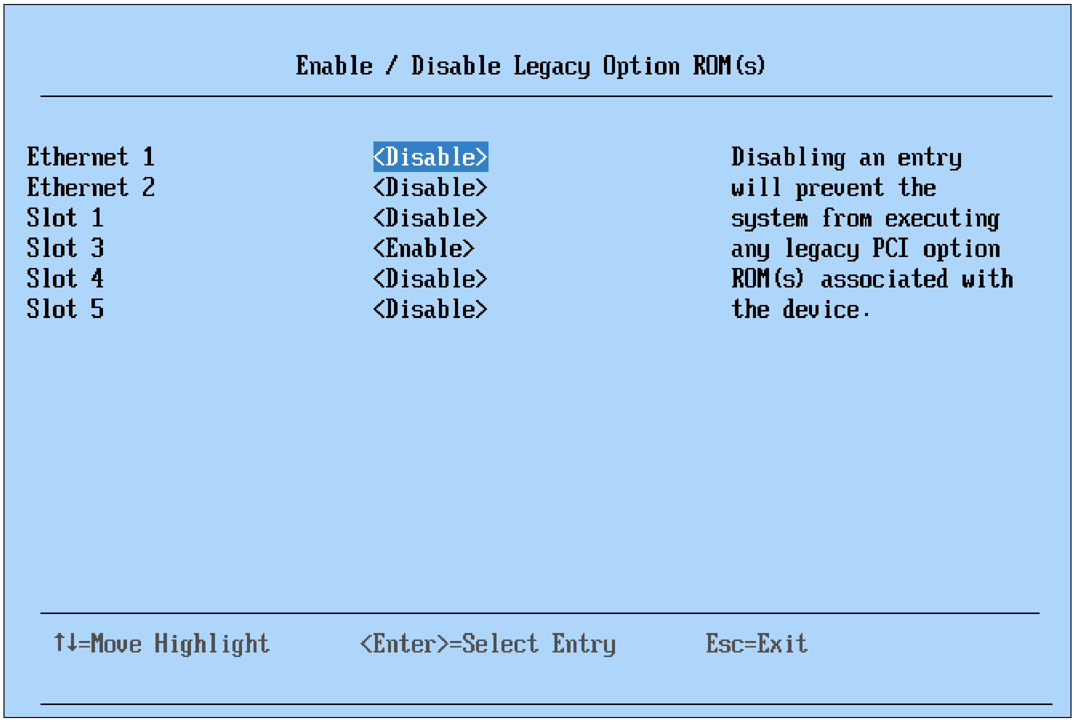


Figure 7-19 Legacy Option ROM states

When booting from SAN with multiple paths for redundancy, you need to enable the legacy option ROM for both HBAs.

The default UEFI mode

On the x3690 X5, you can sequence the order in which the UEFI will search the various attached devices to locate a boot device. You can shorten the time that it takes to perform the search by moving the adapter that contains the boot device to the top of the list.

In UEFI mode, PXE boot can be disabled for the onboard network interface card (NIC) ports through F1-Setup by selecting **System Settings** → **Network** → **PXE Configuration** and then by selecting the port from which you want to disable PXE boot. Figure 7-20 on page 324 shows the panel that you use to disable PXE boot on one of the two onboard network ports.

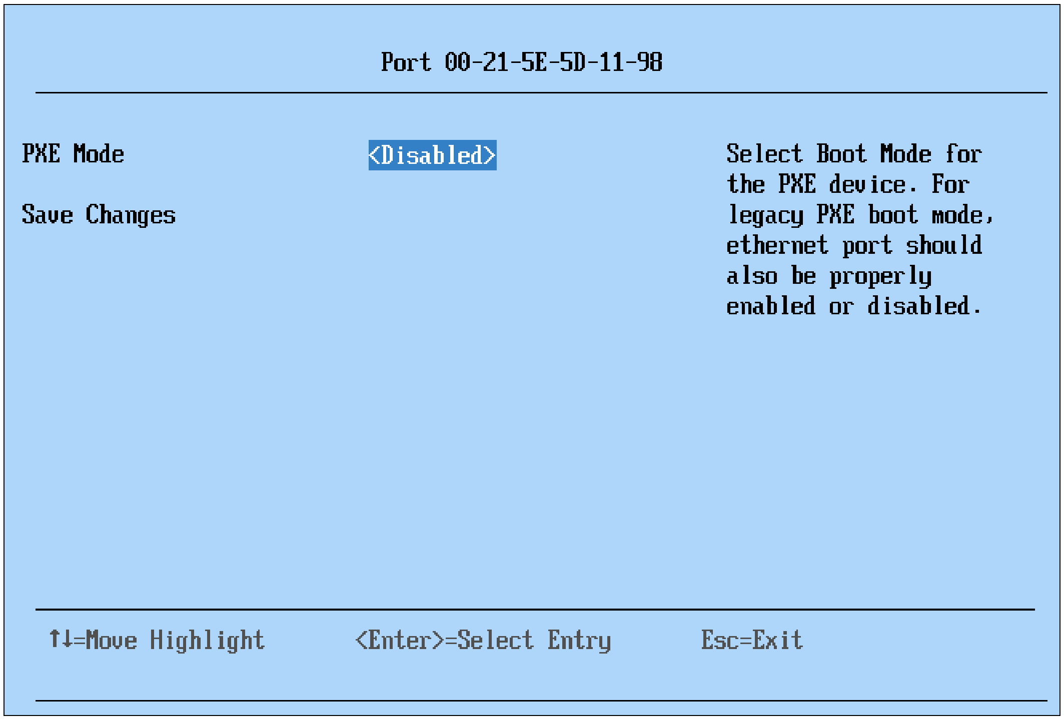


Figure 7-20 Disabling PXE boot of the onboard network ports

Other PCI adapters can have their boot option ROM disabled from within their configuration panels. To access individual adapter configuration panels from F1-Setup, select **System Settings** → **Adapters and UEFI Drivers**. Figure 7-21 shows the available selections on this panel.

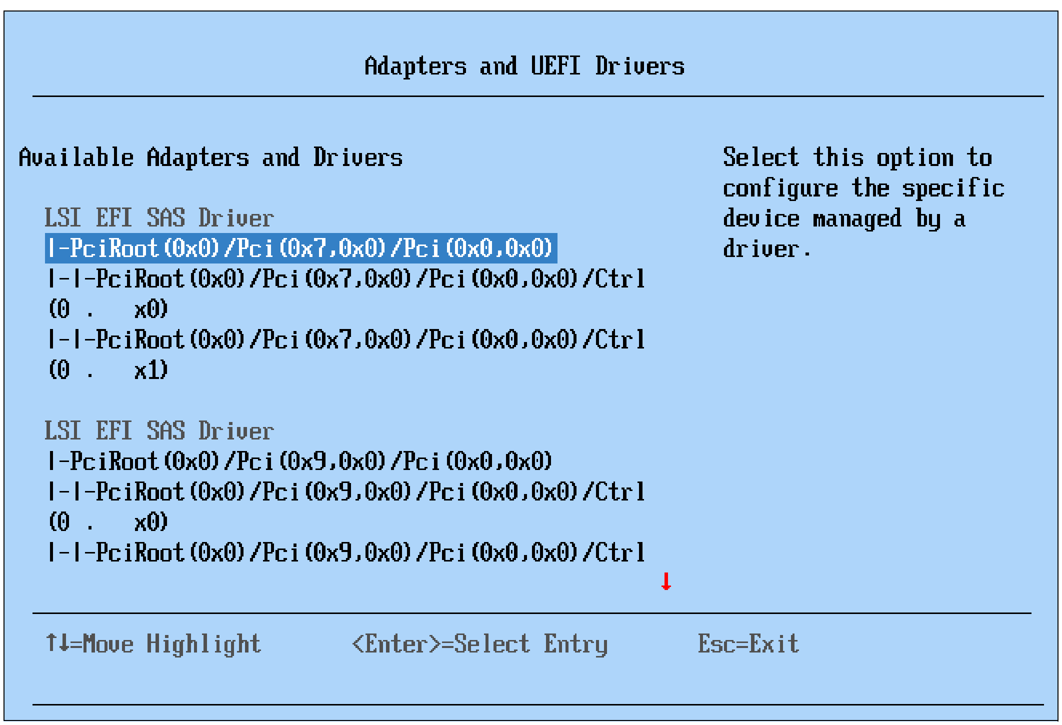


Figure 7-21 Accessing adapter-specific configuration information

To enter the configuration of a specific adapter, select the PciROOT directly beneath the adapter name. When you have multiple controllers of the same type, selecting the PciROOT of any of the same adapter types will select all of them and present a panel that allows you to select the specific adapter from within the configuration routine of the adapter type. For example, Figure 7-22 demonstrates this process when two ServeRAID adapters are installed in the server.

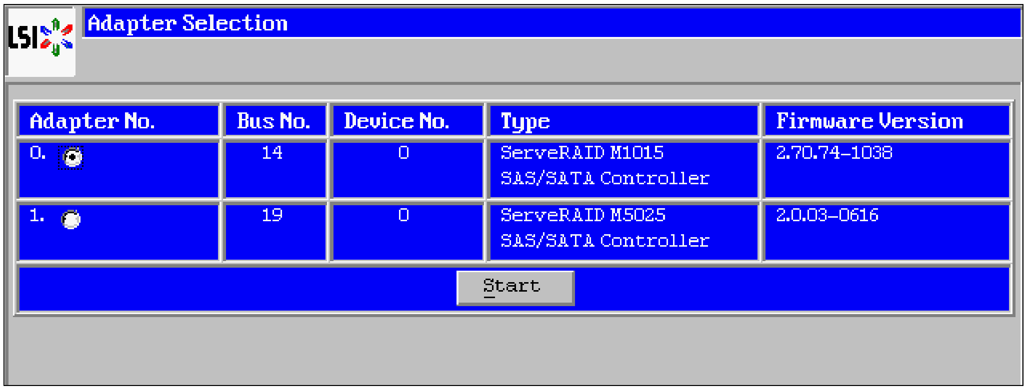


Figure 7-22 LSI Adapter Selection panel from within the LSI configuration window

For adapters that are similar to the ServeRAID or Fibre Channel HBAs, look for the controller configuration or properties. Inside the controller properties for the ServeRAID controllers, you see a panel similar to the panel that is shown in Figure 7-23.

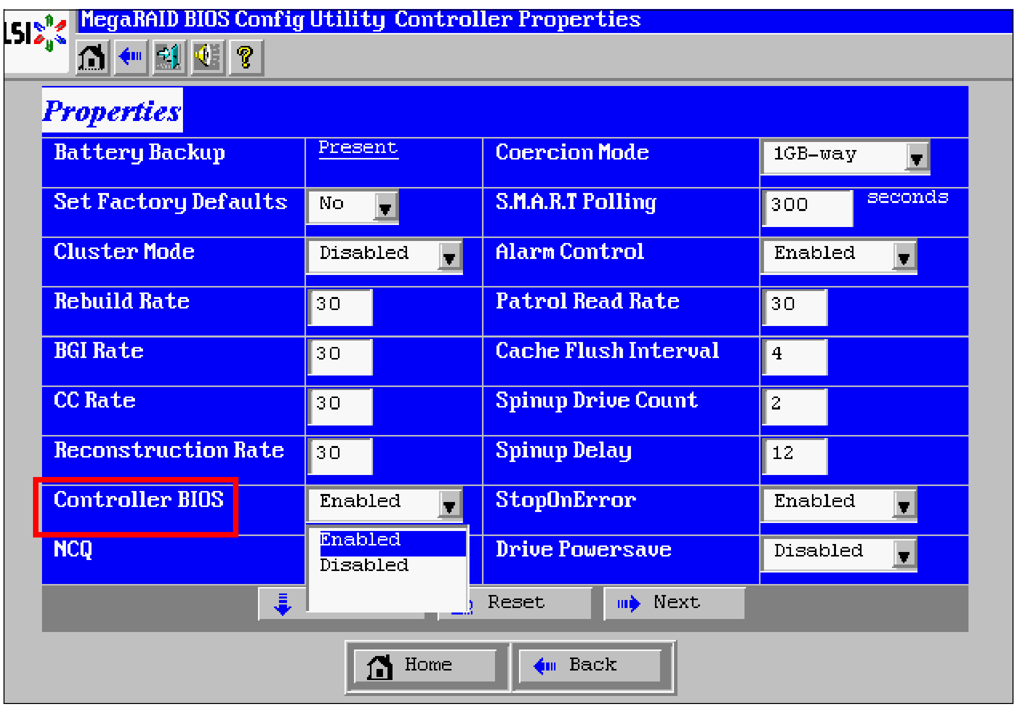


Figure 7-23 ServeRAID BIOS Configuration Utility Controller Properties: Disabling Boot ROM

Although not obvious by the description, disabling the Controller BIOS only disables the Boot ROM execution during POST. All of the other operating characteristics of the adapter BIOS and firmware remain intact.

ROM execution order

Regardless of whether you are booting in Legacy Only mode or UEFI mode, you can control from which device you want to boot. This capability is important when multiple storage adapters are installed in the server.

To control the boot sequence from within F1-Setup, select **System Settings → Device and I/O Protest Option ROM Execution Order**. Figure 7-24 shows what the panel looks like after pressing Enter on the list of possible choices. Use the up and down arrow keys to select a specific entry to move and use the plus or minus keys to move the item up or down in the list.

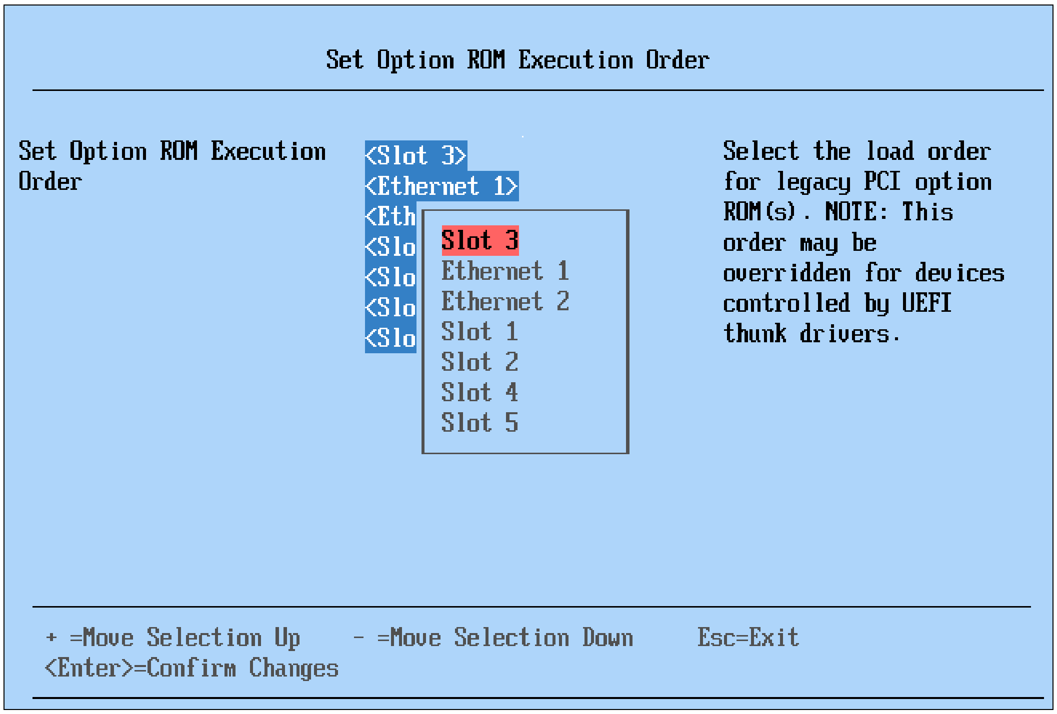


Figure 7-24 Set Option ROM Execution Order maintenance panel

7.6 Power supply considerations

The System x3690 X5 ships with two power supplies that will support a single processor, the system board memory, and all of the PCI slots, except when a high-performance video card is installed.

If you plan to install a high-performance video adapter, or the second processor and the mezzanine memory board, you must install the optional HE Redundant Power Supply Kit, part number 60Y0327, plus another optional power supply, part number 6070332. Installing these parts means that you will then have four power supplies inside the server. Without these options installed, the server will report the following warning in the system event log:

“`Non-redundant: Sufficient Resources for Redundancy Degraded`”

The MAX5 also has two power supplies. If power fails in the MAX5, the server will power off. Powering the server back on will not be possible unless the MAX5 has power or until both QPI cables have been disconnected from the server, after removing the ac power from the server. Ensure that both the MAX5 and the server to which it is attached are plugged into the same common power sources.

Ensure that half of the power supplies of both units are plugged into one utility power source and the remaining half are plugged into a separate utility power source. This approach eliminates the possibility of a single breaker or circuit fault taking down the entire server.

Think of the power supplies in your server like shock absorbers in a car. They are designed to absorb and overcome a wide variety of power conditions that can occur from an electric utility company, but like shock absorbers on a car, they will eventually begin to fail when fed a steady diet of unstable power. The times of their failures will most likely not coincide with a planned maintenance window. Therefore, ensure that the two halves of power supplies are plugged into two separate UPS sources to filter out all of the moderate to severe power fluctuations that occur.

Figure 7-25 depicts the power supplies of the combined x3690 X5 and the MAX5 and to which UPS power sources they must be attached to be fully redundant.

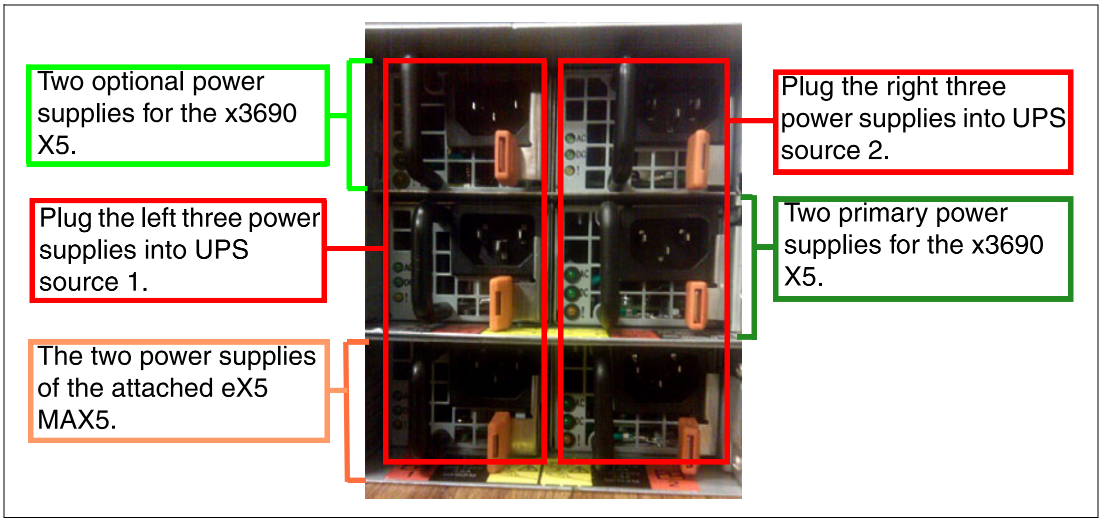


Figure 7-25 Recommended power cabling for the x3690 X5 and the MAX5

7.7 Using the Integrated Management Module

For any successful implementation of a server, there must be provisions set aside to provide access to perform troubleshooting or routine maintenance. The x3690 X5 comes standard with the Integrated Management Module (IMM). The IMM is a separate, independent operating environment that activates and remains active while the server is plugged into a good ac power source. The IMM monitors the hardware components of the server and the environment in which the server operates, looking for a potential hardware fault.

You can access part of the information that is stored in the IMM by using F1-Setup by selecting **System Settings** → **Integrated Management Module**. Figure 7-26 on page 328 shows the first panel of the Integrated Management Module configuration panel.

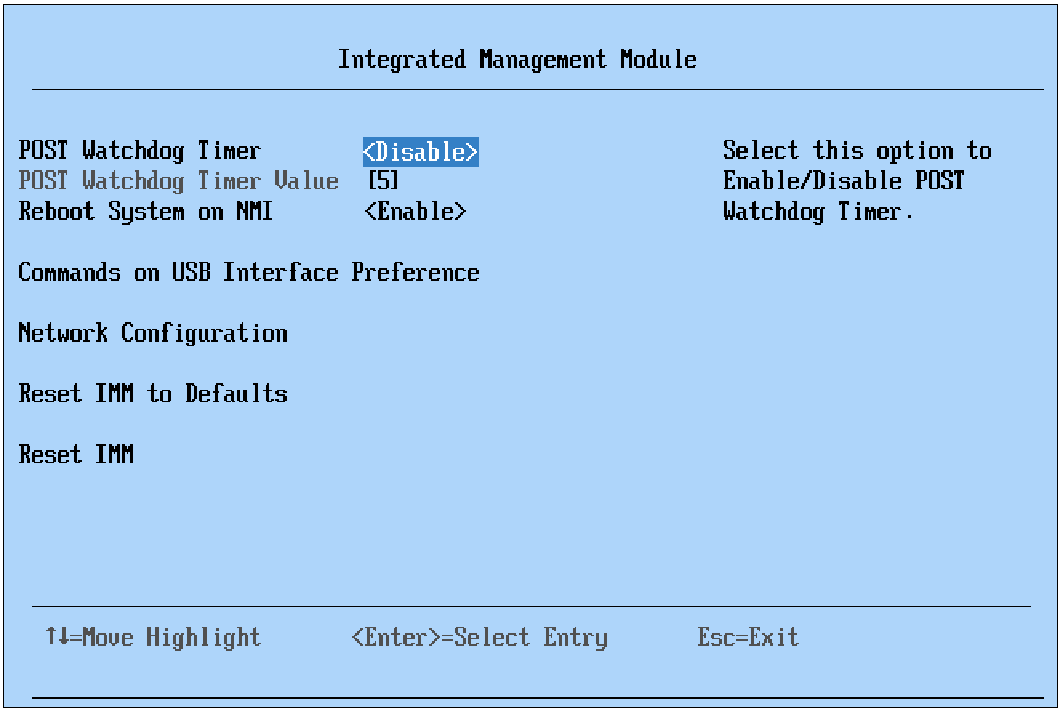


Figure 7-26 Integrated Management Module configuration panel

7.7.1 IMM network access

The greatest strength of the IMM is the ability to completely monitor and manage the server from over the network. The configuration of the IMM determines the amount of functionality that you have through this remote access.

IMM default configuration

The default network connection for the IMM on the x3690 X5 is through the *system management port* on the back of the server. The following settings are the default settings of the IMM from the factory:

- Network IP: Dynamic Host Configuration Protocol (DHCP)-assigned, or if DHCP is not available:
  - IP Address: `192.168.70.125`
  - Subnet Mask: `255:255:255:0`
  - Gateway: `0.0.0.0`
- Default user ID: *`USERID`*
- Default password: `PASSW0RD` where the 0 is a zero.

> **Resetting user credentials:** If you do not know the user credentials to enable you to connect to the IMM remotely, you can reset all passwords by using the **Reset IMM to Defaults** selection in Figure 7-26 on page 328.

7.7.2 Configuring the IMM network interface

The IMM provides two paths to establish a network connection between you and the IMM by setting either Dedicated or Shared for the Network Interface Port in the Network Configuration panel of F1-Setup. In F1-Setup, you can access this panel by selecting **System Setting** → **Integrated Management Module** → **Network Configuration**, as shown in Figure 7-27.

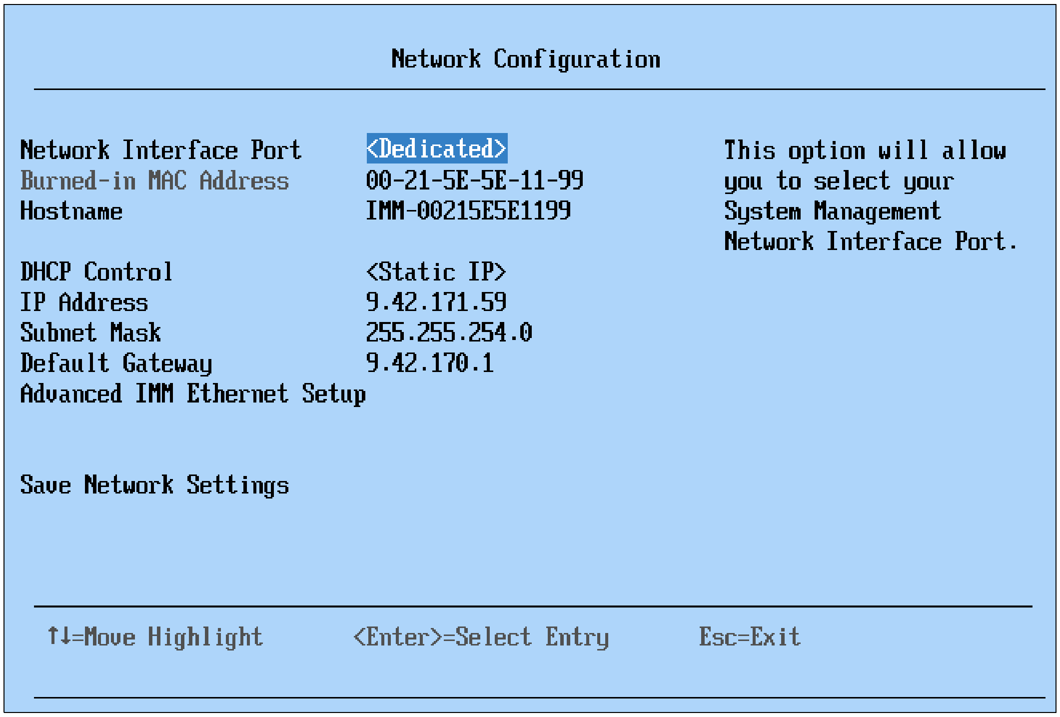


Figure 7-27 IMM Network Configuration panel

When configured as **Dedicated**, you connect to the network via the system management port. As shown in Figure 7-28, the port is located from the rear of the server on the left side of the video port. Using this port allows for easier separation of public and management network traffic. You can separate public and management network traffic when you connect your public network port to switch ports that belong to a public access virtual LAN (VLAN) and the management port is connected to a switch port that is defined by a separate management VLAN.

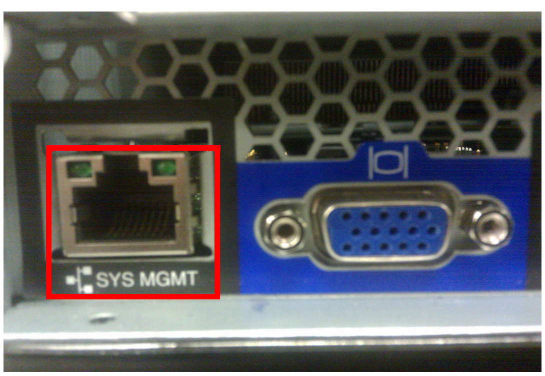


Figure 7-28 Dedicated 10/100 IMM system management port

When configured as **Shared**, the server shares network traffic on the second onboard Ethernet port, which is the port that is closest to the power supply, as shown in Figure 7-29 on page 330.

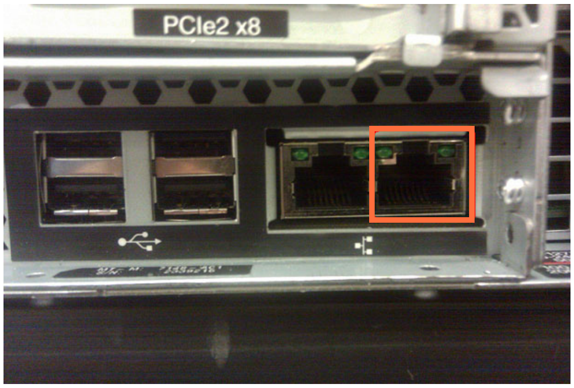


Figure 7-29 Onboard Ethernet port used when IMM Network Interface is Shared

Although this design eliminates a physical switch port and patch cable configuration, both the media access control (MAC) address for the second Ethernet port and the MAC address for the IMM will address through this network port. Therefore, at least two separate IP addresses are assigned to the same physical port. Sharing the port also prevents you from configuring the two onboard Ethernet ports in a network team using 802.3ad load balancing.

Using 802.3ad load balancing results in dropped packets for the IMM MAC address. Smart load balancing and failover are still available network teaming options. However, keeping the public traffic from the management traffic become more difficult. To maintain this separation, you must use network teaming software to establish a VLAN to be used by the server to send public tagged traffic to the network switch. You must configure the switch port as a trunk port to support both the public-tagged VLAN traffic, plus the untagged traffic for the management. You must define the management VLAN as the native VLAN on the switch port, so that its untagged traffic from the switch will be accepted by the IMM MAC and dropped by the second Ethernet port's MAC. Although this configuration is more complex, it is a common practice in virtualized servers running on a common set of host computers.

Although the IMM uses a dedicated RISC processor, there are limitations to the amount of network traffic to which the IMM can be exposed before complex functions, such as booting from a remote DVD or USB storage, become unreliable because of timing issues. Although the OS has all of the necessary drivers in place to deal with these timing issues, the UEFI is not as tolerant.

Therefore, for maintaining secured access, keep the IMM on a separate management network.

7.7.3 IMM communications troubleshooting

The *User's Guide for Integrated Management Module - IBM BladeCenter and System x* is an excellent guide to help you with every aspect of configuring and using the IMM. It is available at this website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5079770

Most communication errors are due to network switch configuration options, such as blocked ports or VLAN mismatches. A first step in determining this type of problem is to connect directly to the IMM port with a mobile computer and Ethernet patch cable to see if you can ping and then start a web session. The management port is a 10/100 Ethernet port, so if your mobile computer does not have a 10/100/1000 Ethernet port on it, you need to replace the patch cable with a 10/100 crossover cable. Only 1Gb Ethernet ports have the ability to auto-negotiate medium-dependent interface crossover (MDIX) when they auto-negotiate speed and duplex settings.

If you can ping the IMM, you have a good direct network link. If the web session fails, perform the following steps:

1. Try a separate web browser.
2. Directly access the Integrated Management Module configuration panel and reset the IMM in F1-Setup by selecting **System Settings** → **Integrated Management Module** → **Reset IMM**. Wait about 5 minutes for the IMM to complete enough of its reboot to allow you to ping it. This IMM reset will have no effect on the OS that is running on the server.
3. Try clearing your web browser cache.
4. Load factory default settings back on the IMM through the F1-Setup by selecting **System Settings** → **Integrated Management Module** → **Reset IMM to Defaults**. The IMM will need to be reset again after the defaults are loaded.
5. Contact IBM support.

7.7.4 IMM functions to help you perform problem determination

This section provides additional problem determination tips for the IMM. This section covers the following topics:

- "System Status" on page 331
- "Virtual light path diagnostics" on page 333
- "Hardware event log" on page 334
- "Remote control" on page 335

System Status

The first panel that you see after completing the login and the session time-out limits panel is the System Status panel, as shown in Figure 7-30 on page 332. This panel provides a quick summary review of the hardware status of the server. A green circle indicates all of the hardware is functioning well from a strictly hardware point of view.

The IMM can check on the status of server components, ServeRAID controllers, and PCIe interfaces to most PCIe adapters. It does not check on the functional status of most PCIe adapters with regard to their hardware connections to external devices. You need to refer to the system event log from within the OS or the switch logs of the network and Fibre Channel switches to which the server is connected to resolve connectivity issues there.

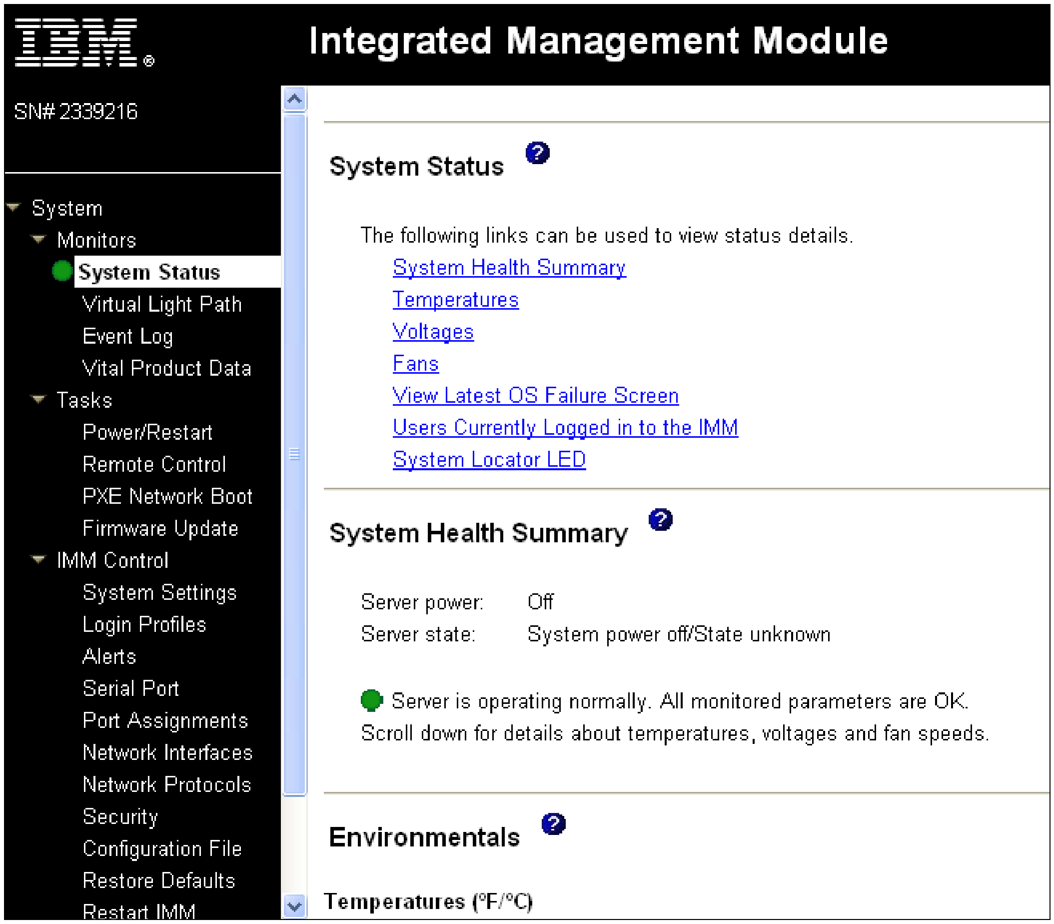


Figure 7-30 Integrated Management Module System Status

When an actual hardware error is detected in the server, the system status is represented by a red X and the System Health Summary provides information about the errors that are presently unresolved in the server, as shown in Figure 7-31 on page 333.

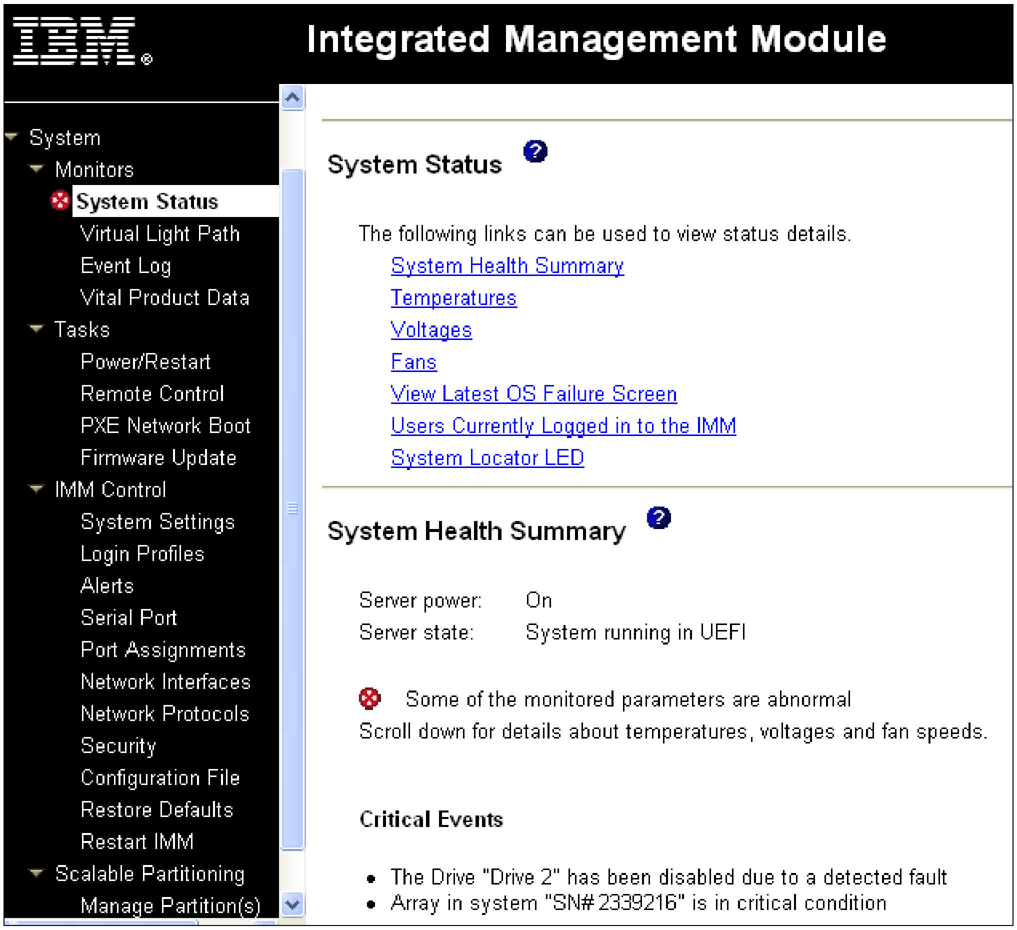


Figure 7-31 IMM System Status with a hard drive failure

Virtual light path diagnostics

It is easy to pinpoint a hardware fault when standing in front of the server, by noticing the first tier of light path diagnostics, with the error light on the front of the operator panel and on the rear of the server.

Pulling out the front operator panel reveals the second tier of light path diagnostics, as shown in Figure 7-32, that indicates the hardware subsystem that is experiencing the error.

Figure 7-32 Tier 2 of light path diagnostics

Servers are not typically located near the people who manage them. To help you see the event from a remote location, the IMM provides the capability to view all tiers of light path diagnostics, as shown in Figure 7-33.

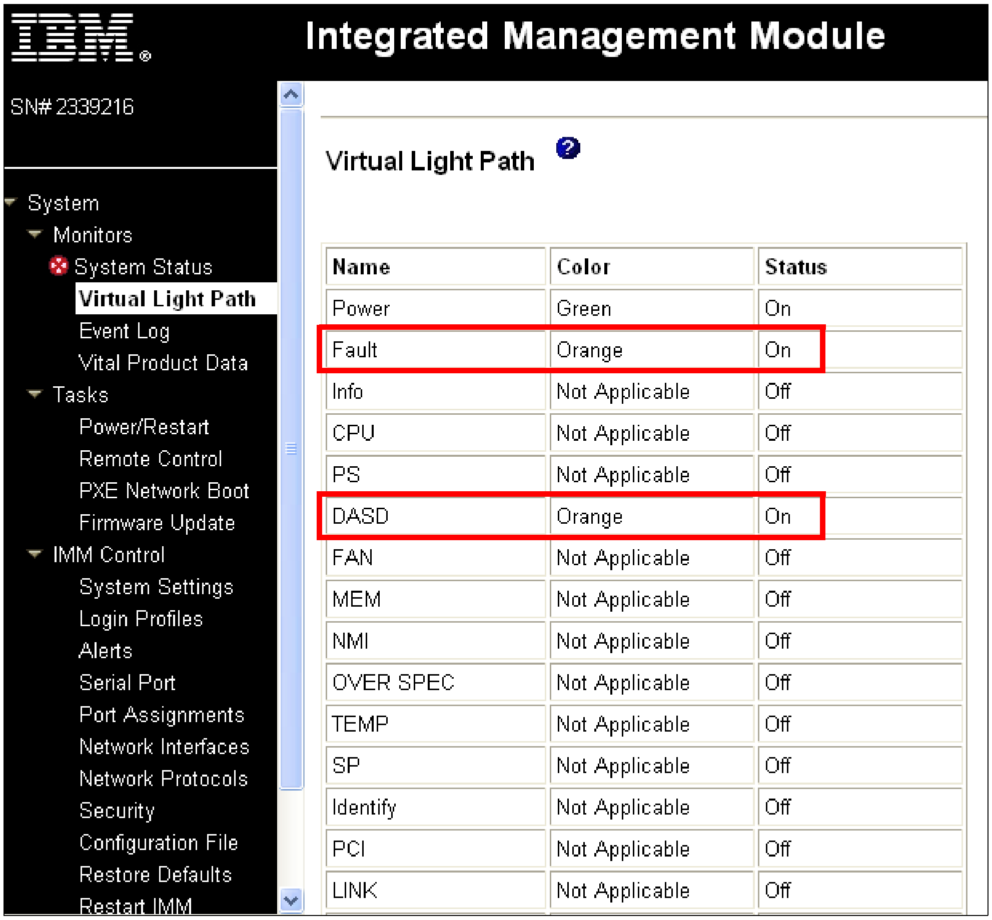


Figure 7-33 Integrated Management Module Virtual light path diagnostics

Hardware event log

For more detailed information, including the events that led up to a failure, you have access to the hardware event log. Not every event in the hardware event log is an event needing attention, but the log can provide insight about the cause or conditions that led up to a failure. You can save the event log to a text file to send to IBM support.

Figure 7-34 on page 335 shows the IMM Event Log for the hard drive failure.

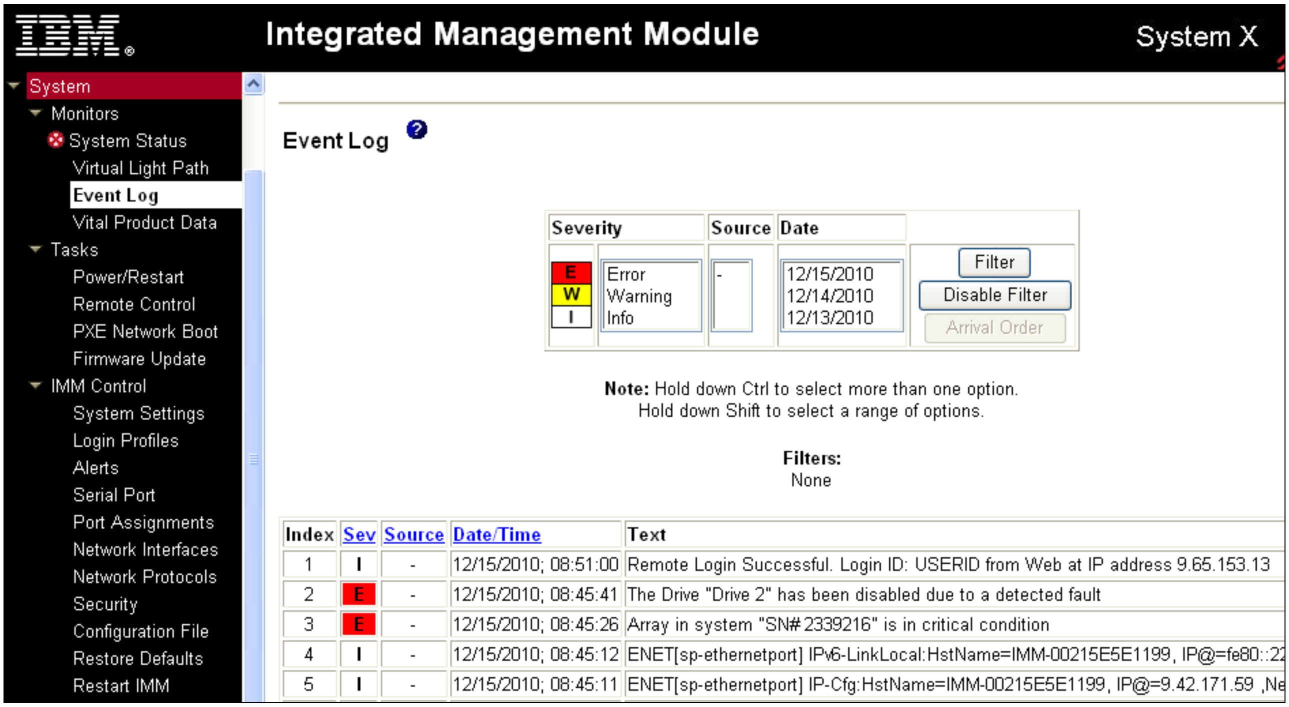


Figure 7-34 Integrated Management Module hardware Event Log

Remote control

Certain problems require that you view OS logs or enter UEFI via F1-Setup to detect them or fix them. For remotely managed servers, you have the Remote Control feature of the Integrated Management Module. Figure 7-35 shows the available options for starting a Remote Control session.

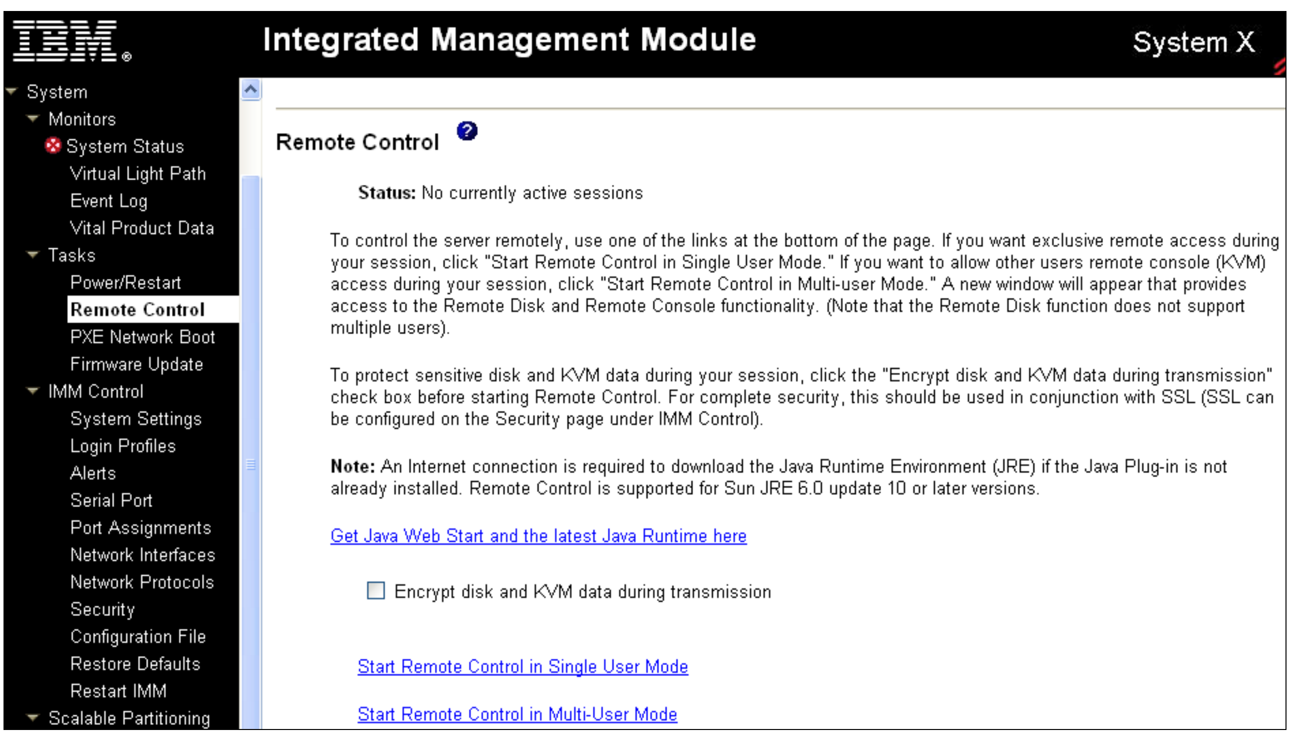


Figure 7-35 Integrated Management Module remote control session start-up window

The Remote Control feature provides the following advantages:

- It gives you the same functionality that you have with a keyboard, mouse, and video panel directly connected to the server.
- You can encrypt the session when using the Remote Control feature over public networks.
- You have the ability to use local storage or ISO files as mounted storage resources on the remote server that you are using. These storage resources can be unmounted, changed, and remounted throughout the session, as needed.
- Combined with the Power/Restart functions of the IMM, you can power down, reboot, or power on the server while maintaining the same remote control session.

Depending on the application that you are accessing through the IMM Remote Control feature, you might find that the mouse pointer is difficult to control. You can fix this problem in the Video Viewer by selecting **Tools** → **Single Cursor**, as shown in Figure 7-36.

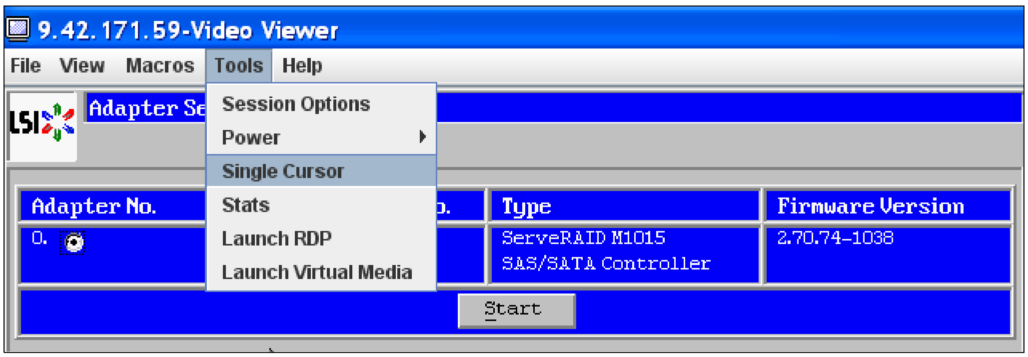


Figure 7-36 Fixing the mouse pointer in the Remote Control Video Viewer

7.8 UEFI settings

Hardware settings for the IBM System x3690 X5 are accessible through the UEFI. It also provides lower-level hardware settings that will most likely be transparent to the OS. UEFI replaces the old BIOS firmware interface.

You can obtain more information about UEFI at the following website:

http://www.uefi.org

New IBM System x models, including the x3690 X5, implement UEFI to take advantage of its advanced features. The UEFI page is accessed by pressing F1 during the system initialization process, as shown on Figure 7-37.

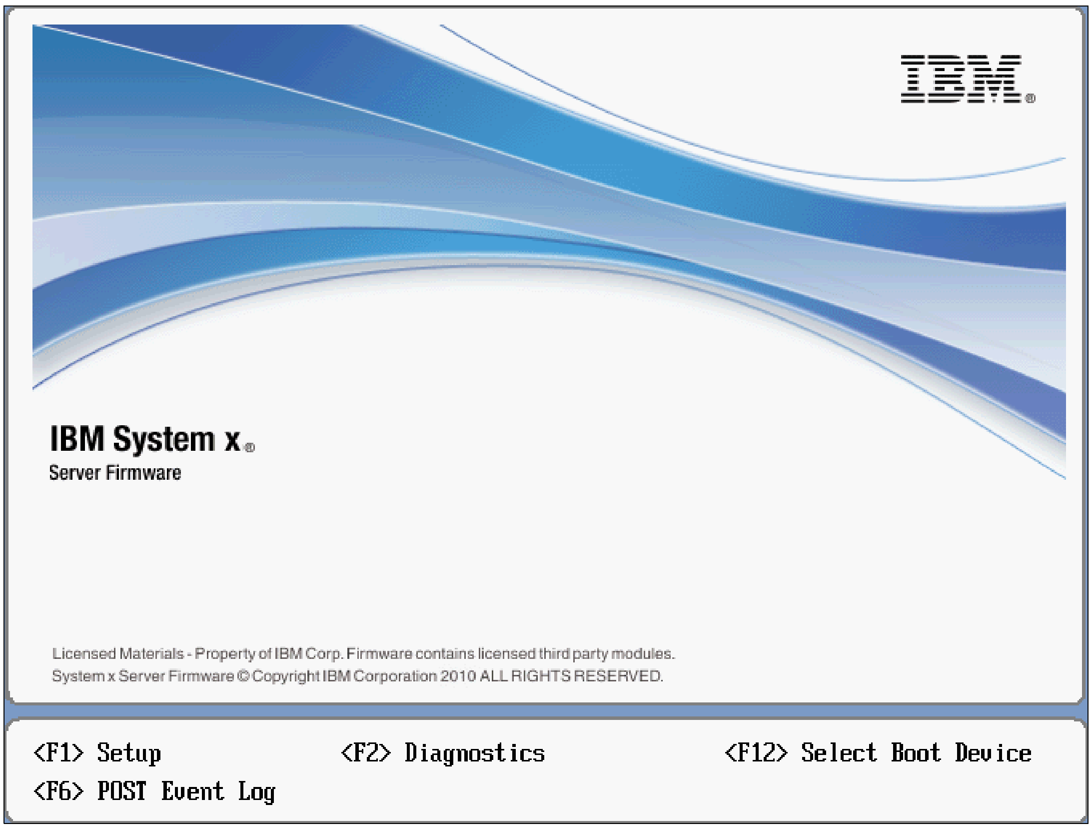


Figure 7-37 Press F1 during system start-up to access UEFI

Figure 7-38 on page 338 shows the UEFI System Configuration and Boot Management panel.

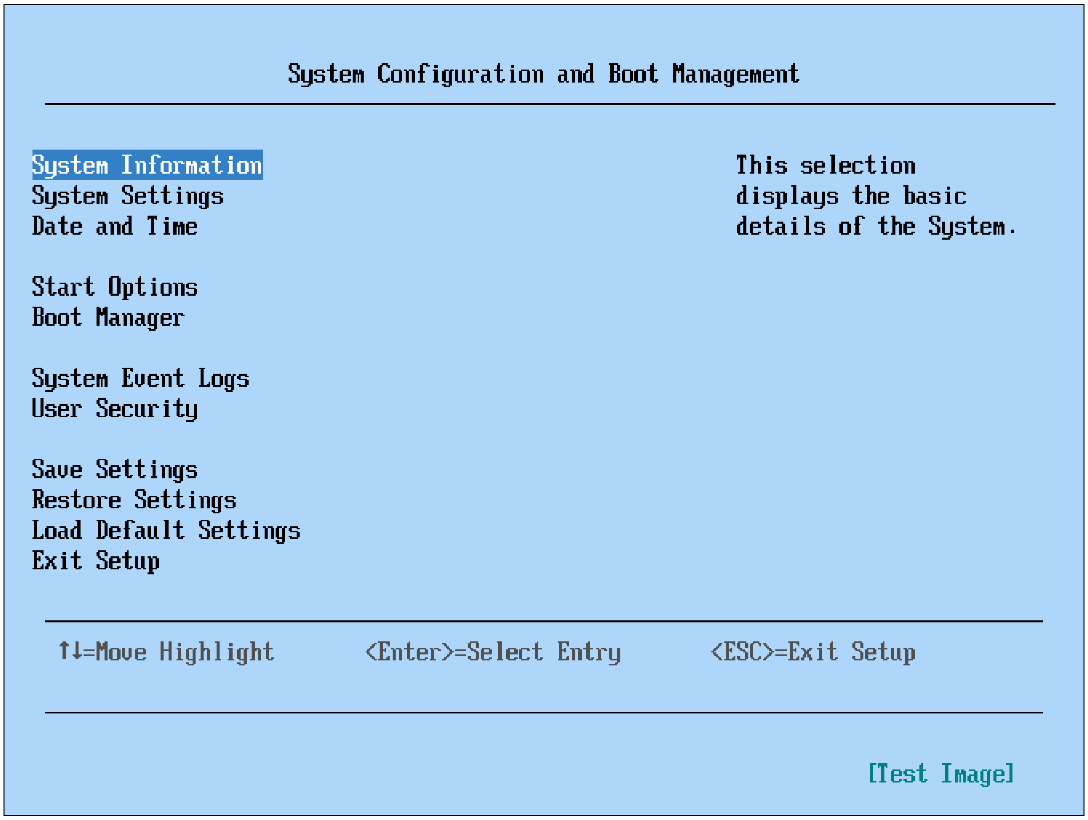


Figure 7-38 UEFI System Configuration and Boot Management settings panel

7.8.1 Scaled system settings

When you attach the MAX5 memory expansion unit, additional settings are enabled in the UEFI. Specifically, it adds the MAX5 Memory Scaling option in **System Settings** → **Memory**. All other settings behave similarly in the single-node configuration and the memory-expanded configuration. This additional option is shown in Figure 7-39 on page 339.

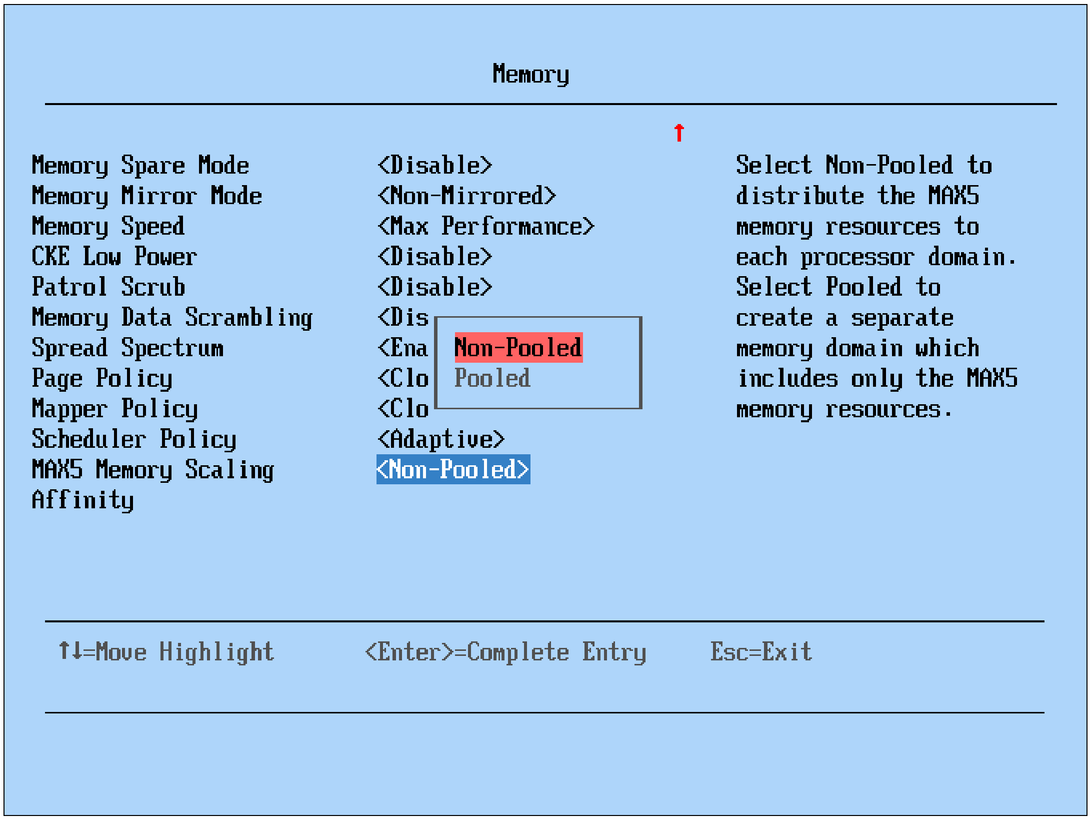


Figure 7-39 MAX5 Memory Scaling option in System x3690 X5 UEFI

The MAX5 Memory Scaling setting provides two options to determine how the system will present the memory capacity in the MAX5 unit to the running OS:

- Non-Pooled

  The default option splits the memory in the MAX5 and assigns part of it to each of the installed processors.

- Pooled

  This option presents the additional memory in the MAX5 as a separate pool of memory without its being assigned to any particular processor.

7.8.2 Operating system-specific settings

Determining what memory scaling option to use depends on the installed OS. The x3690 X5 only supports the x64 (AMD64/EM64T) version of the supported OSs in order to take advantage of the server's memory expansion capability.

Windows Server setting

The following Windows Server versions are supported:

- Windows Server 2008 R2
- Windows Server 2008 x64 Edition

When running Windows Server on the memory-expanded x3690 X5, set the MAX5 Memory Scaling option to **Non-Pooled**.

You can obtain a complete list of supported Windows Server OSs at IBM ServerProven:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/microsoft.html

Linux Server settings

The following Linux Server versions are supported:

- Red Hat Enterprise Linux 6 x64 Edition
- Red Hat Enterprise Linux 5 x64 Edition
- SUSE Linux Enterprise Server 11 for AMD64/EM64T
- SUSE Linux Enterprise Server 10 for AMD64/EM64T

Linux Server supports both MAX5 Memory Scaling modes; however, to get optimal performance, set the MAX5 Memory Scaling to **Pooled**.

You can obtain a complete list of supported Linux Server OSs at IBM ServerProven:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/redchat.html
http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/suseclinux.html

VMware vSphere settings

The following versions of VMware ESX Server are supported to run x3690 X5 expanded with MAX5:

- VMware ESXi Server 4.1
- VMware ESX Server 4.1

Systems running VMware ESX Server must use the Non-Pooled mode in the MAX5 Memory Scaling option.

You can obtain a complete list of supported VMware ESX Server OSs at IBM ServerProven:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/vmware.html

7.8.3 Power and performance system settings

The UEFI default settings are configured to provide optimal performance within a reasonable power consumption plane and are suitable for general server usage in most cases.

However, System Settings enable users to fine-tune the performance and power consumption ratio. The easiest way to fine-tune the performance and power consumption ratio is by determining the Operating Mode of the system by selecting **System Settings** → **Operating Modes** and making a selection in the Choose Operating Mode field, as shown in Figure 7-40 on page 341.

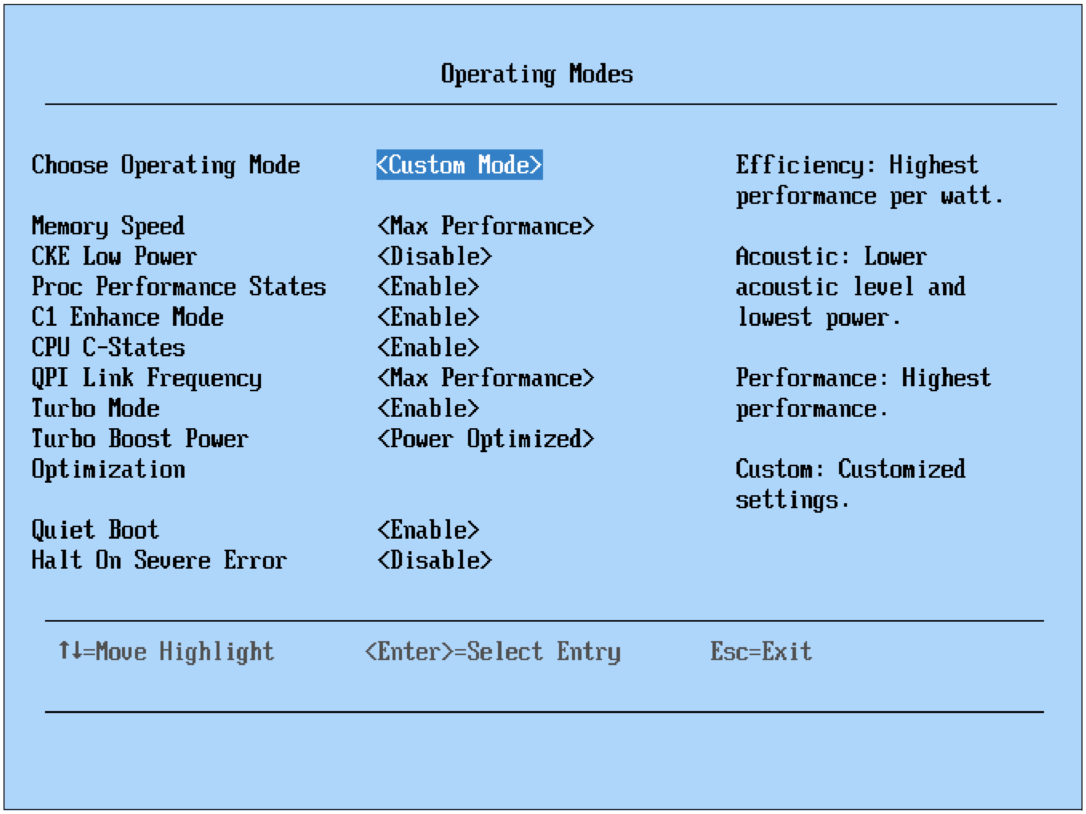


Figure 7-40 UEFI Operating Modes panel

As seen in Figure 7-40, the Operating Modes settings provide the user with several pre-configured Processor and Memory settings that are defined by system characteristics, whether it is Efficiency, Acoustic, Performance, or Custom. The default is Custom.

- Acoustic Mode emphasizes power-saving operations generating less heat and noise, at the expense of performance.
- Efficiency Mode provides the best balance of performance and power consumption.
- Performance Mode provides the best system performance at the expense of power efficiency.
- Custom Mode allows individual settings of the performance-related options.

Table 7-4 on page 342 displays comparisons of the operating modes for the IBM System x3690 X5. We recommend the available default Custom Mode setting, because it provides high system performance with acceptable power consumption. Also, consider Efficiency Mode to gain the best performance per watt operations.

Acoustic Mode is an available option when it is necessary to operate the system with minimal power consumption. Use Performance Mode when you want to get the best possible performance from the system.

Table 7-4 Comparison of operating modes

Settings	Efficiency	Acoustic	Performance	Custom (default)
Memory Speed	Power Efficiency	Minimal Power	Max Performance	Max Performance
CKE Low Power	Enabled	Enabled	Disabled	Disabled
Proc Performance States	Enabled	Enabled	Enabled	Enabled
C1 Enhance Mode	Enabled	Enabled	Disabled	Enabled
CPU C-States	Enabled	Enabled	Disabled	Enabled
QPI Link Frequency	Power Efficiency	Minimal Power	Max Performance	Max Performance
Turbo Mode	Enabled	Disabled	Enabled	Enabled
Turbo Boost Power Optimization	Power Optimized	None	Traditional	Power Optimized

A complete description of power-related and performance-related system settings is available in 2.7.3, “Performance-related individual system settings” on page 43.

System settings for specific server workloads

Setting the individual performance options of the x3690 X5 to achieve the desired system performance might not be a simple task, because many factors affect the overall result. A guideline for three specific server workloads is provided here as a starting point. These workloads represent server uses that most of the time require non-default settings to achieve optimal performance:

- **Virtualization**

  Several virtual OSs run on top of a hypervisor to gain better system consolidation and management. In this workload, server resources are shared dynamically among the virtual OSs. Popular virtualization software, such as ESXi, Hyper-V, and RHEV, falls into this category.

- **Low Latency**

  Data throughput and response time are of utmost importance. Typical uses are financial applications, such as trading, and media streaming servers.

- **High Performance Computing (HPC)**

  Extremely high performance is demanded from the server for advanced calculations. Typical applications are scientific and mathematical applications.

Table 7-5 provides recommended guidelines for system settings for specific server workloads.

Table 7-5 Guidelines for system settings for specific workloads

Setting	Virtualization	Low latency	HPC
Turbo Mode	Enabled	Disabled	Disabled
Turbo Boost Power Optimization	Power Optimized	Automatically disabled	Automatically disabled
Processor Performance States	Enabled	Disabled	Disabled
CPU C-States	Enabled	Disabled	Enabled
C1 Enhanced Mode	Enabled	Disabled	Enabled
Processor Data Prefetch	Enabled	Enabled	Enabled

Setting	Virtualization	Low latency	HPC
Hyper-Threading	Enabled	Disabled	Enabled
Execute Disable Bit	Enabled	Disabled	Disabled
Intel Virtualization Technology	Enabled	Disabled	Disabled
QPI Link Frequency	Max Performance	Max Performance	Max Performance
IMM Thermal Mode	Performance	Performance	Performance
CKE Low Power	Disabled	Disabled	Disabled
Memory Speed	Max Performance	Max Performance	Max Performance
Page Policy	Closed	Closed	Closed
Mapper Policy	Closed	Closed	Closed
Patrol Scrub	Disabled	Disabled	Disabled
Demand Scrub	Enabled	Disabled	Disabled

7.8.4 Optimizing boot options

UEFI systems, including the x3690 X5, can take more time to start up when scaled with the MAX5 expansion. The number of installed adapters directly affects the time that it takes for UEFI to initialize. A simple step to reduce the overall system start-up time is to remove any unnecessary boot devices from the boot order list. In particular, ensure that the PXE network boot is removed from the boot order list if it is not used. PXE network boot can increase your start-up time substantially, especially if it is also placed in the incorrect order in the boot order list.

Use the following steps to remove entries from the Boot Order menu and to configure the correct boot order:

1. Power on the system and press F1 on the UEFI splash panel to enter the UEFI System Configuration and Boot Management settings panel.
2. Navigate to the Boot Manager menu by selecting **Boot Manager** from the System Configuration and Boot Management menu. The Boot Manager panel displays, as shown in Figure 7-41 on page 344.

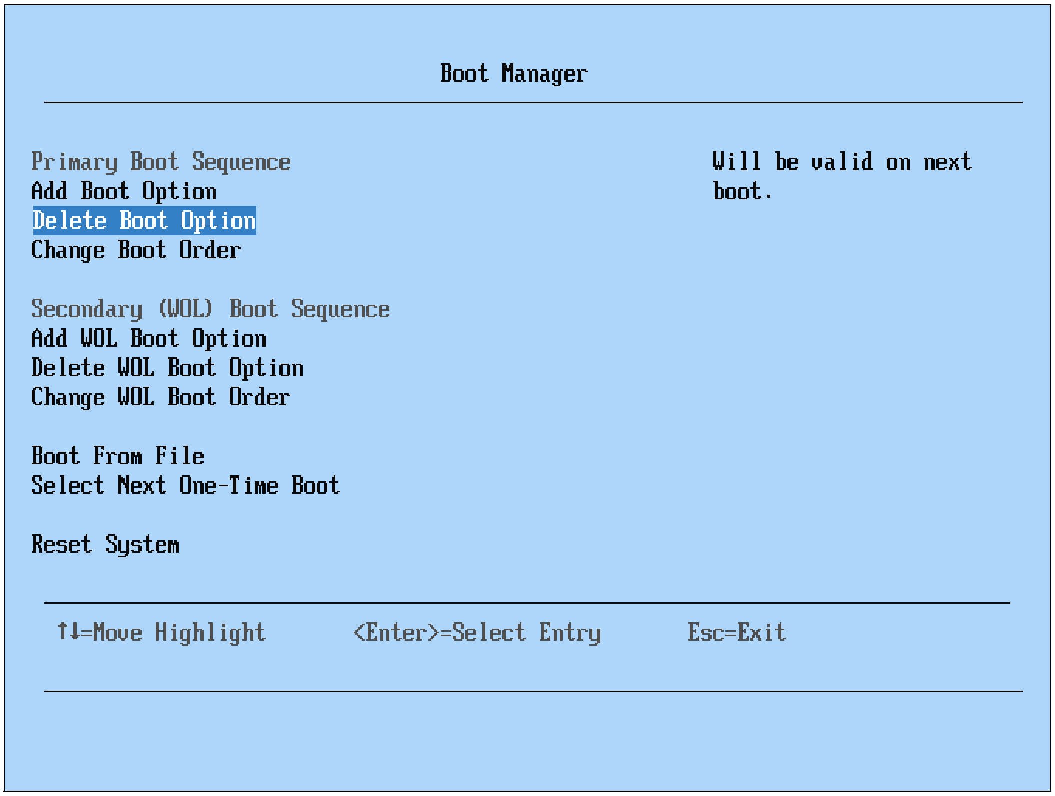


Figure 7-41 UEFI Boot Manager menu

3. Select **Delete Boot Option** from the Boot Manager menu.
4. Select all items that you do not want to boot from by using the Spacebar to select them. When you have selected all of the items, scroll down to the end of the page using the down arrow key and select **Commit Changes**, as shown in Figure 7-42 on page 345. In the example in Figure 7-42 on page 345, **Hard Disk 0** has also been selected for removal. The system is running Windows Server 2008 R2 with GUID Partition Table (GPT) disk support from UEFI, and therefore, Windows Boot Manager is used to boot the OS, not Hard Disk 0.

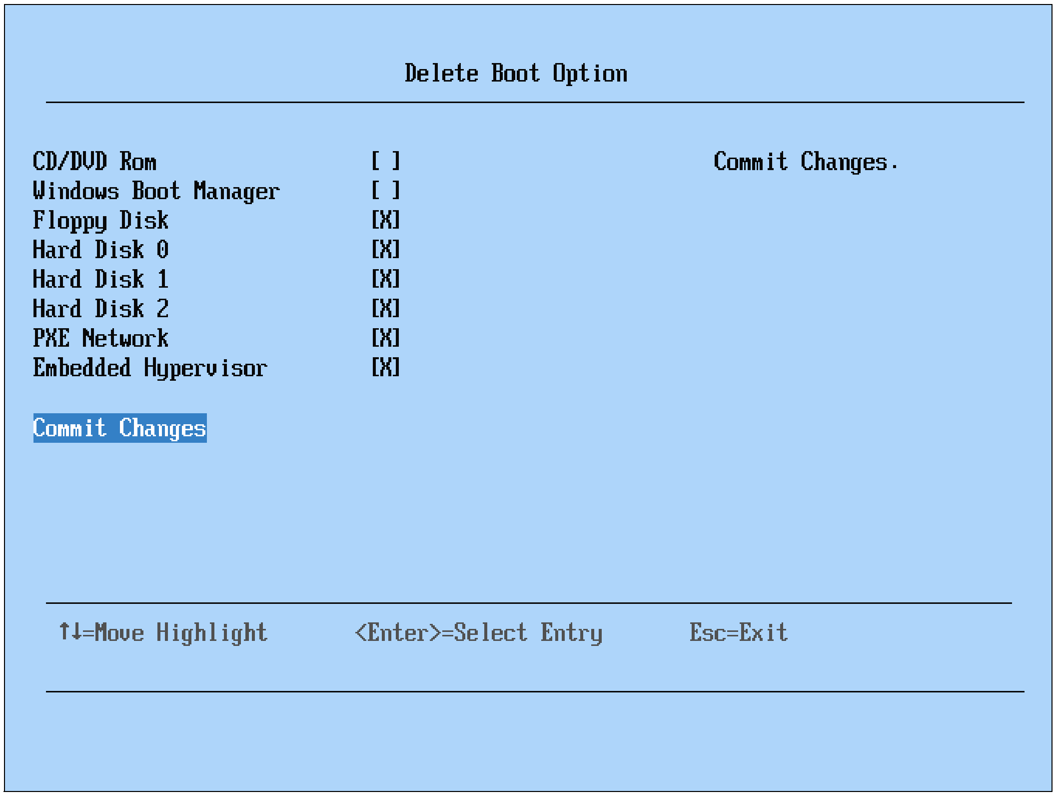


Figure 7-42 Deleting boot options

5. Press Esc when you have finished to return to the Boot Manager menu.
6. Select **Change Boot Order** from the Boot Manager menu.
7. Press Enter to make the device list active.
8. Use the up and down arrow keys to navigate to the device for which you want to change the order. After you highlight the device, use the - or Shift and + keys to move the device up or down the list. You can then perform the same actions to move other devices up or down the list. Press Enter when done.
9. Use the down arrow key to highlight **Commit Changes** and press Enter to commit the changes that you have made.
10. Press Esc to return to the Boot Manager menu.
11. Press Esc again to exit to the System Configuration and Boot Management menu.
12. Press Esc again to exit the UEFI and press the Y key to exit and save any changes that you have made. The x3690 X5 then proceeds to boot normally.

> **Tip:** It is common practice to place CD/DVD Rom higher than the default OS in the boot order list to accommodate system tools media that require booting from CD/DVD.

More information about editing and cleaning the boot options is available in 7.5.3, “Cleaning up the boot sequence” on page 322.

7.9 Operating system installation

Scaling an x3690 X5 with MAX5 allows the OS to access additional system memory that is installed in the MAX5. Most of the configurations are done on the hardware and firmware level, so that the OS installation method is similar to that of a non-scaled system.

We describe the following topics in this section:

7.9.1 Installation media

The installation process requires an installation media to proceed, which usually comes in CD or DVD form. However, in cases where CD/DVD media devices are not available or the process needs to be performed remotely, we can take advantage of several features that are included in the IBM System x3690 X5.

Preboot eXecution Environment (PXE) network boot

The onboard Broadcom 5709 Gigabit Ethernet supports the PXE network boot, making it possible to boot into the PXE and access installation files from a remote location.

System x3690 X5 can boot from PXE simply by setting the boot options to **PXE Network** in the UEFI. The default system settings, as shown in Figure 7-43 on page 347, are to boot to PXE network after boot attempts to CD/DVD Rom, Floppy Disk, and Hard Disk 0.

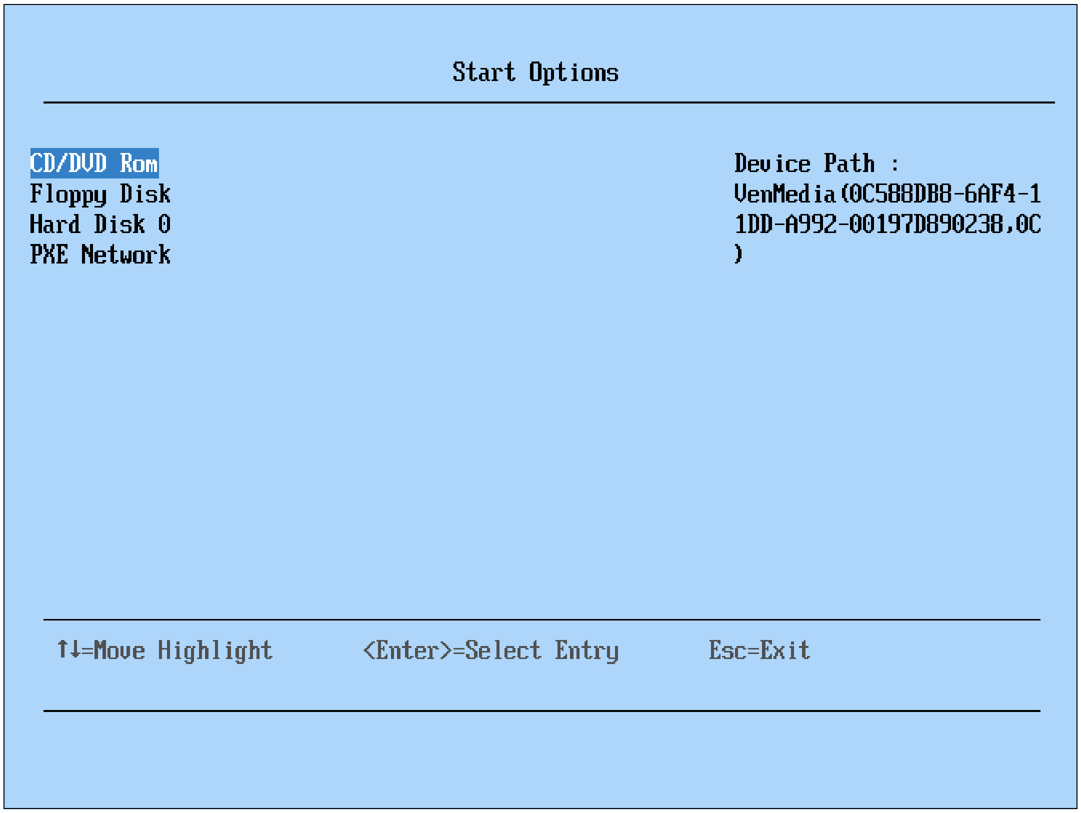


Figure 7-43 Default UEFI boot order

In a typical configuration, these server functions are required to allow the server to perform PXE network boot and installation:

- DHCP server, providing IP address
- Trivial File Transfer Protocol (TFTP) server, providing initial boot and installation environment
- File server, providing OS installation files

Use the guides that are provided in the following list to perform the PXE network boot installation that is specific to each OS:

- Windows Server 2008 R2

  http://technet.microsoft.com/en-us/library/cc772106(WS.10).aspx

- Red Hat Enterprise Linux 6

  http://docs.redhat.com/docs/en-US/Red_Hat_Enterprise_Linux/6/html/Installation_Guide/index.html

- SUSE Linux Enterprise Server 11

  http://www.novell.com/documentation/sles11/book_sle_deployment/?page=/documentation/sles11/book_sle_deployment/data/pre_sle.html

- VMware vSphere ESXi 4.1

  http://www.vmware.com/pdf/vsphere4/r41/vsp_41_esxi_i_vc_setup_guide.pdf

Image loading using IMM Remote Control

Integrated Management Module (IMM) on IBM System x3690 X5 is equipped with a Remote Control function that is capable of mounting an installation image on the management workstation. It is accessed using the IMM **Remote Control** option and either Single User Mode or Multi-User Mode, as shown in “Local USB port” on page 349. Using this function opens the Virtual Media and Video Viewer of the controlled system.

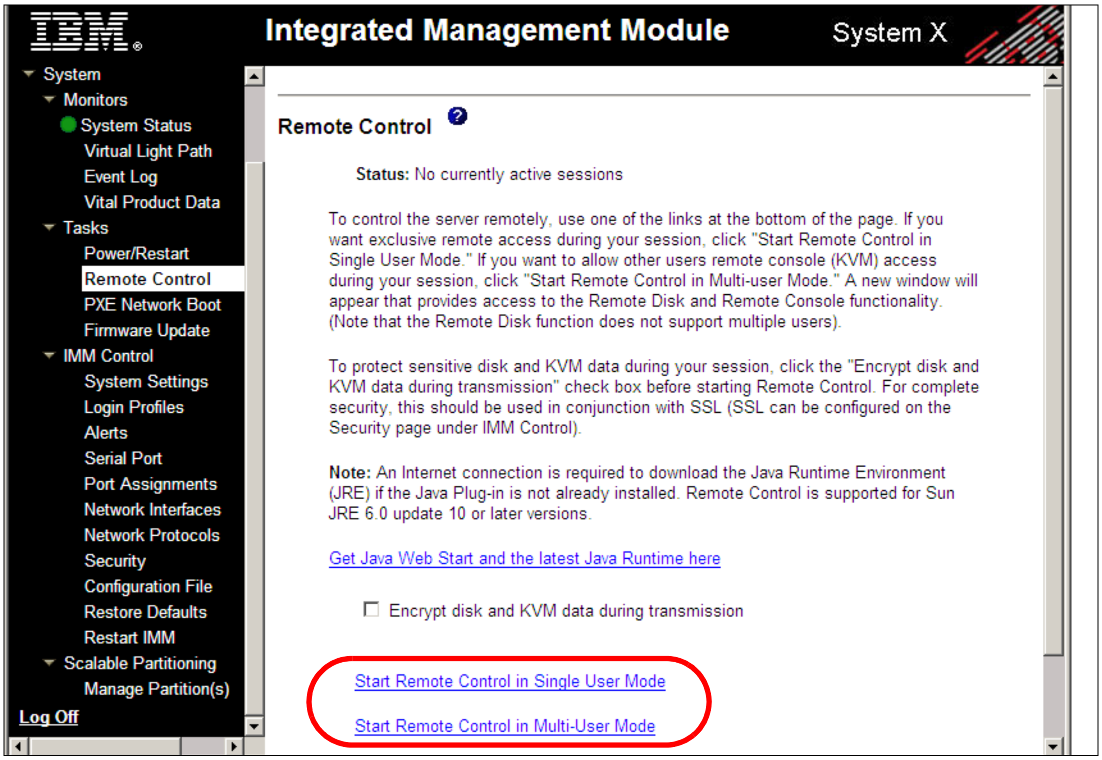


Figure 7-44 Remote Control function in the IMM

> **Tip:** When using Microsoft Internet Explorer (IE) version 7 or 8, if the Remote Control window does not open and you get no warning that a pop-up was blocked, it is possible that IE advanced security settings are blocking the Java application from starting.
>
> A work-around for this issue is to hold the Ctrl key while clicking **Start Remote Control in Single User Mode** or **Start Remote Control in Multi-User Mode** in the IMM Remote Control window.
>
> RETAIN tip H196657 describes other work-arounds:
>
> `http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5083262`

Figure 7-45 on page 349 shows the Virtual Media Session interface where we can mount a selected installation image.

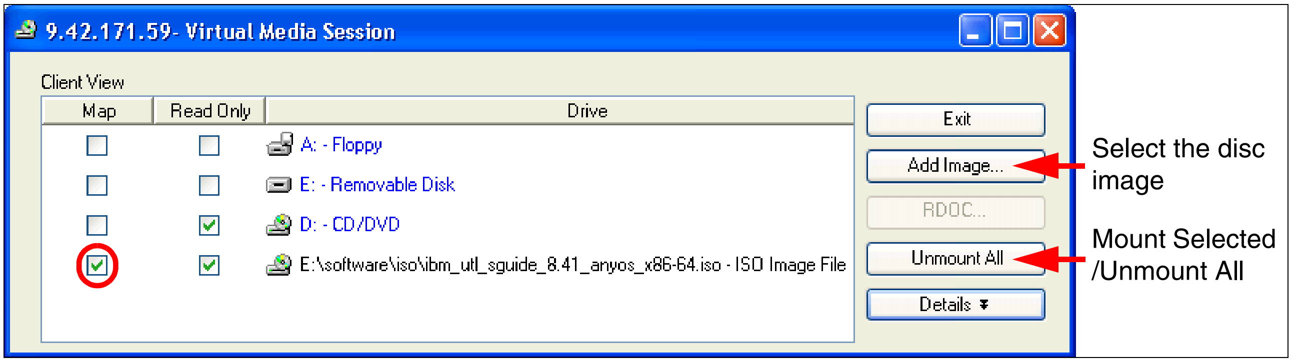


Figure 7-45 IMM Virtual Media Session interface

Follow this sequence to mount an image as virtual CD/DVD media at the controlled system:

1. Click **Add Image**.
2. Choose an image in `.iso` or `.img` format.
3. Check the box under Map of the image to mount.
4. Click **Mount Selected**.

Simply click **Unmount All** to unmount the image.

You can also mount a local Floppy, CD/DVD, or Removable Disk to a remote system using similar methods.

Local USB port

You can use the local USB port to attach a USB flash drive that contains the OS installation files. There are several methods to create a bootable flash drive.

For VMware, you can use the embedded hypervisor key, which is pre-installed with ESXi, and you do not need to install VMware. For more information about the embedded hypervisor key, see 2.9.1, “VMware ESXi” on page 50.

For Linux, look on the vendor websites. They contain information about installation with a USB flash drive. For example, the following websites provide details for using a USB key as an installation medium:

- Installing Red Hat Linux from a USB flash drive:

  `http://ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101131`

- How to create a bootable USB drive to install SLES:

  `http://www.novell.com/support/php/search.do?cmd=displayKC&docType=kc&externalId=3499891`

You can also use the ServerGuide Scripting Toolkit to create a bootable USB flash drive, as explained in the next section.

ServerGuide Scripting Toolkit

As described in 9.9, “IBM ServerGuide Scripting Toolkit” on page 507, you can use the ServerGuide Scripting Toolkit to customize your OS deployment. You can use the ServerGuide Scripting Toolkit for Windows, Linux, and VMware. This section contains information about deployment to allow you to begin using the Toolkit as quickly as possible.

For the full details, see the *IBM ServerGuide Scripting Toolkit, Windows Edition User's Reference* and *IBM ServerGuide Scripting Toolkit, Linux Edition User's Reference* at the following website:

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

Windows installation

This section describes the process to install the ServerGuide Scripting Toolkit, create a deployment image for Windows Server 2008 R2 Enterprise Edition, and copy this image to a USB key for deployment. To configure a USB key for deployment, you need the following devices:

- A system running Windows Vista, Windows Server 2008, Windows 7, Windows Server 2008 R2, Windows 2.1 Preinstallation Environment (PE), or a Windows 3.0 PE session
- A USB key with a storage capacity at least 64 MB larger than your Windows PE image, but not less than 4 GB

Procedure

We use this procedure:

1. Install the ServerGuide Scripting Toolkit.
2. Create a deployment image.
3. Prepare the USB key.

Installing the ServerGuide Scripting Toolkit

You must install the English language version of the Windows Automated Installation Kit (AIK) for the Windows 7 family, Windows Server 2008 family, and Windows Server 2008 R2 family, which is available at the following website:

`http://www.microsoft.com/downloads/en/details.aspx?familyid=696DD665-9F76-4177-A811-39C26D3B3B34&displaylang=en`

Follow these steps to install the ServerGuide Scripting Toolkit, Windows Edition:

1. Download the latest version from the following website:

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

2. Create a directory, for example, `C:\sgshare`.
3. Decompress the *`ibm_utl_sgtkwin_X.XX_windows_32-64.zip`* file to the directory that you created, for example, `C:\sgshare\sgdeploy`.

Creating a deployment image

Follow these steps to create a Windows installation image:

1. Start the Toolkit Configuration Utility in the `C:\sgshare\sgdeploy` directory.
2. Select **Add Operating System Installation Files**, as shown in Figure 7-46 on page 351.

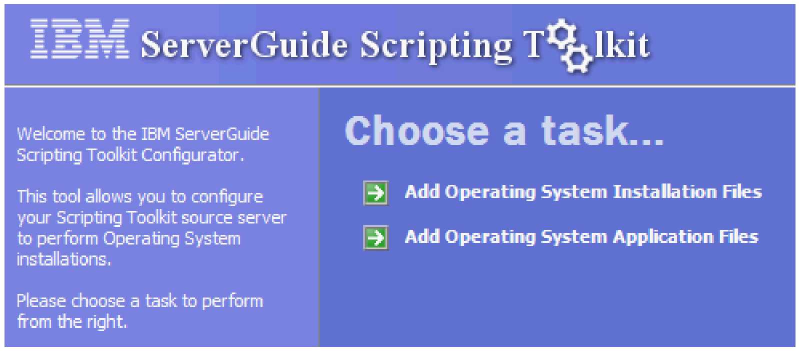


Figure 7-46 IBM ServerGuide Scripting Toolkit

3. Choose the OS type that you want and click **Next**, as shown in Figure 7-47.

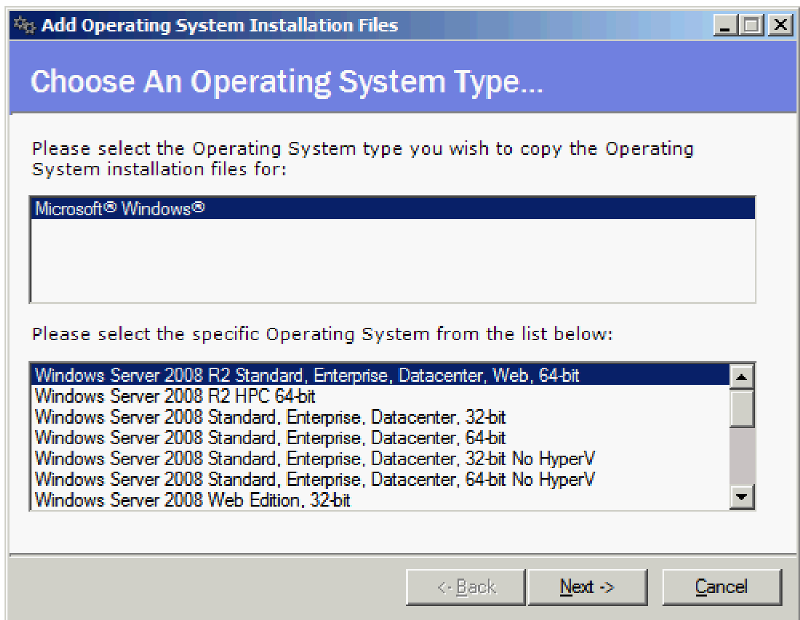


Figure 7-47 Select the OS type

4. Insert the correct OS installation media or select the folder that contains the installation files for the source, as shown in Figure 7-48 on page 352. If necessary, modify the target and click **Next**.

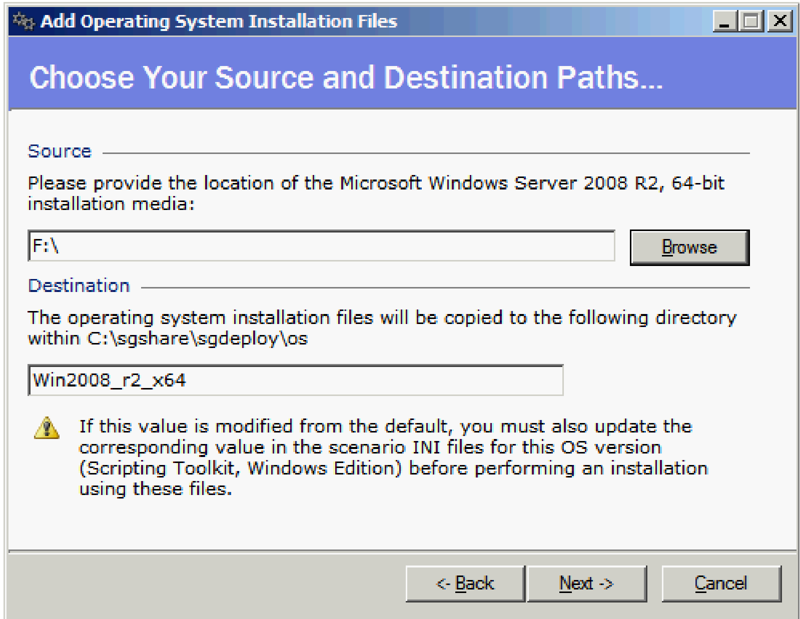


Figure 7-48 Define the source and target

5. When the file copy process completes successfully, as shown in Figure 7-49, click **Finish**.

Figure 7-49 Copy process completes successfully

6. Open a command prompt and change to the `C:\sgshare\sgdeploy\SGTKWinPE` directory. Use the following command to create the Windows installation image:

```
SGTKWinPE.cmd ScenarioINIs\Local\Win2008_R2_x64_EE.ini
```

7. When the process completes, as shown in Figure 7-50 on page 353, your media creation software is started to create a bootable media from the image. Cancel this task.

```
18:26:21 - Creating the WinPE x64 ISO...

18:27:07 - The WinPE x64 ISO was created successfully.

*** WinPE x64 ISO:  c:\sgshare\sgdeploy\WinPE_ScenarioOutput\Local_Win2008_R2_x6
4_EE\WinPE_x64.iso

18:27:07 - Launching the registered software associated with ISO files...

*** Using ISO File:  c:\sgshare\sgdeploy\WinPE_ScenarioOutput\Local_Win2008_R2_x
64_EE\WinPE_x64.iso

18:27:08 - The WinPE x64 build process finished successfully.

SGTKWinPE complete.
```

Figure 7-50 Build process is finished

Preparing the USB key

Follow these steps to create a bootable USB key with the Windows installation image that was created in “Creating a deployment image” on page 350:

1. Insert your USB key.

2. Use **diskpart** to format the USB key using FAT32. All files on the USB key will be deleted. At the command prompt, type the commands that are listed in Figure 7-51.

```
C:\>diskpart
Microsoft DiskPart version 6.1.7600
Copyright (C) 1999-2008 Microsoft Corporation.
On computer:

DISKPART> list disk

  Disk ###  Status         Size     Free     Dyn  Gpt
  --------  -------------  -------  -------  ---  ---
  Disk 0    Online          271 GB      0 B        *
  Disk 1    Online          135 GB      0 B
  Disk 2    Online         7839 MB      0 B

DISKPART> select disk 2
Disk 2 is now the selected disk.

DISKPART> clean
DiskPart succeeded in cleaning the disk.

DISKPART> create partition primary
DiskPart succeeded in creating the specified partition.

DISKPART> select partition 1
Partition 1 is now the selected partition.

DISKPART> active
DiskPart marked the current partition as active.

DISKPART> format fs=fat32
  100 percent completed
DiskPart successfully formatted the volume.

DISKPART> assign
DiskPart successfully assigned the drive letter or mount point.

DISKPART> exit
```

Figure 7-51 Using diskpart to format the USB memory key

3. Copy the contents from `C:\sgshare\sgdeploy\WinPE_ScenarioOutput\Local_Win2008_R2_x64_EE\ISO` to the USB key. The USB key now contains the folders and files that are shown in Figure 7-52.

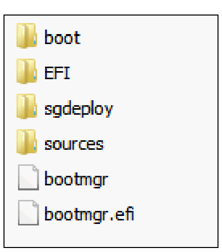


Figure 7-52 Contents of the USB key

4. Boot the target system from the USB key. The deployment will execute automatically.

RAID controller: If the target system contains a RAID controller, RAID will be configured as part of the installation.

Linux and VMware installation

The procedure for the Linux and VMware installation is similar to the Windows procedure:

1. Install the ServerGuide Scripting Toolkit.
2. Create a deployment image.
3. Prepare a USB key.

For more information, see the *IBM ServerGuide Scripting Toolkit, Linux Edition User's Reference* at the following website:

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

Preboot eXecution Environment (PXE)

The Preboot eXecution Environment (PXE) is an environment to boot computers using a network interface for operating system deployment. All eX5 systems support PXE.

For example, you can use the ServerGuide Scripting Toolkit. For more information, see the *IBM ServerGuide Scripting Toolkit User's Reference* at the following website:

http://www.ibm.com/support/docview.wss?uid=psg1SERV-TOOLKIT

Tivoli Provisioning Manager for OS Deployment

IBM Software has an offering for users needing advanced features in automating and managing the remote deployment of OSs and virtual images, in the form of Tivoli Provisioning Manager for OS Deployment. Tivoli Provisioning Manager for OS Deployment is available in a stand-alone package and as an extension to IBM Systems Director. You can obtain more information about these offerings at the following websites:

- Tivoli Provisioning Manager for OS Deployment:

  http://ibm.com/software/tivoli/products/prov-mgr-os-deploy/

- Tivoli Provisioning Manager for OS Deployment IBM Systems Director Edition:

  http://ibm.com/software/tivoli/products/prov-mgr-osd-isd/

7.9.2 Integrated virtualization hypervisor

The IBM USB Memory Key for virtualization is available as an option for System x3690 X5. It is preloaded with VMware ESXi embedded hypervisor software and attached to an internal USB connector on the x8 low profile PCIe riser card. Using this tool eliminates the need to perform the hypervisor installation and provides added performance and reliability as opposed to using standard mechanical drives.

Tips:

- To boot the system to the integrated hypervisor option, add **Embedded Hypervisor** to the system boot options in the UEFI.
- The integrated virtualization option is preloaded with the ESXi embedded hypervisor; therefore, no installation activity is required. However, to run it on a memory-expanded x3690 X5, you must enable the `allowInterleavedNUMAnodes` parameter. See "VMware vSphere ESXi 4.1" on page 358 for details.

See 4.13, “Integrated virtualization” on page 174 for ordering information for the integrated virtualization option.

7.9.3 Windows Server 2008 R2

Windows Server 2008 R2 is currently the latest version of the popular Windows Server OS. It provides enhancements in several areas, including virtualization, web, and management.

The following list provides general tips about Windows Server 2008 R2 installation on the memory-expanded x3690 X5:

- We recommend using IBM ServerGuide to simplify the installation of Windows Server 2008 R2. IBM ServerGuide integrates the RAID array configuration and driver installation into the OS installation process. You can download the IBM ServerGuide image from the following website:

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

  Note the supported system of each version. At the time of this writing, IBM ServerGuide v8.41 64 bit provides the latest support for x3690 X5.

  We describe using the IBM ServerGuide to perform the Windows Server installation on IBM eX5 servers in 9.8, “IBM ServerGuide” on page 501.

- Windows Server 2008 R2, Standard x64 Edition only supports 32 GB of memory. Use at least Enterprise Edition to take full advantage of the MAX5 expansion memory capability.

  You can read information about Windows Server memory limits at the following website:

  http://msdn.microsoft.com/en-us/library/aa366778.aspx

- When using ServerGuide, the RAID array configuration is set in the Configure RAID Adapter panel. Choose **Keep Current Adapter Configuration** if the RAID array has been set previously and the next installation will use that current configuration.
- In the ServerGuide partition setup panel, selecting the **Clear All Disks** check box clears the contents of all disks attached to the system, including SAN storage logical drives that are mapped to the system.

  Figure 7-53 on page 357 shows the partition setup panel with the Clear All Disks option. Proceed with caution when using this option.

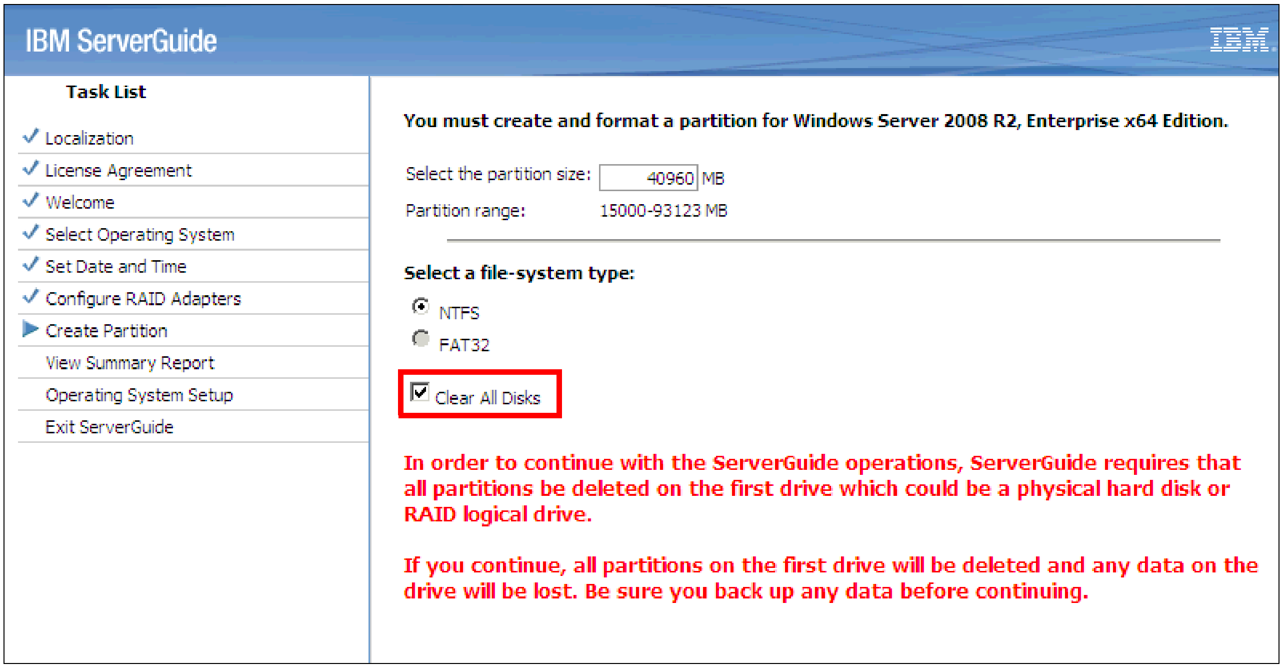


Figure 7-53 IBM ServerGuide partition setup window

We recommend that you map SAN storage logical drives after the OS installation process completes, except in Boot from SAN configurations.

- You can access more information about the Windows Server 2008 R2 installation and configuration at the following website:

  `http://www.microsoft.com/windowsserver2008/en/us/product-documentation.aspx`

- If Windows Server 2008 R2 cannot be started and the system has 1 TB or more of memory installed, try to apply the Microsoft KB98059 that can be downloaded from the following website:

  `http://support.microsoft.com/kb/980598`

- If a system with 128 GB or more of memory installed and running Windows Server 2008 R2 with Hyper-V receives error message "`STOP 0x000000A`", try to apply the Microsoft KB979903 Hotfix that can be downloaded from the following website:

  `http://support.microsoft.com/kb/979903`

- We recommend that you set Flow Control to **Enabled** in the Broadcom NetXtreme II Properties to ensure that no data transmission issue exists between servers with large performance gaps.

  The following steps describe this task:

  a. Select **Start** → **Administrative Tools** → **Computer Management** → **Device Manager**.

  b. Expand **Network adapters**.

  c. Double-click **Broadcom BCM5709C NetXtreme II GigE**.

  d. Select **Advanced** → **Flow Control**.

  e. Set Value to **Rx & Tx Enabled** and click **OK**.

- You can download Windows Server 2008 R2 Trial Software from the following website:

  `http://www.microsoft.com/windowsserver2008/en/us/trial-software.aspx`

7.9.4 Red Hat Enterprise Linux 6 and SUSE Linux Enterprise Server 11

IBM System x3690 X5 with MAX5 memory expansion fully supports the latest version of two major Linux server distributions: RHEL6 from Red Hat and SLES11 from Novell.

The installation of RHEL6 and SLES11 is a straightforward process, because all of the required drivers for System x3690 X5 are included in the kernel. However, use the following general tips when installing either of these OSs:

- The x3690 X5 supports only the 64-bit version of RHEL6 and SLES11.
- RHEL6 and SLES11 are UEFI-compliant OSs. Each product adds its boot file on top of the system's UEFI boot order.
- We recommend not to attach SAN storage to the server during the installation process. Attach the SAN storage after the OS is up and running.
- You can access more information about the installation and configuration of Red Hat Enterprise Linux at the following website:

  https://access.redhat.com/knowledge/docs/manuals/Red_Hat_Enterprise_Linux/
- You can download Red Hat Enterprise Linux 6 evaluation software from the following website:

  https://access.redhat.com/downloads/
- You can access more information about the installation and configuration of SUSE Linux Enterprise Server at the following website:

  http://www.novell.com/documentation/suse.html
- You can download SUSE Linux Enterprise Server 11 evaluation software from the following website:

  http://www.novell.com/products/server/eval.html

7.9.5 VMware vSphere ESXi 4.1

The installation of VMware vSphere ESXi 4.1 on an x3690 X5 with an attached MAX5 requires the installation media of the OS and an available logical drive. It is possible to use a hard drive RAID array or a USB flash drive for the logical drive. The small footprint of the ESXi makes it possible to run ESXi from a USB flash drive, which eliminates the need for a dedicated RAID controller and addresses the reliability problem of a mechanical drive. This way, you avoid the relatively higher investment of SSD drives.

An IBM System x3690 X5 attached to MAX5 allows the use of interleaved NUMA nodes. In order to boot and run ESXi on this system, you must enable the `allowInterleavedNUMAnodes` boot option.

Use the following steps to install ESXi 4.1 on an x3690 X5 server scaled with MAX5:

1. Boot the server using ESXi installation media.
2. At the VMware VMvisor Boot Menu panel, place the cursor on **ESXi Installer**, and then press Tab to edit the boot options.
3. Add the `allowInterleavedNUMAnodes=TRUE` parameter at the boot options line, so that the line reads:

   ```
   > mboot.c32 vmkboot.gz allowInterleavedNUMAnodes=TRUE --- vmkernel.gz ---
   sys.vgz --- cim.vgz --- ienviron.vgz --- install.vgz
   ```

4. Press Enter after doing the modification.

   Figure 7-54 shows the modified boot menu that allows the ESXi installation to proceed with the enabled interleaved NUMA nodes parameter.

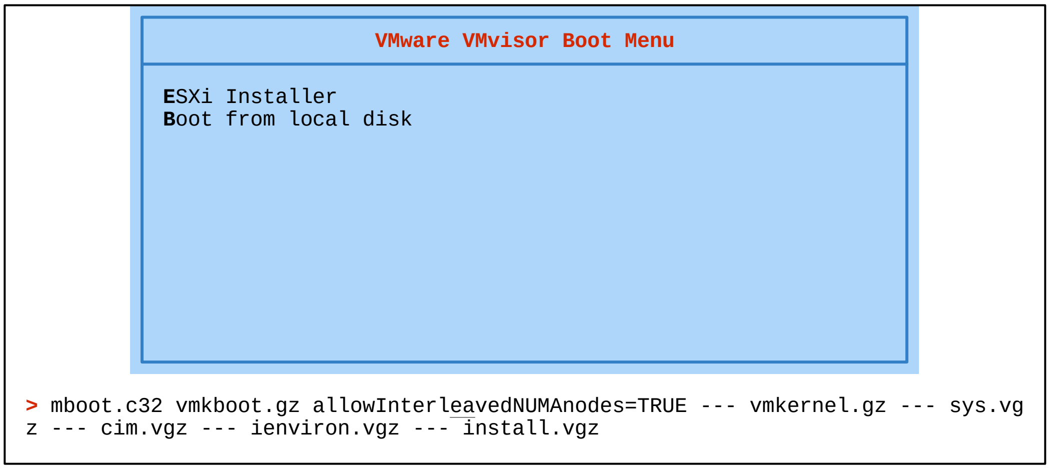


Figure 7-54 Modified VMware VMvisor Boot Menu installation panel

Without this additional parameter, the VMware ESXi 4.1 installation fails with the "`Interleaved NUMA nodes are not supported`" error message, as shown in Figure 7-55.

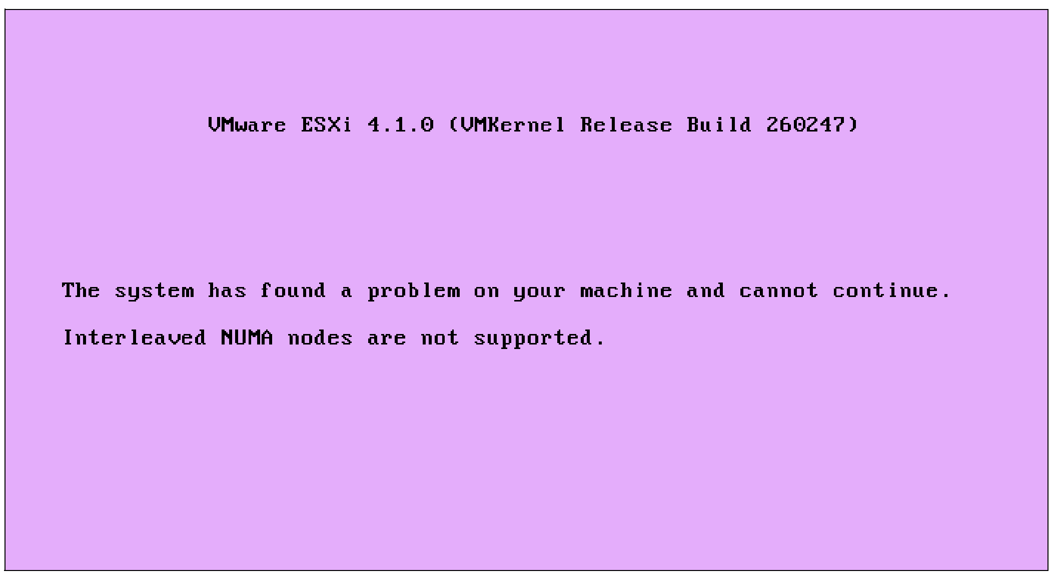


Figure 7-55 Interleaved NUMA nodes are not supported error panel

5. Continue with the ESXi 4.1 installation.
6. To successfully first boot, the `allowInterleavedNUMAnodes` parameter must be set to `TRUE` in the Loading VMware Hypervisor panel. Perform the following steps:
   a. Press Shift+O while the gray bar progresses to add a boot parameter.
   b. Type the following command at the prompt:

      ```
      esxcfg-advcfg -k TRUE allowInterleavedNUMAnodes
      ```

   c. Press Enter to continue the boot process.

   Figure 7-56 on page 360 shows the edited ESXi Loading VMware Hypervisor panel.

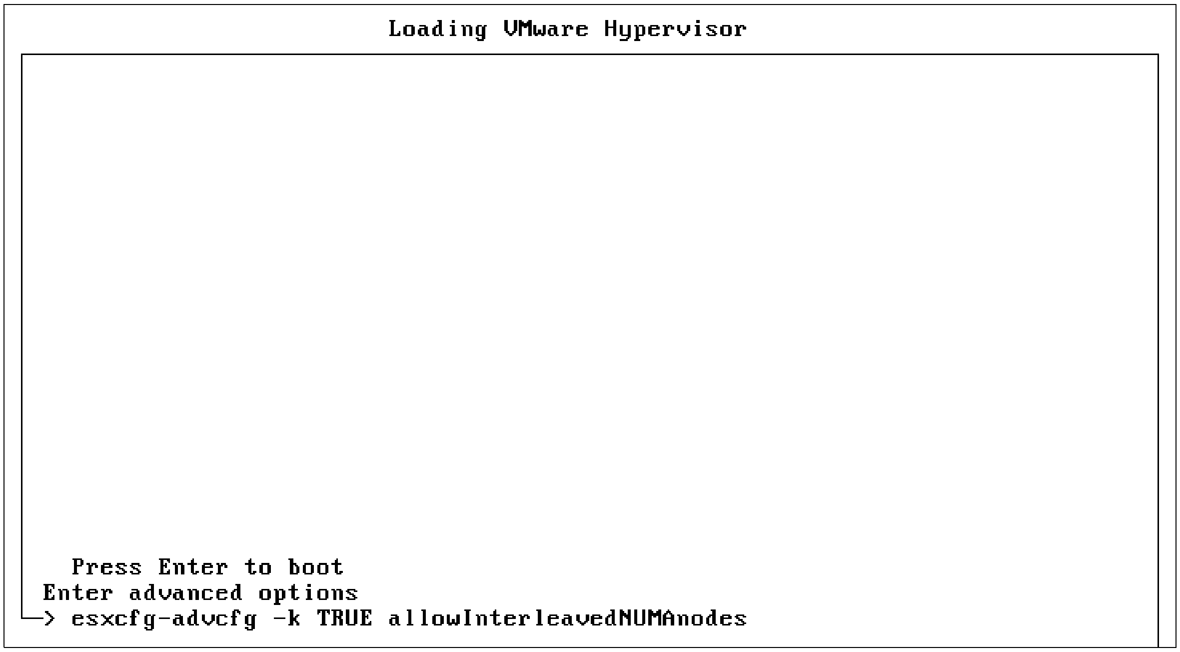


Figure 7-56 Edited ESXi hypervisor loading panel

7. Set the `allowInterleavedNUMAnodes` parameter to `TRUE` in all subsequent system reboots. You are required to have VMware vSphere Client to perform the following steps:

   a. After the ESXi hypervisor is running, connect to the System x3690 X5 using VMware vSphere Client.

   b. Click the **Configuration** tab.

   c. Choose **Advanced Settings** in the Software section.

   d. Click **VMkernel**.

   e. Select the **VMkernel.Boot.allowInterleavedNUMAnodes** check box to enable this parameter and click **OK**.

   Figure 7-57 on page 361 shows the required option to set in the Software Advanced Settings panel.

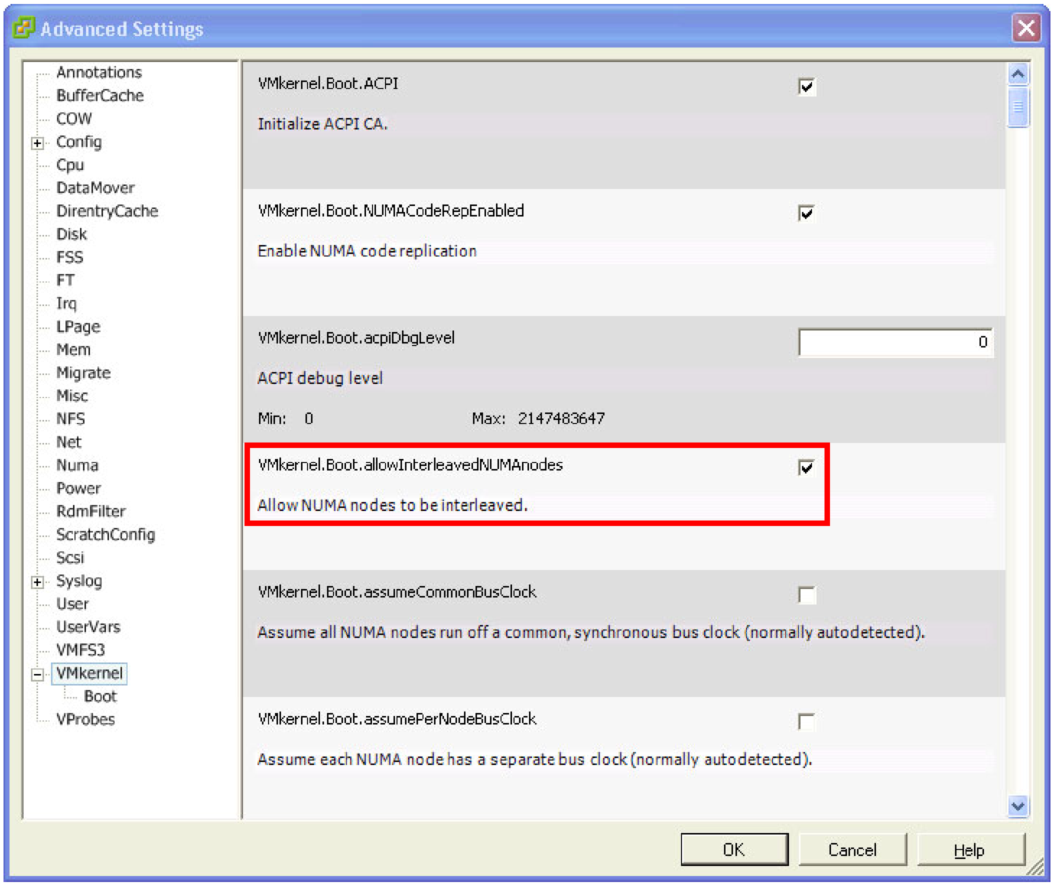


Figure 7-57 VMware ESXi 4.1 Software Advanced Settings panel

In addition to these required steps, use the tips in the following list for the VMware vSphere ESXi installation on the x3690 X5:

- VMware vSphere 4.1 is the last release to support ESX; all subsequent releases of vSphere will only support ESXi. Therefore, we recommend that you use the ESXi hypervisor rather than ESX. The following websites provide more information about this topic, as well as a comparison between ESX and ESXi:
  - `http://www.vmware.com/products/vsphere/esxi-and-esx/`
  - `http://kb.vmware.com/kb/1023990`
- Separate VMware vSphere editions support separate amounts of physical memory. To support the maximum memory capability of the memory-expanded x3690 X5, which is 1 TB, the Enterprise Plus Edition is necessary. Other editions only support up to 256 GB of physical memory.

  You can see a comparison of vSphere editions at this website:

  `http://www.vmware.com/products/vsphere/buy/editions_comparison.html`
- Installing VMware ESXi on System x3690 X5 requires that an equal amount of memory is installed for each system processor. See “MAX5 memory population order” on page 137 for guidelines about memory installation on the x3690 X5 scaled with MAX5.
- ESXi installation fails with the “`NUMA node 1 has no memory`” error message if a processor in the system has no memory installed.

- Create a drive RAID array before installation. The ServeRAID management interface is accessible via UEFI settings, by selecting **System Settings** → **Adapters and UEFI Drivers**.
- When running the ESXi hypervisor from a USB flash drive, plug it in to the internal USB slot to prevent physical disruption. The internal USB slot is located on the x8 low profile PCI riser card.
- ESXi formats all local disks in the server when it is initially booted. Remove any local and SAN-attached disks from the system if they contain important data.
- If ESXi is installed to a USB flash drive, you must add **USB Storage** to the system boot option. Placing it on the top of the boot order list decreases system boot time.
- VMware vSphere 4.1 offers no support for UEFI and therefore is categorized as Legacy OS. Moving **Legacy Only** up in the boot order list decreases system boot time.
- If a USB keyboard stops functioning at any time during the ESXi installation, try plugging it in to a separate USB port.
- You can access more information about the VMware vSphere installation and configuration at the following website:

  `http://www.vmware.com/support/pubs/`

7.9.6 VMware vSphere ESX 4.1

The installation of VMware vSphere ESX 4.1 on an x3690 X5 with an attached MAX5 requires the installation media of the OS and an available logical drive. Unlike the ESXi with its small footprint, you typically install the ESX on a RAID disk array.

An x3690 X5 that is attached to MAX5 allows the use of interleaved NUMA nodes. To boot and run ESX on this system, you must enable the `allowInterleavedNUMAnodes` boot option.

Use the following steps to install ESX 4.1 on the x3690 X5 server scaled with MAX5:

1. Boot the host from the ESX installation media.
2. Press F2 when the installation options panel displays, as shown in Figure 7-58 on page 363.

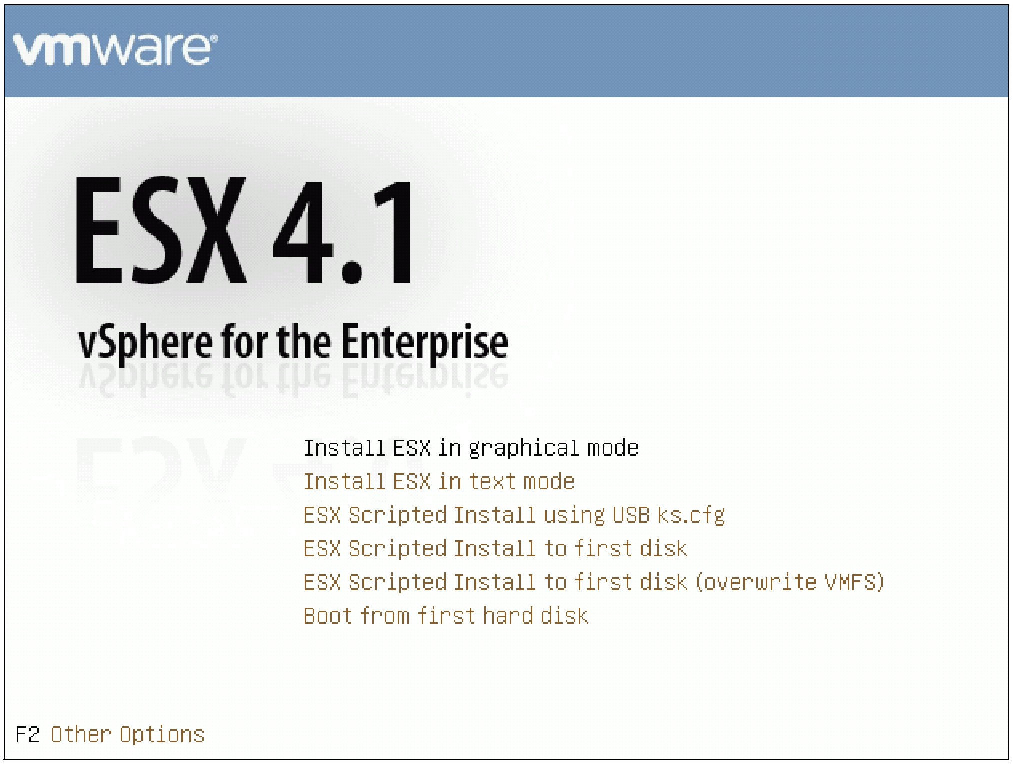


Figure 7-58 ESX installation options window

3. The Boot Options line appears on the panel. Add the `allowInterleavedNUMAnodes=TRUE` parameter at the boot options line, so that the line reads in the following manner:

   ```
   Boot Options initrd=initrd.img debugLogToSerial=1 vmkConsole=false mem=512M
   quiet allowInterleavedNUMAnodes=TRUE
   ```

   The edited result looks like Figure 7-59 on page 364. Press Enter to proceed.

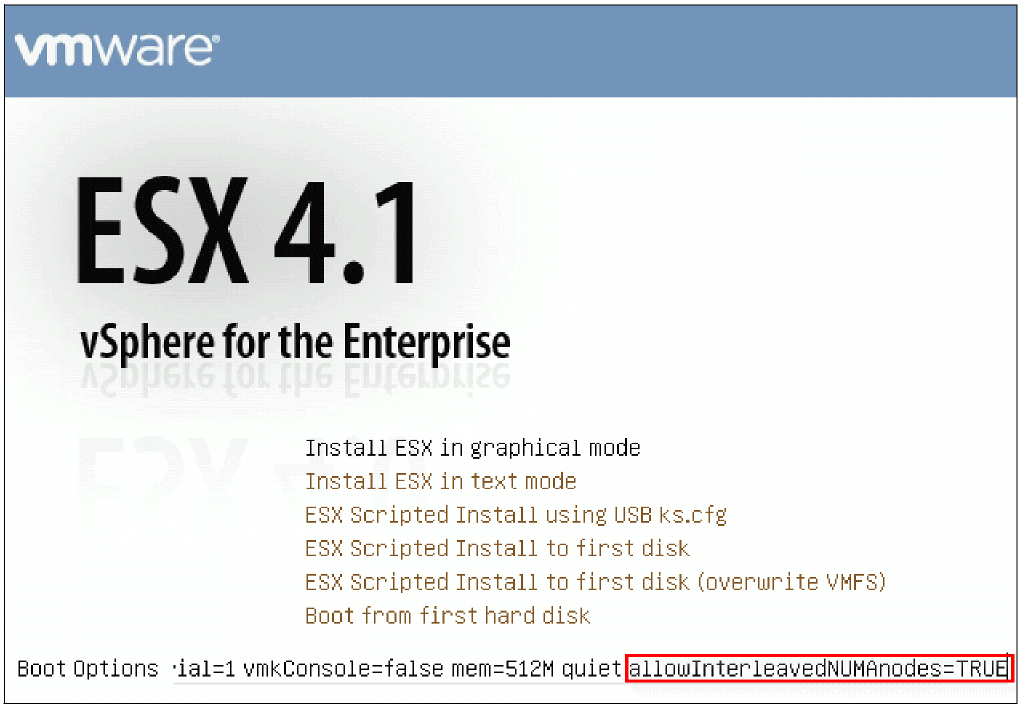


Figure 7-59 Edited ESX installation Boot Options

4. Proceed with ESX installation until you get to the Setup Type page, as shown in Figure 7-60. Click **Advanced setup**, clear **Configure bootloader settings automatically (leave checked if unsure)**, and click **Next**.

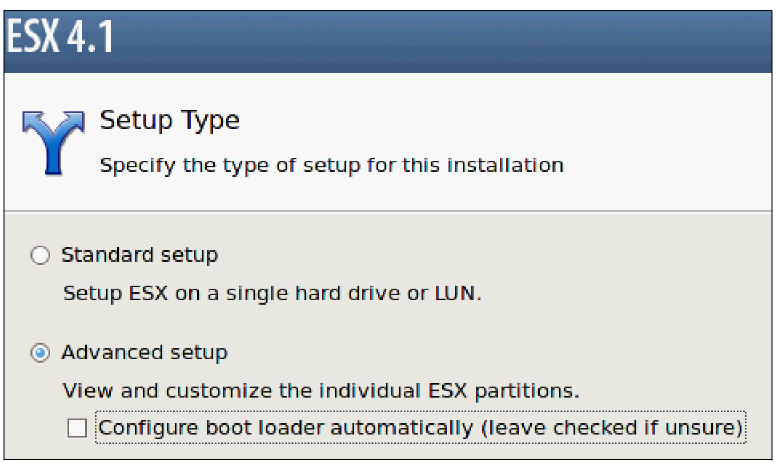


Figure 7-60 ESX installation Setup Type

5. Continue with the installation until you get to the Set Bootloader Options page, as shown in Figure 7-61 on page 365, and type the following parameter in the Kernel Arguments field:

```
allowInterleavedNUMAnodes=TRUE
```

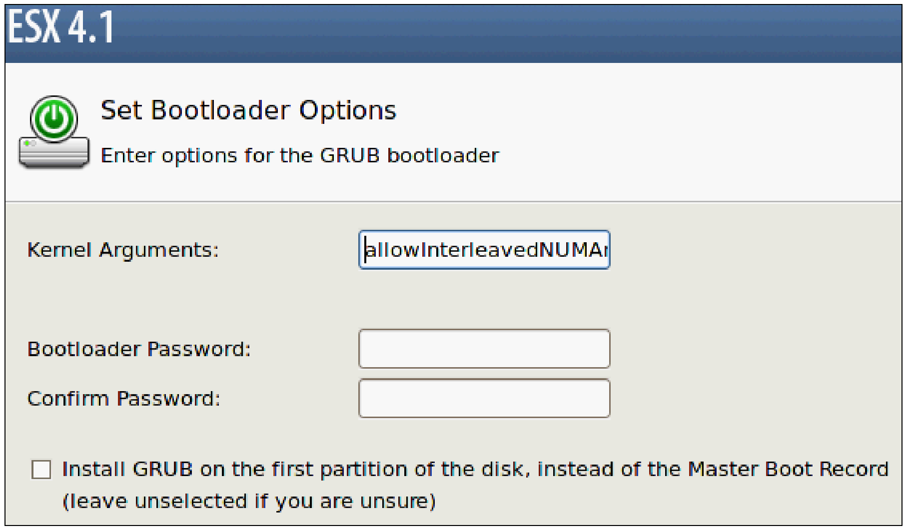


Figure 7-61 Set Bootloader Options

6. Continue with the ESX installation.

7.9.7 Downloads and fixes for the x3690 X5 and MAX5

It is common during the support lifetime of a product that updates are released to provide users with enhanced capabilities, extended functions, and problem resolutions. Most of the updates are in the form of firmware, drivers, and OS patches.

We recommend that you perform a scheduled review of available updates to determine if they are applicable to your user production environment.

Server firmware

Software that resides on flash memory and controls the lower-level functions of the server hardware is known as server *firmware*. An IBM System x server, such as the x3690 X5, can run a number of types of firmware that are in charge of the server components.

The following types of firmware make up the primary firmware for the IBM System x3690 X5:

- Unified Extensible Firmware Interface (UEFI)
- Integrated Management Module (IMM)
- Field-Programmable Gate Array (FPGA)
- Preboot Dynamic System Analysis (DSA)

Other firmware controls each server component and corresponds to the manufacturer and model of the device, such as firmware for NIC, RAID controller, and so on.

IBM provides the firmware updates. You can download the firmware updates from the IBM website, including proven firmware from other manufacturers to apply on IBM systems.

We describe several methods of performing firmware updates on IBM eX5 servers in 9.10, “Firmware update tools and methods” on page 509.

Tip: It is a recommended practice to update all System x firmware to the latest level prior to performing OS and application installation.

Useful links:

- IBM *Bootable Media Creator (BoMC)* is a tool that simplifies the IBM System x firmware update process without an OS running on the system. You can obtain more information about this tool at the following website:

  http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-BOMC

- You can access the IBM System x3690 X5 firmware and drivers quickly at the following website:

  http://www.ibm.com/support/fixcentral/systemx/quickorder?parent=ibm~Systemx3690X5&product=ibm/systemx/7148&&platform=All&function=all&source=fc

- The following website contains a guide that describes firmware update best practices for UEFI-based IBM System x servers:

  http://ibm.com/support/entry/myportal/docdisplay?lndocid=MIGR-5082923

Device drivers

Device drivers are software that controls hardware server components on the OS level. They are specific to the OS version, so critical device drivers are included with the installation media.

Device driver updates are provided by IBM, OS vendors, and component device vendors. They are primarily downloadable from each company's support website.

Whenever possible, we recommend acquiring tested and approved driver updates from IBM.

You can get more information at these useful links:

- The Windows Server installation process using IBM ServerGuide performs IBM driver updates after the OS installation is complete. You can access this tool at the following website:

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

- IBM UpdateXpress is a tool that simplifies the IBM System x firmware and driver update running a supported OS. You can access more information about this tool at the following website:

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

- You can access the IBM System x3690 X5 firmware and drivers quickly at the following website:

  http://www.ibm.com/support/fixcentral/systemx/quickorder?parent=ibm~Systemx3690X5&product=ibm/systemx/7148&&platform=All&function=all&source=fc

Operating system updates, fixes, and patches

The performance and reliability of a scaled x3690 X5 are tightly related to the OS running on it. IBM supports an assortment of modern and widely used OSs that are capable of utilizing the system's potential. Each vendor supports its OS by releasing updates, fixes, and patches that provide enhanced functionality and fixes to known problems. Certain updates, fixes, and patches are only applicable to certain configurations, but other updates, fixes, and patches apply to all configurations. Each OS vendor's support website has extensive information about these updates, fixes, and patches.

System update resources

Table 7-6 on page 367 lists important resources for system updates with the available support in each website.

Table 7-6 Internet links to support and downloads

Vendor	Product	Address	Available support
IBM	Systems support	http://www.ibm.com/systems/support/	Docs, firmware, and drivers
IBM	x3690 X5 support	http://www.ibm.com/support/entry/portal/Overview/Hardware/Systems/System_x/System_x3690_X5	
IBM	ServerGuide	http://www.ibm.com/support/entry/portal/docdisplay?lndocid=SERV-GUIDE	Installation tool
IBM	UpdateXpress	http://www.ibm.com/support/entry/portal/docdisplay?lndocid=SERV-XPRESS	Firmware and driver update tool
IBM	Bootable Media Creator	http://www.ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-BOMC	Firmware update tool
Microsoft	Windows Server	http://support.microsoft.com/ph/14134	Docs, drivers, and OS update
Red Hat	RHEL	https://www.redhat.com/support/	Docs, drivers, and OS update
Novell	SLES	http://www.novell.com/support/	Docs, drivers, and OS update
VMware	vSphere	http://downloads.vmware.com/d/	Docs, drivers, and OS update

7.9.8 SAN storage reference and considerations

The IBM System x3690 X5 with its MAX5 memory expansion capability is considered a high-end server in the IBM System x product line. Target workloads for this design include virtualization and database applications. In both designations, it is typical for the user to attach SAN storage to the server for data storage.

SAN storage attachment

A storage area network (SAN) enables a storage system to be accessible by separate servers as though it is a local disk. This SAN approach provides easier centralized management of the storage system. The following list describes the typically used SAN protocols, each with their own characteristics:

- Fibre Channel (FC)

  This is the most prominent interface that is used for SAN connectivity in mission-critical data centers. It has proven reliability and performance in highly demanding applications. Storage data communication uses FC protocol over fibre cable. Typical requirements for this configuration require FC HBA on the servers and FC SAN switches connecting to the storage system with the available FC interface (IBM System Storage DS3400, DS4000 series, DS5000 series, and so on).

- Internet SCSI (iSCSI)

  This interface is often considered as a more affordable alternative to FC because storage data communication uses the TCP/IP protocol over the standard network interface. However, when paired with fast 10 Gb Ethernet, iSCSI has proven to be able to compete with FC performance. Typical requirements include iSCSI HBA or standard Ethernet port with software iSCSI initiator and network infrastructure. Attached storage systems require an available iSCSI interface (IBM System Storage DS3300, DS4000 series, DS5000 series, and so on.)

- Fibre Channel over Ethernet (FCoE)

  This relatively new interface was introduced to provide an easier method for FC storage environments to take advantage of the high performance 10 Gb Ethernet interface. Data communication uses FC protocol over a 10 Gb Ethernet interface. System setup for FCoE requires a Converged Network Adapter (CNA) on the servers to pass both network and FC storage data via 10 Gb network infrastructure. The attached storage system, for example, IBM System Storage DS3400, DS4000 series, DS5000 series, and so on, requires the available FC interface.

  FCoE competes with iSCSI to be the predominant SAN protocol that works on the Ethernet interface.

- Serial Attached SCSI (SAS)

  For smaller environments not requiring complex SAN configuration, using a SAS cluster is a more cost-effective alternative to an FCoE or iSCSI. Although performance, manageability, and distance are limited compared to the other solutions, servers can access their individual portions of the storage system. Typical requirements are the SAS HBA on the servers and a storage system with a SAS interface, for example, IBM System Storage DS3200 and others.

Booting from SAN

The implementation of the SAN storage environment presents the possibility of running the OS from the SAN. This feature offers several advantages:

- Reduces downtime and increases host application availability. If a server fails, it can be replaced with another server with minimal effort.
- Utilizes SAN and System Storage advanced features. Most storage and SAN solutions offer better features compared to using local disk.
- Improves system and host reliability by removing mechanical parts from the server.

> **Tip:** Prepare the hardware infrastructure and cabling before performing a SAN boot OS installation, and make sure to provide only one path to the storage during the process. Use the proper zoning configuration or disconnect the physical cable to provide only one path to the storage during the process.

IBM Redbooks references

In regard to System x3690 X5 connectivity with SAN storage, many IBM Redbooks publications are available. The following books describe IBM System Storage products and their various implementations, including with the IBM System x product lines:

- *IBM System Storage Solutions Handbook*, SG24-5250

  This book provides overviews and pointers for information about the most current IBM System Storage products.

  `http://www.redbooks.ibm.com/abstracts/sg245250.html`

- *Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*, SG24-6116

  This book consolidates critical information while also covering procedures and tasks that you are likely to encounter on a daily basis when implementing an IBM/Brocade SAN.

  `http://www.redbooks.ibm.com/abstracts/sg246116.html`

- *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363

  This book represents a compilation of best practices for deploying and configuring IBM Midrange System Storage servers, which include the DS4000 and DS5000 family of products.

  http://www.redbooks.ibm.com/abstracts/sg246363.html

- *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065

  This book introduces the IBM System Storage DS3000, providing an overview of its design and specifications, and describing in detail how to set up, configure, and administer it.

  http://www.redbooks.ibm.com/abstracts/sg247065.html

- *Implementing an IBM/Cisco SAN*, SG24-7545

  This book consolidates critical information while discussing procedures and tasks that are likely to be encountered on a daily basis when implementing and IBM/Cisco SAN.

  http://www.redbooks.ibm.com/abstracts/sg247545.html

- *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659

  This book describes the concepts, architecture, and implementation of the IBM XIV Storage System.

  http://www.redbooks.ibm.com/abstracts/sg247659.html

- *IBM Midrange System Storage Hardware Guide*, SG24-7676

  This book consolidates, in one document, detailed descriptions of the hardware configurations and options that are offered as part of the IBM Midrange System Storage servers, which include the IBM System Storage DS4000 and DS5000 families of products.

  http://www.redbooks.ibm.com/abstracts/sg247676.html

7.10 Failure detection and recovery

A memory-expanded System x3690 X5 connects to the MAX5 via a QPI connection. Therefore, a problem with the MAX5 is equivalent to a problem with the system itself. System alerts and reporting on the x3690 X5 will respond to failures on the MAX5.

7.10.1 System alerts

On the system hardware level, use the System x3690 X5 alerting method to analyze failure on the MAX5. Use the following tools:

- Light path diagnostics (LPD)

  Light path diagnostics is a system of LEDs on the various external and internal components of the server. When an error occurs, LEDs are lit throughout the server. By viewing the LEDs in a particular order, you can identify the source of the error.

- Integrated Management Module (IMM)

  The System x3690 X5 IMM informs you of the current system status at the System Status panel, as shown in Figure 7-62 on page 370. Use IMM to investigate server conditions when an error or abnormal status is observed.

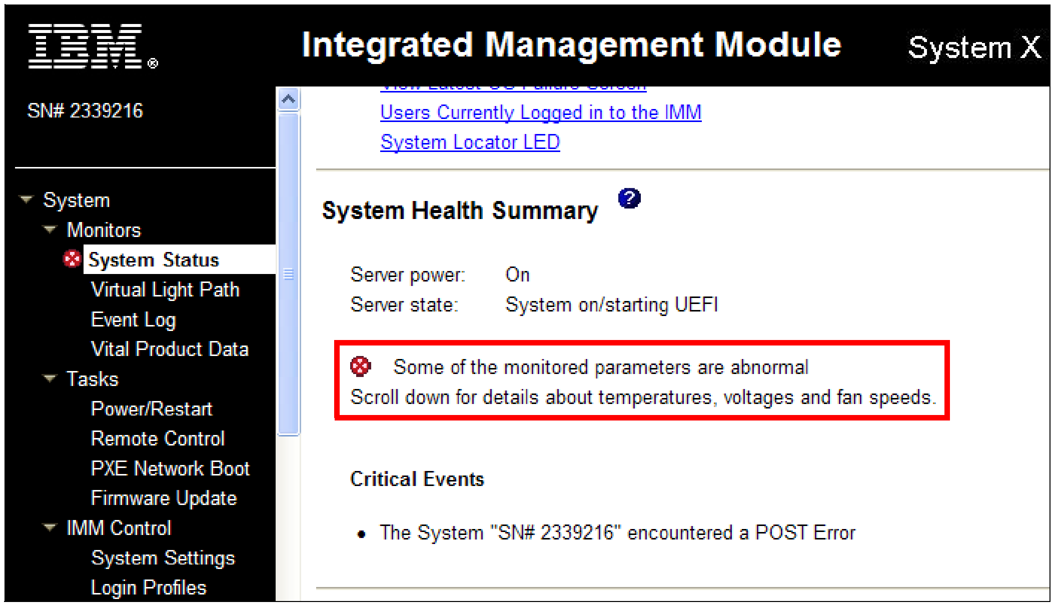


Figure 7-62 Example of IMM error status

- System event logs

  System event logs in UEFI and IMM maintain histories of detected error events.

- IBM Dynamic System Analysis (DSA)

  IBM DSA is a tool to collect and analyze complete system information to aid in diagnosing system problems. Among other things, DSA result files contain information about hardware inventory, firmware level, and system event logs that assist in system problem determination.

  You can access the DSA application and user guide at the following website:

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

> **Tips:**
>
> - Use DSA *Portable* to run the application without having to install it.
> - Ensure that the Intelligent Platform Management Interface (IPMI) driver and RAID Manager agent are installed in the system prior to running DSA in order to get a complete result file.
> - The DSA result file is stored in `C:\IBM_Support\` for Windows and in `/var/log/IBM_Support/` for Linux.

- IBM Electronic Service Agent (ESA)

  IBM ESA performs system monitoring and automatic problem reporting to IBM when a hardware error is detected. This no-charge software can run as stand-alone software or as an extension of IBM Systems Director.

  See 9.6, “IBM Electronic Services” on page 493 for more information about IBM ESA.

- *Problem Determination and Service Guide*

  This guide describes the diagnostic tests, troubleshooting procedures, and an explanation of error messages and error codes to solve many problems without needing outside assistance.

You can download the *Problem Determination and Service Guide* for the IBM System x3690 X5 from the following website:

http://www.ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5085205

7.10.2 System recovery

In the unlikely event of a MAX5 power loss or severe failure that prevents the unit from functioning, the memory-expanded x3690 X5 powers down. Perform the following steps to immediately resume the x3690 X5 system operation while the MAX5 is being repaired:

1. Remove the power to the x3690 X5 and the MAX5.
2. Unplug the QPI scalability cable and the MAX5 for repair.
3. Plug the x3690 X5 power back in.
4. Allow the IMM to boot and then power on the server.

After the MAX5 expansion is functioning properly, schedule downtime and reattach it to the x3690 X5 using the QPI scalability cable. Follow the guidelines for the MAX5 expansion attachment in 7.4, “MAX5 considerations” on page 311.

Important: The attachment and detachment of MAX5 require that the system is down and that the power is removed from the System x3690 X5.

8

IBM BladeCenter HX5

In this chapter, we take a closer look at the three scalable configuration options that are offered with the HX5. We describe the requirements that must be met for each of the scaling options with reference to hardware and firmware. We provide recommendations for hardware component placement, Unified Extensible Firmware Interface (UEFI) settings, and setup through the Advanced Management Module (AMM). We also give a brief overview of the operating system (OS) installation.

This chapter contains the following topics:

Figure 8-1 shows the three scalable configurations of the HX5.

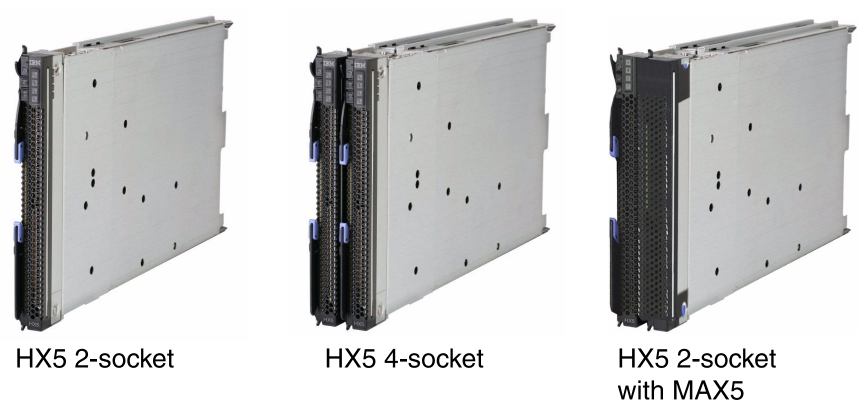


Figure 8-1 The scalable configurations of the IBM BladeCenter HX5

8.1 Before you apply power for the first time after shipping

Before you begin populating your server with processors, memory, and expansion adapters, and before you install an operating system, perform the recommendations that are described in the following sections:

- 8.1.1, "Verifying that the components are securely installed" on page 374
- 8.1.2, "Clearing CMOS memory" on page 375
- 8.1.3, "Verifying the server boots before adding options" on page 376

8.1.1 Verifying that the components are securely installed

Perform the following tasks to ensure that all of the electrical components in your server have proper connectivity:

- Inspect heat sinks to ensure that they are secure.
- Verify that the dual inline memory modules (DIMMs) are mounted in the correct locations and are fully plugged in with their retain clips in the closed position.
- Inspect all expansion adapters to ensure that they are securely plugged into their slots.
- Perform a visual inspection of the system board and all components and ensure that there are no bent pins, loose expansion cards, or loose connectors.

> **Power sharing cap:** If installing a MAX5, ensure that the power sharing cap (as shown in Figure 8-2) is removed when adding the MAX5 and is not discarded. If the MAX5 is ever removed (in order to form a stand-alone system or a 2-node system), the power sharing cap is required for the server to boot to power-on self test (POST). If not using a MAX5, ensure that the power sharing cap is in place and securely pushed down.

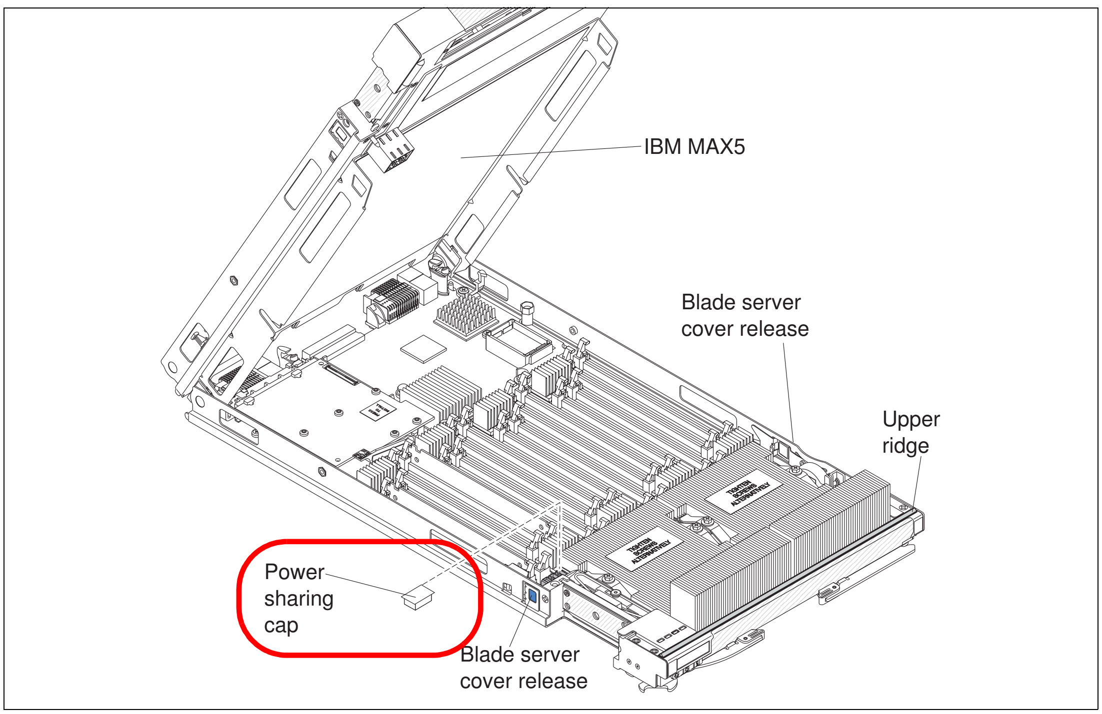


Figure 8-2 If installing an HX5, remove the power sharing cap, but do not discard it

8.1.2 Clearing CMOS memory

When a server is shipped from one location to another location, you have no idea what the server was exposed to. For all you know, it might have been parked next to a large magnet or electric motor and everything in the server that stores information magnetically has been altered, including the CMOS memory. IBM does not indicate on the shipping carton that magnetic material is enclosed, because the information is readily recoverable.

Booting the server to the F1 system configuration panel and selecting **Load Default Settings** restores the default values for the items that you can change in the configuration. This option does not change the settings of registers that are used by the Integrated Management Module (IMM) and the UEFI. These registers define the system state of the server. When they become corrupt, the server can display the following symptoms:

- Fail to power on.
- Fail to complete POST.
- Turn on amber light path diagnostic lights that describe conditions that do not really exist.
- Reboot unexpectedly.
- Fail to detect all of the installed CPU, memory, PCIe adapters, or physical disks.

You cannot restore or modify these internal registers to their defaults by using the F1 system configuration panel; however, you can restore these internal registers to their defaults by clearing the CMOS memory.

There are two key ways to clear the CMOS. One method is by jumper and the second method is by removing the CMOS battery. We describe each method in the following sections.

Jumper method

Use the following procedure to clear the CMOS using the jumper method:

1. Power-cycle the real-time clock (RTC). Locate the switch block SW2 in the middle of the system board at the extreme rear of the server, as shown in Figure 8-3. You must remove all of the expansion cards to reach the jumper block.

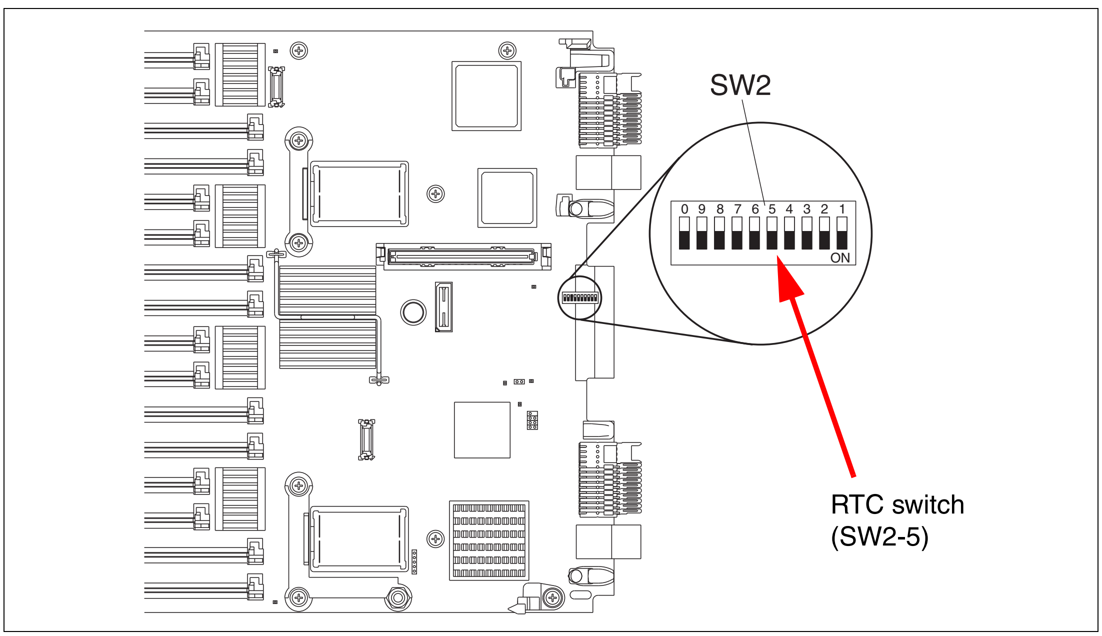


Figure 8-3 Switch block SW2 location

2. The numbers on the switch block represent the OFF side of the switch. Locate switch 5, which is the RTC switch. To clear the CMOS, remove the plastic sheet from the jump block, as shown in Figure 8-3 on page 375. Set switch 5 on switch block SW2 to the ON position for 5 seconds and then return it back to the OFF position.

CMOS battery removal option

Use the following procedure to clear the CMOS by removing the CMOS battery:

1. Remove the ac power from the server.
2. Locate the CMOS battery, as highlighted in Figure 8-4.

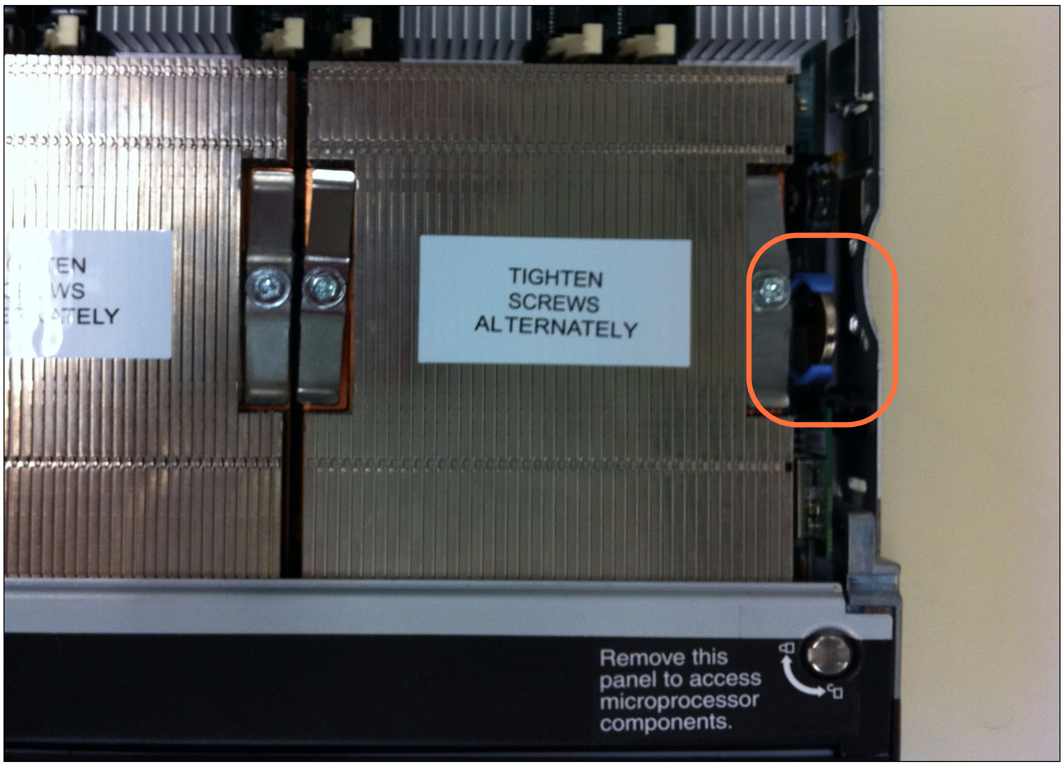


Figure 8-4 CMOS battery location

3. Use your finger to pry up the battery on the side toward the second processor. It is easy to lift the battery out of the socket.

> **Light path diagnostics:** The light path diagnostics (LPD) lights are powered from a separate power source (a capacitor) than the CMOS memory. LPD lights remain lit for a period of time after the ac power and the CMOS memory battery have been removed.

4. Wait 30 seconds and then insert one edge of the battery, with the positive side up, back into the holder.
5. Push the battery back into the socket with your finger and it will clip back into place.

8.1.3 Verifying the server boots before adding options

When you have ordered options for your server that have not yet been installed, it is a good idea to make sure that the server will complete the POST properly before you start to add the options. Performing this task makes it easier to compartmentalize a potential problem with an installed option rather than having to look at the entire server to try to find a good starting point during problem determination.

8.2 Planning to scale: Prerequisites

As described in 5.8, “Scalability” on page 188, the HX5 blade architecture allows for three scalable configurations:

- Single-node server: A single HX5 server with two processor sockets
- Two-node server: Two HX5 servers connected to form a single image 4-socket server
- HX5+MAX5: A single HX5 server with two processor sockets, plus a MAX5 memory expansion blade that is attached to the server.

When making the decision to scale with either a MAX5 or by adding a second blade, you must meet several key prerequisites before proceeding:

- Processors supported and requirements to scale
- Minimum memory requirement
- Required firmware of each blade and the AMM

We describe these prerequisites in the following sections.

8.2.1 Processors supported and requirements to scale

Two-node configurations can include four processors (two in each node) or two processors (one in each node). When implementing a 4-processor configuration, all processors must be identical and can be any of the supported Intel Xeon 7500 processors that are listed in 5.9, “Processor options” on page 192. Intel Xeon 6500 series processors are not supported in 4-processor configurations. When implementing a 2-node configuration with one processor in each node, both processors must be identical; however, any of the Xeon 7500 or Xeon 6500 series processors that are listed in 5.9, “Processor options” on page 192 are supported.

If you install two Xeon 6500 processors in either a 1-node or 2-node configuration and then later decide to upgrade to four processors in a 2-node configuration, you must replace the Xeon 6500 processors with Xeon 7500 processors, because Xeon 6500 processors are not designed to scale to 4-way.

Figure 8-5 shows an example of an error that you will see in the error log of the AMM if the processors do not match when trying to scale.

E	Blade_02	11/18/10 08:09:04	0x806f0507	Processor 1, (CPU 1) configuration error
E	Blade_01	11/18/10 08:08:59	0x806f0507	Processor 1, (CPU 1) configuration error

Figure 8-5 Processor mismatch error when trying to scale a 2-node

The HX5+MAX5 configurations support all of the processors that are listed in 5.9, “Processor options” on page 192.

8.2.2 Minimum memory requirement

For a 2-node configuration, the minimum required amount of memory is four DIMMs, two on each blade in slots 1 and 4. However, we do not recommend this amount of memory for performance reasons.

Table 8-1 shows the recommended installation guide to install the memory in a nonuniform memory access (NUMA)-compliant DIMM installation for the HX5 2-node configuration.

Table 8-1 NUMA-compliant DIMM installation for the HX5 2-node

Number of processors	Number of DIMMs	Hemisphere Mode[a]	Processor 1								Processor 2							
			Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer		Buffer	
			DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16
4	8	N	x			x					x			x				
4	16	Y	x			x	x			x	x			x	x			x
4	24	N	x	x	x	x	x			x	x	x	x	x	x			x
4	32	Y	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

a. For more information about Hemisphere Mode and its importance, see 2.3.5, "Hemisphere Mode" on page 26.

For an HX5 with MAX5 configuration, when installing memory in a MAX5, you do not need to follow a specific order. However, we recommend that you install the memory in the blade before populating the memory in the HX5. After you install the memory in the blade, install the memory in the method that is shown in Table 8-2.

Table 8-2 DIMM installation for the MAX5 for IBM BladeCenter

Number of DIMMs	Power domain A								Domain C (½)				Power domain B								Domain C (½)			
	Buffer				Buffer				Buffer				Buffer				Buffer				Buffer			
	DIMM 1	DIMM 2	DIMM 3	DIMM 4	DIMM 5	DIMM 6	DIMM 7	DIMM 8	DIMM 9	DIMM 10	DIMM 11	DIMM 12	DIMM 13	DIMM 14	DIMM 15	DIMM 16	DIMM 17	DIMM 18	DIMM 19	DIMM 20	DIMM 21	DIMM 22	DIMM 23	DIMM 24
2	x							x																
4	x							x					x				x							
6	x							x				x	x				x							x
8	x		x			x		x				x	x				x							x
10	x		x			x		x				x	x		x		x		x					x
12	x		x			x		x		x		x	x		x		x		x			x		x
14	x	x	x			x	x	x		x		x	x		x		x		x			x		x
16	x	x	x			x	x	x		x		x	x	x	x		x	x	x			x		x
18	x	x	x			x	x	x		x	x	x	x	x	x		x	x	x			x	x	x
20	x	x	x	x	x	x	x	x		x	x	x	x	x	x		x	x	x			x	x	x
22	x	x	x	x	x	x	x	x		x	x	x	x	x	x	x	x	x	x	x		x	x	x
24	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

8.2.3 Required firmware of each blade and the AMM

If you plan to have a 2-node HX5 or HX5+MAX5 configuration, be prepared to install the latest versions of all firmware on the blade and the AMM in the chassis. Without the proper versions of firmware, several issues might occur, such as these problems:

- Blades not booting
- Blade or MAX5 not recognized by the AMM
- Incompatibility errors in the AMM, such as the error that is shown in Figure 8-6.

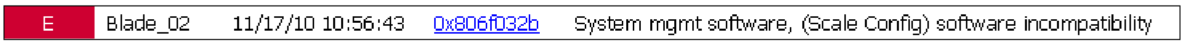
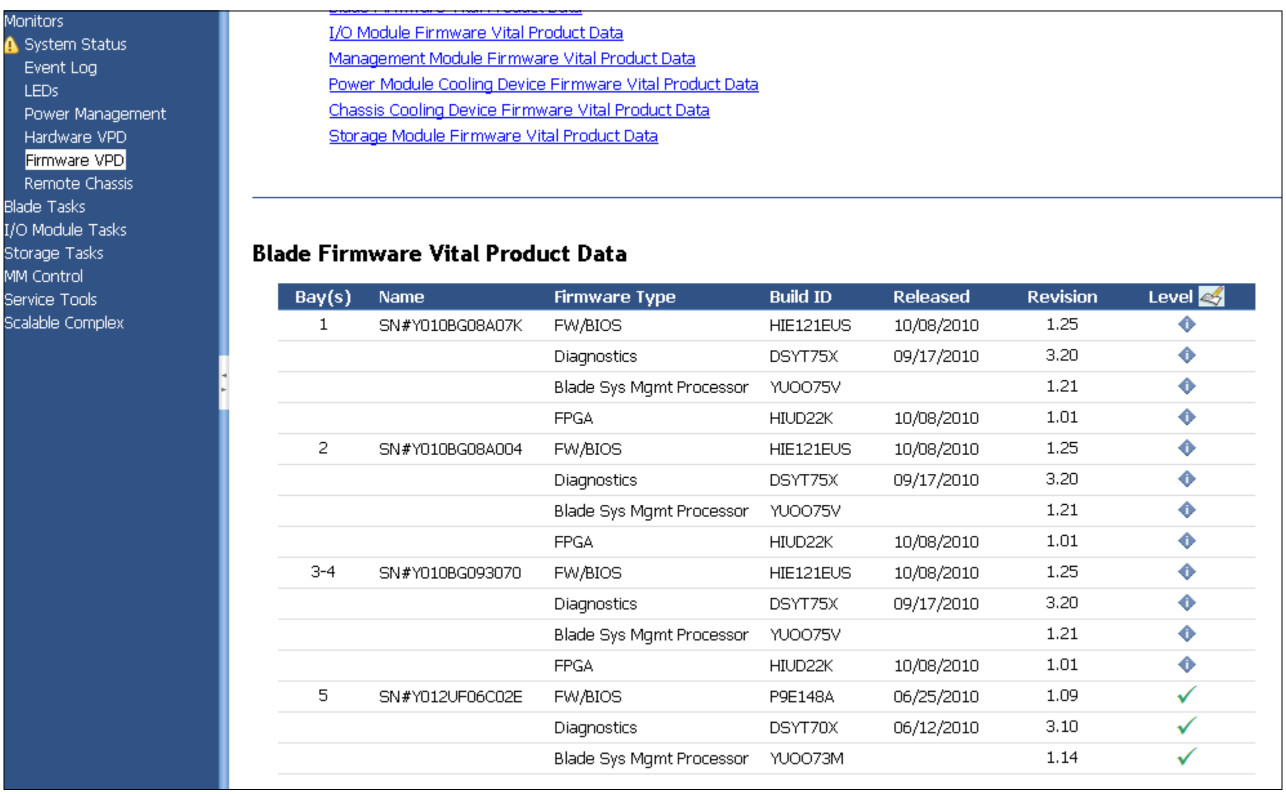


Figure 8-6 Blade incompatibility error when blade firmware is mismatched

Before joining blades together or adding a MAX5, check the firmware versions of both the AMM and the individual blades.

> **MAX5:** MAX5 only has Field Programmable Gate Array (FPGA) and is not a factor.

To locate the current blade and AMM code, you need to log in to the AMM and choose **Monitors** → **Firmware VPD**. Figure 8-7 shows the current firmware versions of each blade.



Bay(s)	Name	Firmware Type	Build ID	Released	Revision	Level
1	SN#Y010BG08A07K	FW/BIOS	HIE121EUS	10/08/2010	1.25	
		Diagnostics	DSYT75X	09/17/2010	3.20	
		Blade Sys Mgmt Processor	YUOO75V		1.21	
		FPGA	HIUD22K	10/08/2010	1.01	
2	SN#Y010BG08A004	FW/BIOS	HIE121EUS	10/08/2010	1.25	
		Diagnostics	DSYT75X	09/17/2010	3.20	
		Blade Sys Mgmt Processor	YUOO75V		1.21	
		FPGA	HIUD22K	10/08/2010	1.01	
3-4	SN#Y010BG093070	FW/BIOS	HIE121EUS	10/08/2010	1.25	
		Diagnostics	DSYT75X	09/17/2010	3.20	
		Blade Sys Mgmt Processor	YUOO75V		1.21	
		FPGA	HIUD22K	10/08/2010	1.01	
5	SN#Y012UF06C02E	FW/BIOS	P9E148A	06/25/2010	1.09	✓
		Diagnostics	DSYT70X	06/12/2010	3.10	✓
		Blade Sys Mgmt Processor	YUOO73M		1.14	✓

Figure 8-7 Blade firmware as shown in the AMM

> **Tip:** In Figure 8-7, `FW/BIOS` refers to the UEFI and `Blade Sys Mgmt Processor` refers to the Integrated Management Module (IMM).

Further down the same Firmware VPD page, the AMM firmware version is also displayed, as shown in Figure 8-8.

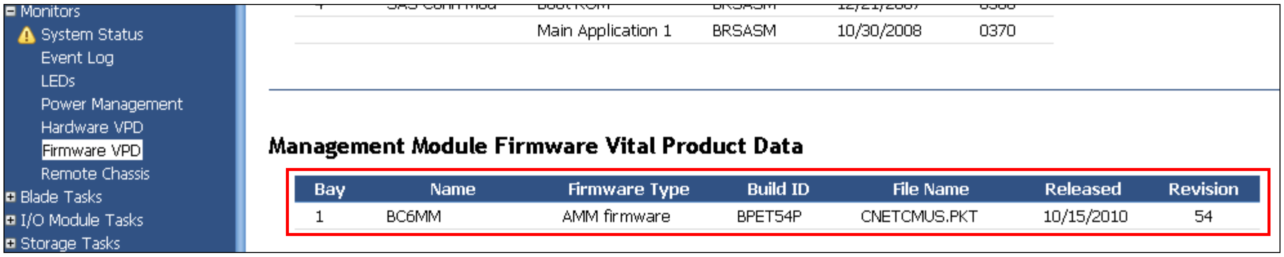


Figure 8-8 AMM firmware version

Table 8-3 shows the minimum firmware levels that the HX5 must be running for a 2-node scaled configuration. It also shows the necessary firmware levels to scale with a MAX5.

Table 8-3 Minimum code levels required before scaling HX5 with either MAX5 or a 2-node scale

	AMM code	UEFI	IMM	FPGA
HX5 single-node	BPET54L	1.00	1.16	1.00
HX5 2-node	BPET54G	1.25	1.20	1.01
HX5 with MAX5	BPET54P	1.25	1.21	1.01

Minimum requirements: The blades and the AMM are required to be, at a minimum, at these code levels in order to scale. However, IBM however recommends that you update to the latest versions, which are available at the following website:

`http://www.ibm.com/support/us/en/`

It is important, in a 2-node scaled configuration, that both HX5s have exactly the same firmware versions for BIOS (UEFI), the blade system management processor (IMM), and FPGA. If they are not at the same version when the QuickPath Interconnect (QPI) wrap card is added, neither blade will boot until the QPI wrap card is removed. See Figure 8-9 on page 380 for an example of the error that you will see.

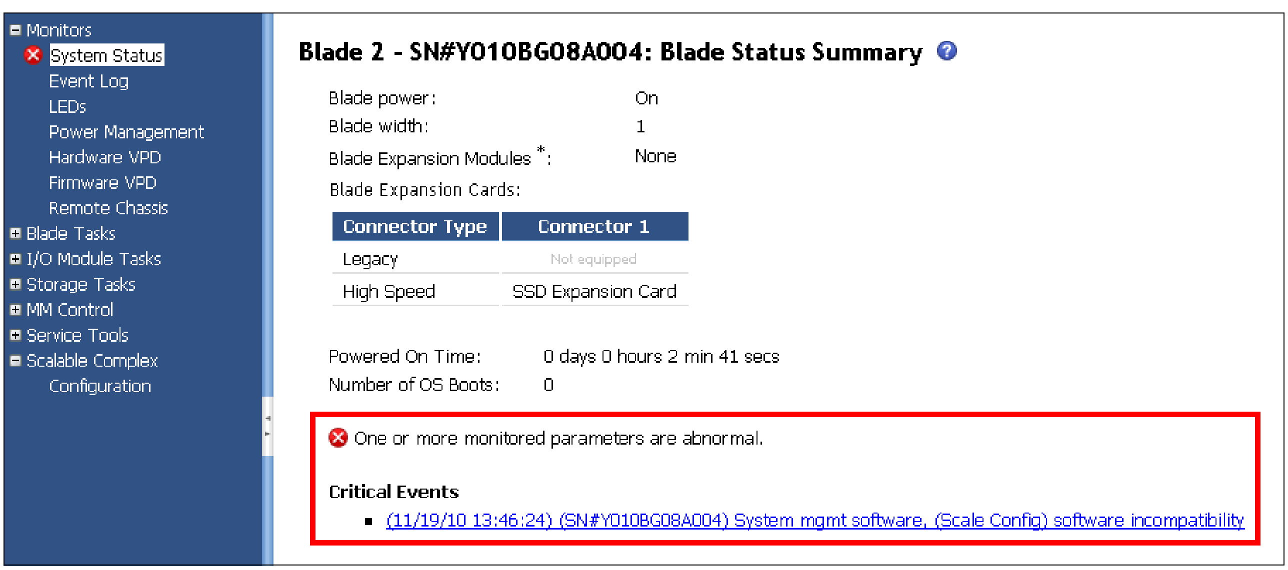


Figure 8-9 Mismatched firmware error as seen from the AMM

Another characteristic of mismatched firmware is the following message when attempting to create the scale in the scalable complex, as shown in Figure 8-10:

`“The complex state data cannot be read at this time. Check the event log to verify the completion of prior complex operations and try again.”`

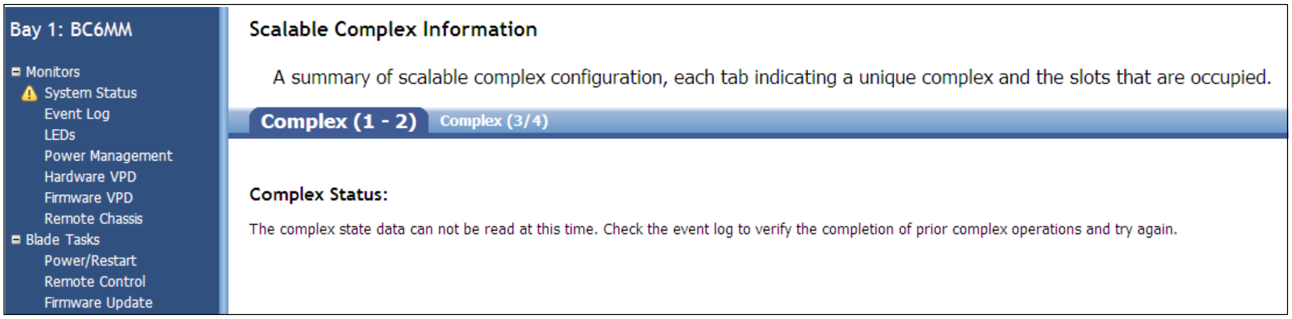


Figure 8-10 Error when trying to create a partition with mismatched firmware on the blade

You must perform the updates on blades individually without the QPI wrap card installed. The AMM might give you the options to flash both blades at one time, as shown in Figure 8-11 on page 381, when the QPI wrap card is installed, but this approach can cause issues and must not be used for the initial setup.

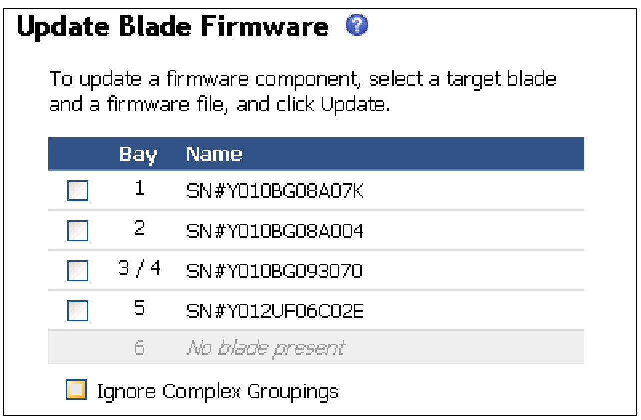


Figure 8-11 Update blade option in blade task: Update Blade Firmware

When the blades are scaled together and operational, you can use any form of updating: Bootable Media Creator, Update Express, AMM, and so on. Both nodes will be updated automatically. Figure 8-12 shows a 2-node scale being updated together from AMM.

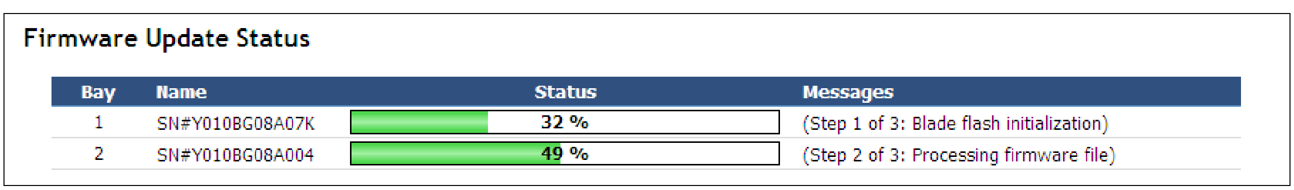


Figure 8-12 Two-node HX5 being updated from AMM

For details about these update methods, see the detailed information about firmware update methods that is available in 9.10, “Firmware update tools and methods” on page 509.

Important tip: When updating a 2-node configuration from the AMM, it is important that you do *not* choose **Ignore Complex Groupings**. This option treats each blade as an individual blade, and both blades might not be updated, which might then lead to the complex not being able to boot. The blades will not boot until the 2-node scalability card is removed and the blades are reflashed individually.

8.3 Recommendations

In this section, we describe the considerations when creating a 2-node scale or adding a MAX5 to an HX5 blade. These settings help you to make decisions for performance and maximum uptime.

8.3.1 Power sharing cap

If installing a MAX5, ensure that you do *not* discard the power sharing cap, as shown in Figure 8-13 on page 383, which is removed when adding the MAX5. If the MAX5 is later removed either to form a stand-alone system or a 2-node configuration, the power sharing cap is required in order for the server to boot to POST. If not using a MAX5, ensure that the power sharing block is in place and securely pushed down.

Power sharing cap: The power sharing cap is sometimes referred to as the *power jumper cap*.

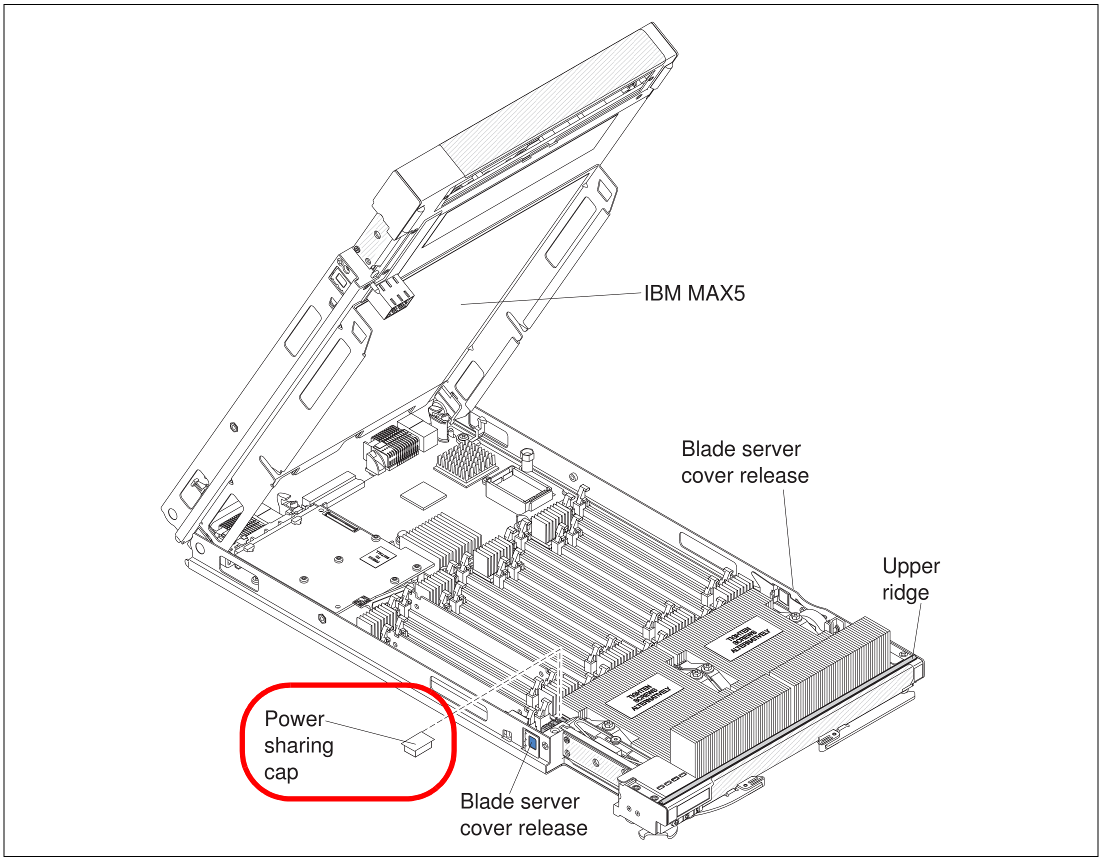


Figure 8-13 Blade showing power sharing cap and location

8.3.2 BladeCenter H considerations

Blade server placement in a chassis is vitally important in balancing the power consumption and ensuring reliability. This section is only relevant if you use a BladeCenter H chassis. If you use a BladeCenter S, this section does not apply.

For maximum redundancy, IBM always recommends that you install all four power supplies in every H chassis. In the BladeCenter H, this configuration is important due to power domains.

In a BladeCenter H, you have two separate power domains:

- Domain 1: Blades 1 - 7
- Domain 2: Blades 8 - 14

With four power supplies, each domain has a redundant pair of power supplies. Not having two power supplies for each domain can cause a loss of blades if the single power supply or its power source fails. If the blades do not power off, they might throttle the processor speed down to a point at which performance is severely degraded.

To determine the current power consumption, use the AMM. Move the blades around until the maximum balance is obtained. Figure 8-14 on page 384 shows the location of the Power Management menu in the AMM.

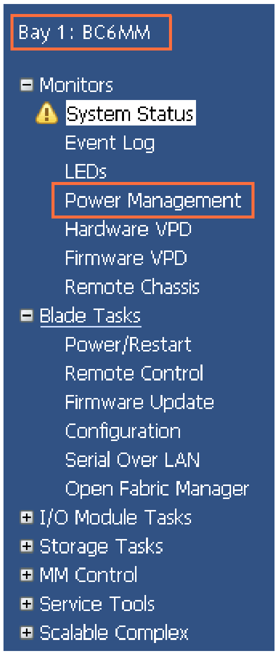


Figure 8-14 Power Management as shown from the AMM

Figure 8-15 shows a breakdown of each domain and the current amount of power that is currently being used for each domain.

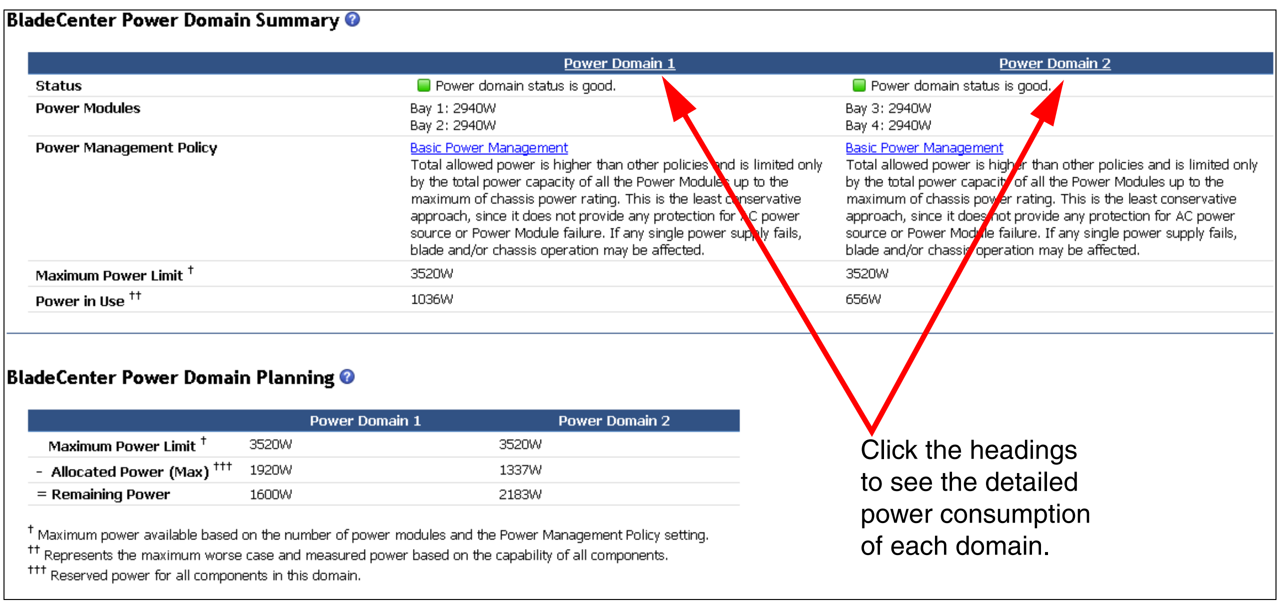


Figure 8-15 Main view when looking at the AMM Power Management option on a BladeCenter H

You can see more details about the power consumption of each domain by clicking each power domain heading, as shown in Figure 8-15. Figure 8-16 on page 385 shows an example of the details that are displayed.

Bay (s)	Status	Module	State	Power In Use	Allocated Power Maximum	Allocated Power Minimum	CPU Duty Cycles
Chassis Components							
		Midplane	On	5W	5W	5W	n/a
1		Media Module	On	5W	5W	5W	n/a
Power Module Cooling Devices							
1		Power Module	On	15W	15W	15W	n/a
2		Power Module	On	15W	15W	15W	n/a
3		Power Module	On	15W	15W	15W	n/a
4		Power Module	On	15W	15W	15W	n/a
Management Modules							
1		BC5MM	On	12W	13W	13W	n/a
2		Advanced Management Module Bay 2 (not present)		0W	8W	8W	n/a
I/O Modules							
1		Ethernet SM	On	22W	23W	23W	n/a
2		Ethernet SM	On	22W	23W	23W	n/a
3		Fibre Channel SM	On	45W	45W	45W	n/a
4		Fibre Channel SM	On	45W	45W	45W	n/a
Blades							
[1]		Blade1-HS22	On	135W	301W	221W	n/a ††
[2]		Blade46-HS22	On	144W	301W	221W	n/a ††
[3]		Blade2-HS22	On	131W	296W	216W	n/a ††
[4]		Blade5-HS22	On	133W	301W	221W	n/a ††
[6]		Blade36	On	133W	193W	151W	(100% ,0%)
[7]		Blade12-HS22	On	138W	301W	221W	n/a ††

Figure 8-16 Power consumption in one power domain

A further consideration is when a 2-node or HX5+MAX5 configuration straddles the two power domains. If you install the primary blade server of a scalable blade complex in blade server bay 7 of a BladeCenter H chassis, the second node or the MAX5 is installed in bay 8. Therefore, the scale is split between two separate power domains.

The following situations can occur if there is a power loss to either power domain, depending on how the scalable blade complex is implemented:

- A loss of power to power domain 1 results in both blade servers in the scalable blade complex going down. If this configuration is a 2-node configuration, you also might get errors on the reboot and the light path diagnostics reporting that a non-maskable interrupt (NMI) error has occurred. This result is normal due to losing processors abruptly. The servers will not come back up until the power is restored or the scalability card is removed.

- If the scalable blade complex is implemented in stand-alone mode, a loss of power to any power domain will still result in both blade servers going down. The blades will not come back up until either the QPI card is removed or the power is restored.

8.4 Local storage considerations and array setup

In an HX5 2-node configuration, you can choose to use the drives on the primary node, the secondary node, or both. In this section, we describe creating an array in the LSI 1064. For drive option and ordering information, see 5.11, “Storage” on page 203.

You can create an array on the primary node, the secondary node, or both. IBM recommends that the operating system is installed on the primary node for maximum performance.

The HX5 offers two methods to create RAID arrays: either via the LSI Setup Utility in UEFI or through the use of IBM ServerGuide.

8.4.1 Launching the LSI Setup Utility

The HX5 offers two ways of launching the LSI Setup Utility.

Using UEFI to launch the LSI Setup Utility

Use the following steps to launch the LSI Setup Utility using UEFI:

1. When the server boots, press F1 when prompted to go to the UEFI menu.
2. In the UEFI System Configuration and Boot Management menu, select **System Settings**, as shown in Figure 8-17.

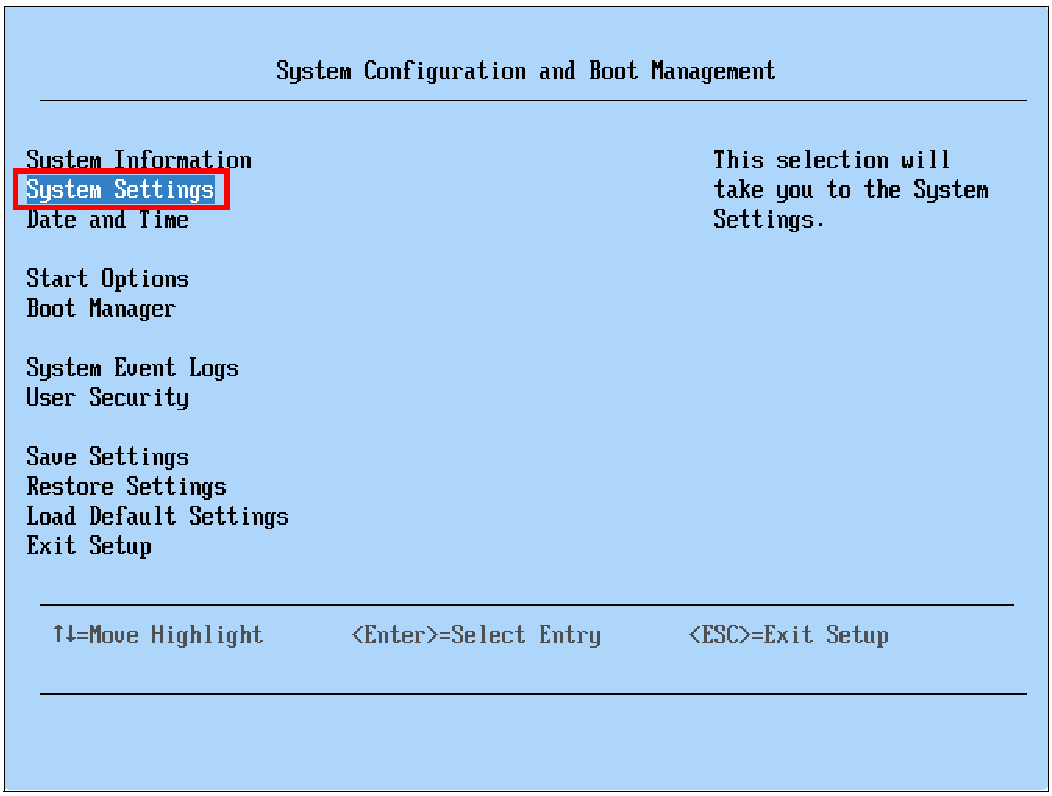


Figure 8-17 UEFI System Configuration and Boot Management menu

3. Select **Adapters and UEFI Drivers** and press Enter.

4. Select **LSI Logic SAS Controller** and press Enter, as shown in Figure 8-18.

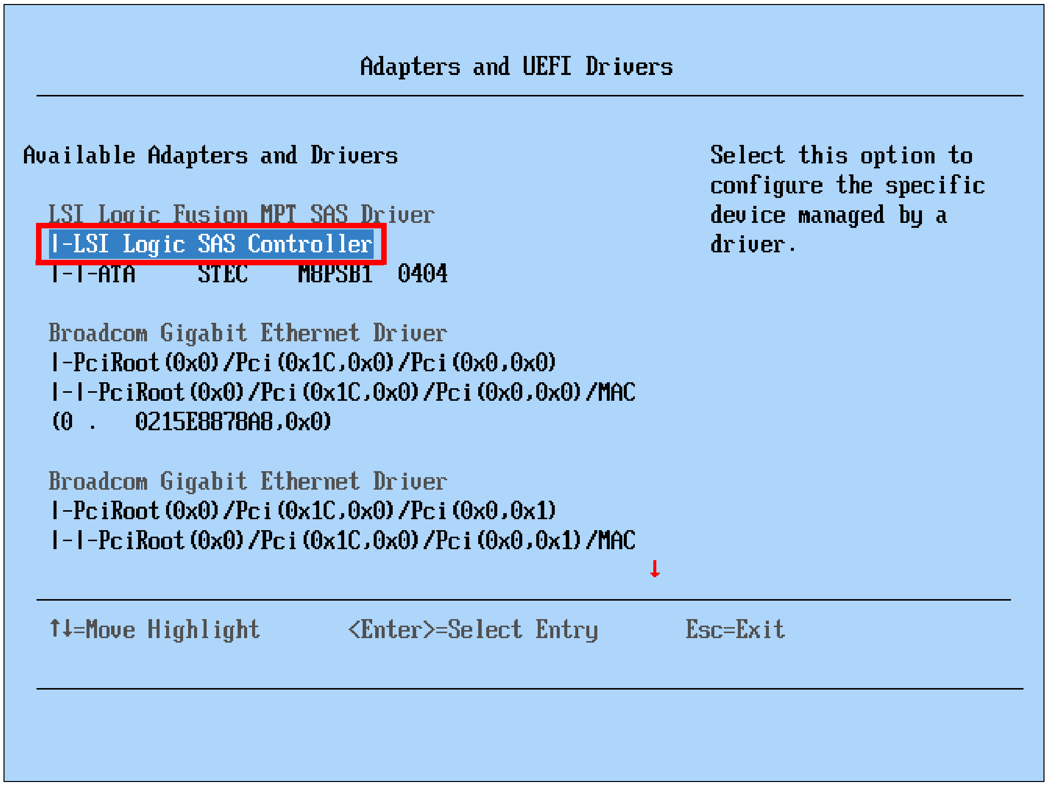


Figure 8-18 Adapter list in UEFI

You can now create an array. We describe creating a RAID-1 array in 8.4.2, “Creating a RAID-1 mirror using the LSI Setup Utility” on page 389.

Enabling Legacy Only mode and pressing Ctrl-C from boot

To see the LSI SAS BIOS option at boot time, you need enable the Legacy Only option in the boot order. Understand that enabling this feature will add additional time to the boot sequence. Follow these steps:

1. When the server performs POST, press F1 to go to the UEFI System Configuration and Boot Management menu.
2. In the UEFI System Configuration and Boot Management menu, select **Boot Manager**, as shown in Figure 8-19 on page 388.

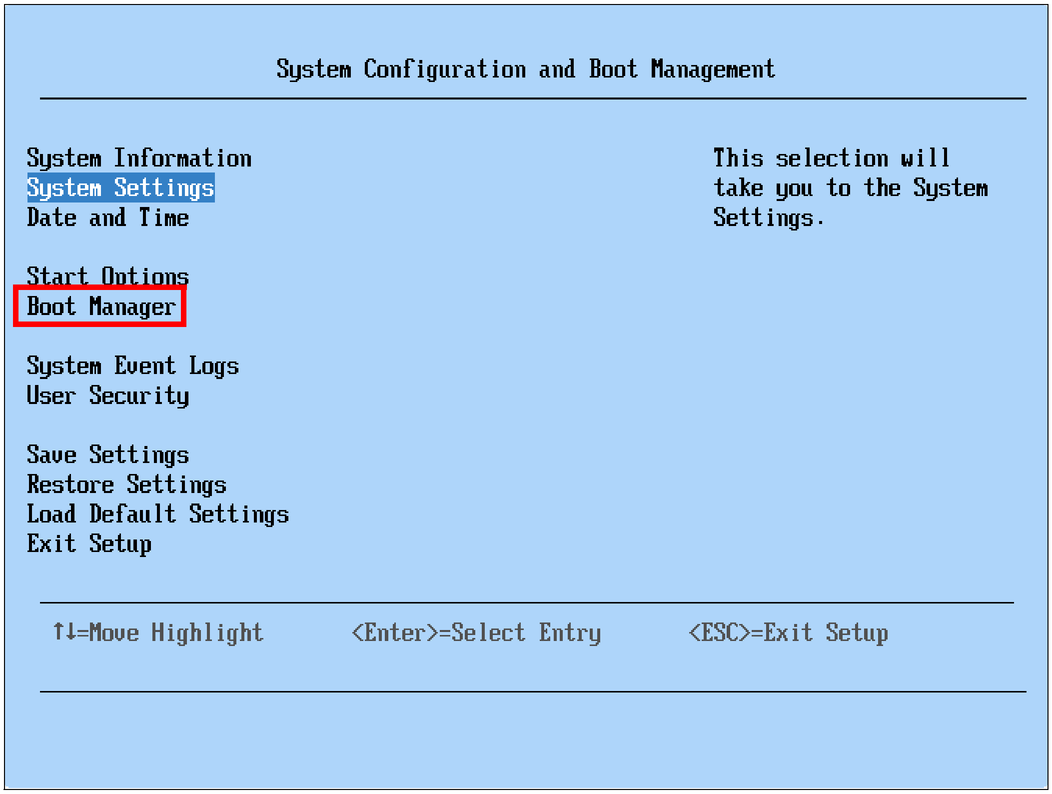


Figure 8-19 UEFI System Configuration and Boot Management menu

3. On the Boot Manager menu, select **Add Boot Option**, as shown in Figure 4 on page 389.

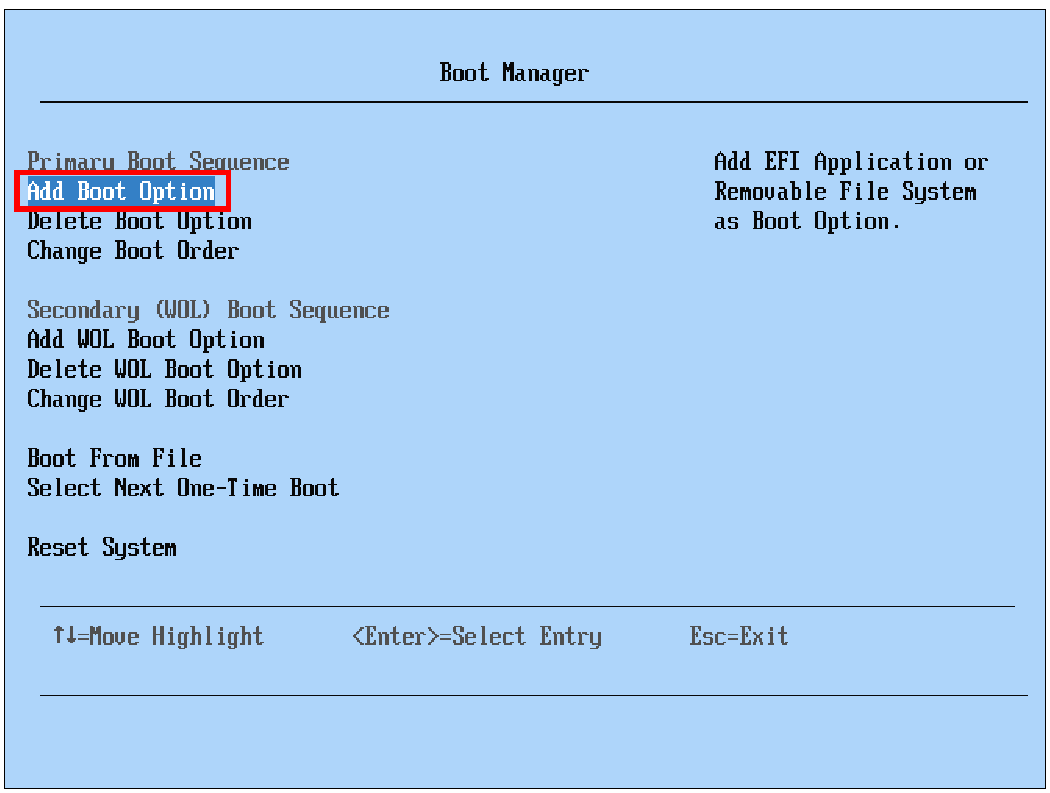


Figure 8-20 UEFI Boot Manager menu

4. Select **Legacy Only**, as shown in Figure 8-21, and press Enter. The selection disappears when selected. If this option is not available, the Legacy Only option might already be part of the boot list.

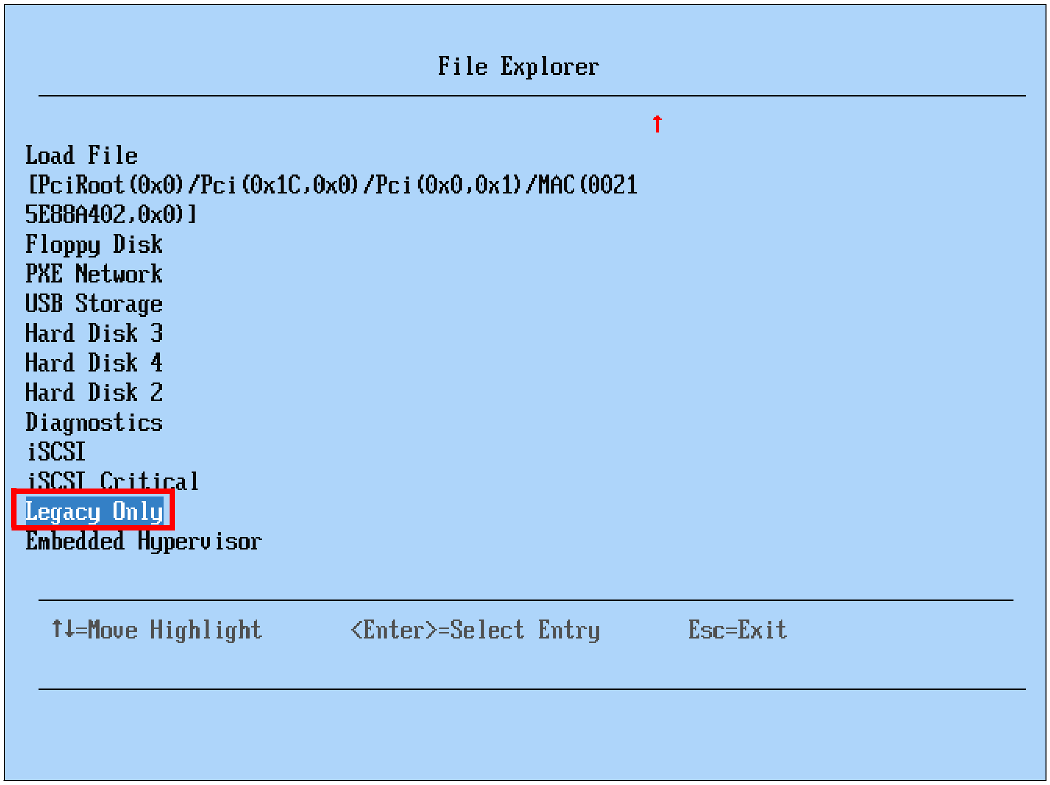


Figure 8-21 Legacy Only option in the boot list

5. When the system reboots, the LSI boot information displays. Press Ctrl-C when prompted, as shown in Figure 8-22.

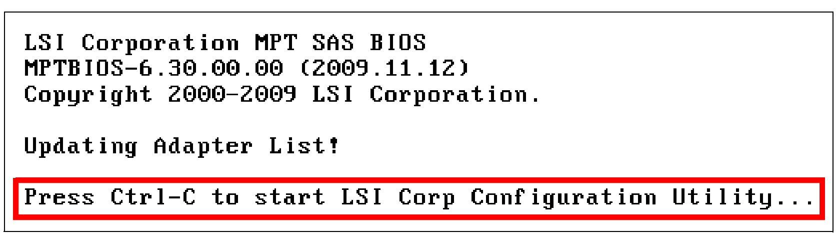


Figure 8-22 Cntrl-C option for LSI on boot

You can now create a RAID array. We describe creating a RAID-1 array in the next section.

8.4.2 Creating a RAID-1 mirror using the LSI Setup Utility

After you have started the LSI Setup Utility, you see the choices that you have to create RAID arrays, as shown in Figure 8-23 on page 390.

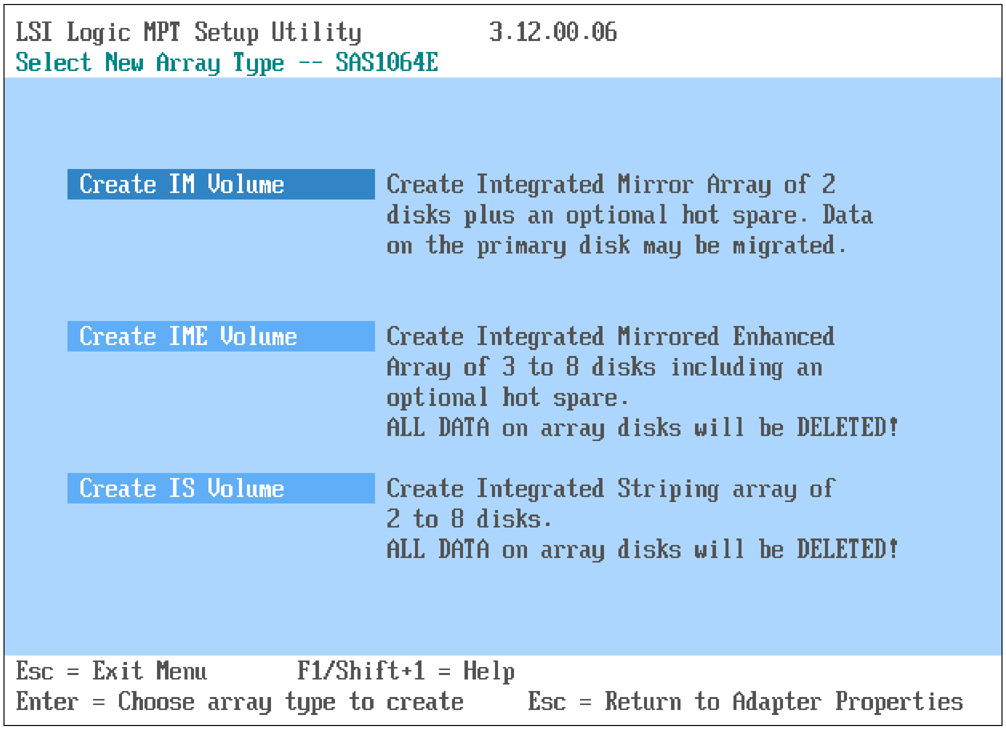


Figure 8-23 LSI array creation options

You can use one of the following options to create arrays:

- Create IM Volume: Creates an integrated mirror (IM) RAID-1 array

  RAID 1 drives are mirrored on a 1 to 1 ratio. If one drive fails, the other drive takes over automatically and keeps the system running. However, in this configuration, you lose 50% of your disk space because one of the drives is a mirrored image.

  The stripe size is 64 Kb and cannot be altered.

  This option also affects the performance of the drives, because all data has to be written twice (one time per drive). See the performance chart in Figure 5-23 on page 207 for details.

- Create IME Volume: Creates an integrated mirrored enhanced (IME) RAID-1E array

  This option requires three drives. This option is not available in the HX5, because the HX5 only has two drives on each node.

- Create IS Volume: Creates an integrated striping (IS) RAID-0 array

  The RAID-0 or IS volume, as shown in LSI, is one of the faster performing disk arrays, because the read and write sectors of data are interleaved between multiple drives. The downside to this configuration is immediate failure if one drive fails. This option has no redundancy.

  In RAID-0, you also keep the full size of both drives. We recommend that you use drives of identical sizes for performance and data storage efficiency.

  The stripe size is 64 Kb and cannot be altered.

In our example, we create a RAID-1 array with two drives. Follow these steps:

1. Select **Create IM Volume** and press Enter. Figure 8-24 on page 391 appears.

```
LSI Logic MPT Setup Utility          v6.04.07.00 (2005.11.03)
Create New Array – SAS1064

       Array Type:              IM
       Array Size (MB)          ------

Slot   Device Identifier                RAID    Hot   Drive       Pred  Size
Num                                     Disk    Spr   Status      Fail  (MB)
  1    IBM-ESXSMAY2036RC    T106        [No]    [No]  --------    No    34715
  0    IBM-ESXSMAY2036RC    T106        [No]    [No]  --------    No    34715

Esc = Exit Menu           F1/Shift+1 = Help
Enter=Select Item Alt+N=Next Array   C=Create an array
```

Figure 8-24 Creating an IM volume

2. In Figure 8-24, click **No** under RAID Disk to form the array. For each disk, you are prompted to confirm by pressing D, as shown in Figure 8-25. Pressing D ensures that you understand that all data on this disk will be deleted.

```
LSI Logic MPT Setup Utility          v6.04.07.00 (2005.11.03)
Create New Array – SAS1064

            M    -    Keep existing data, migrate to an IM array.
                      Synchronization of disk will occur.

            D    -    Overwrite existing data, create a new IM array
                      ALL DATA on ALL disks in the array will be DELETED!!
                      No Synchronization performed

Esc = Exit Menu           F1/Shift+1 = Help
Enter=Select Item Alt+N=Next Array   C=Create an array
```

Figure 8-25 Overwriting existing data

3. Repeat this deletion for the other drive. After you finish, you see Figure 8-26 on page 392. The RAID Disk column now lists each drive as `Yes`.

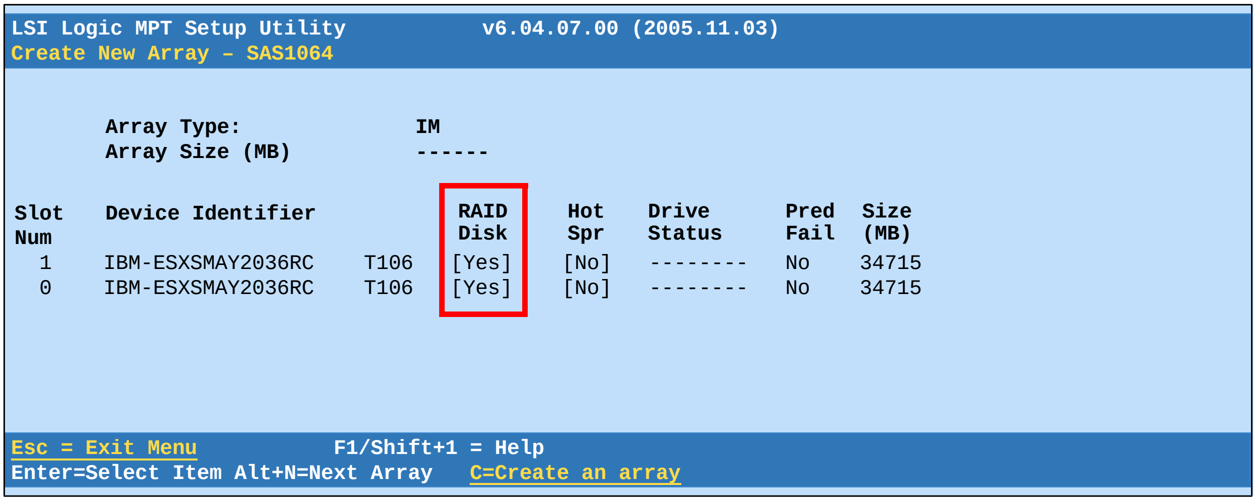


Figure 8-26 Array with both drives created

4. After you have set both disks to `Yes`, press C to create the array. This step completes creating and initializing a RAID-1 array.

8.4.3 Using IBM ServerGuide to configure the LSI controller

IBM ServerGuide also lets you configure the RAID arrays for the onboard LSI controller as part of the IBM ServerGuide setup procedure, as shown in Figure 8-27 on page 393.

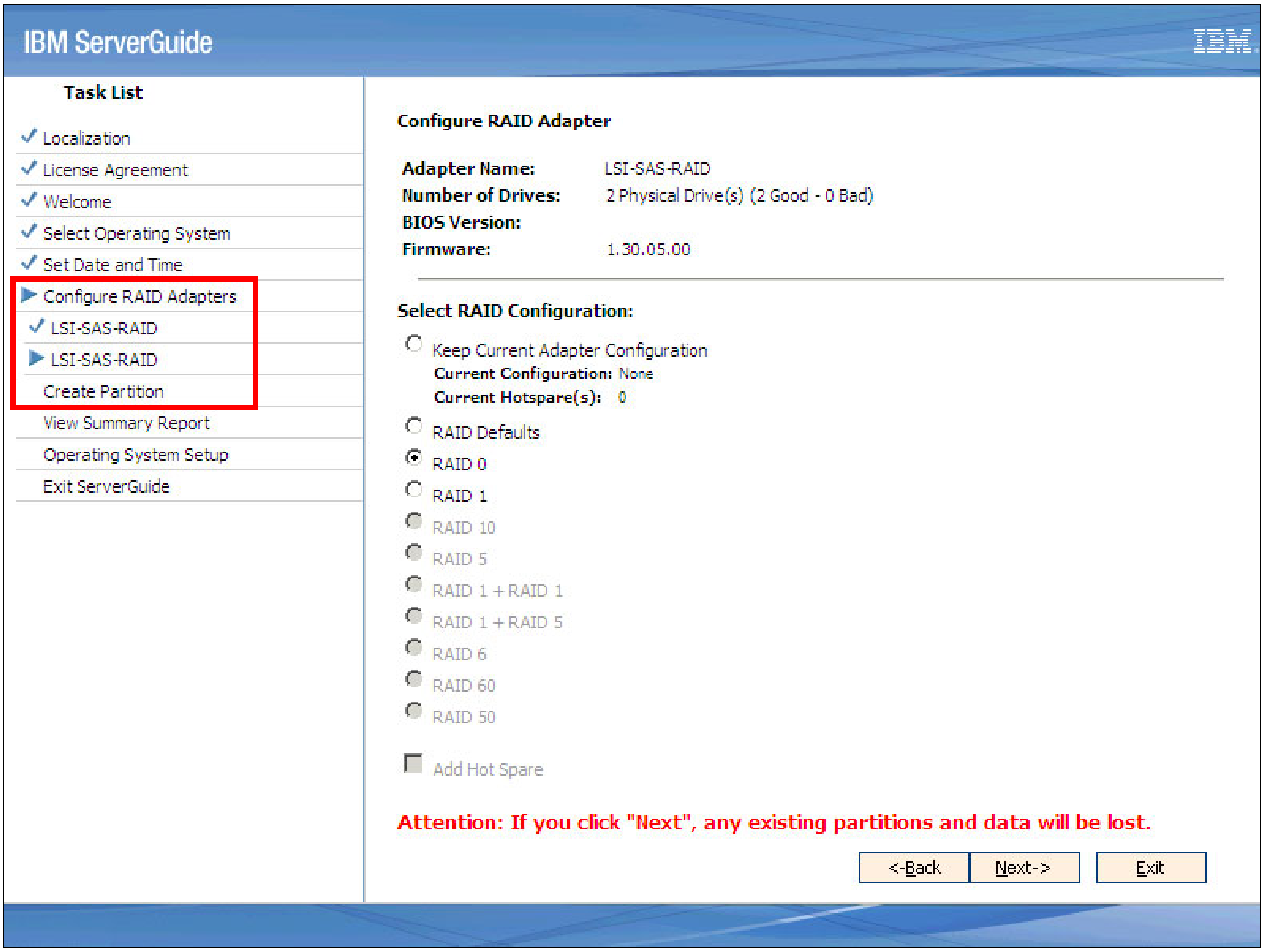


Figure 8-27 LSI setup using ServerGuide

If desired, you can install an operating system on each node. Then, you can use this configuration to boot each blade individually if you need increased flexibility. You can use the temporary boot in stand-alone mode in the complex scaling section, as shown in Figure 8-28 on page 394, which allows you to temporarily boot each node separately without having to delete the 2-node configuration.

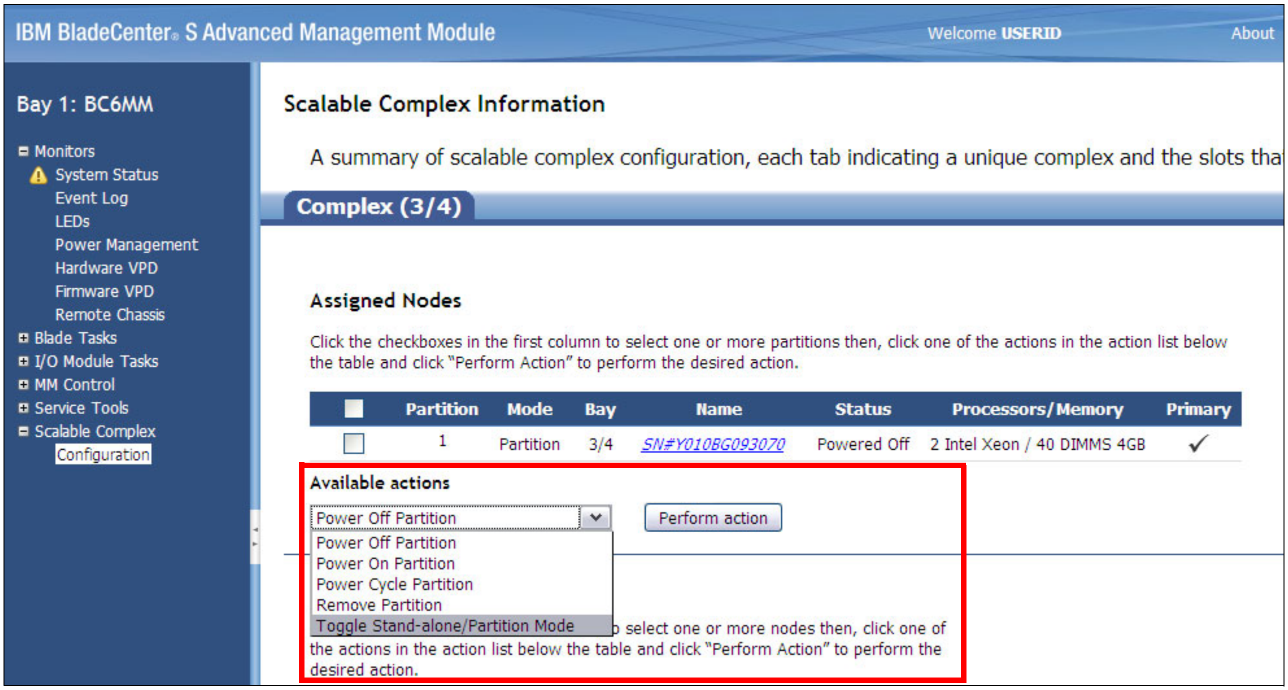


Figure 8-28 Complex scale toggle for stand-alone mode

8.4.4 Speed Burst Card reinstallation

In the event that you decide to split a 2-node complex permanently in order to use the blades independently, always reinstall the Speed Burst Card for performance reasons.

The Speed Burst Card takes the QPI links that are used for scaling two HX5 2-socket blades and routes them back to the processors on the same blade. Without the Speed Burst Card, there is only one QPI link between processor 1 and 2 for exchanging memory information. The Speed Burst Card increases the number of QPI links from 1 to 3, which will significantly increase the bandwidth of data that can be exchanged between the memory of the two processors. For solutions that use intensive memory, the performance boost will be measurable. See the block diagram in Figure 8-29 on page 395.

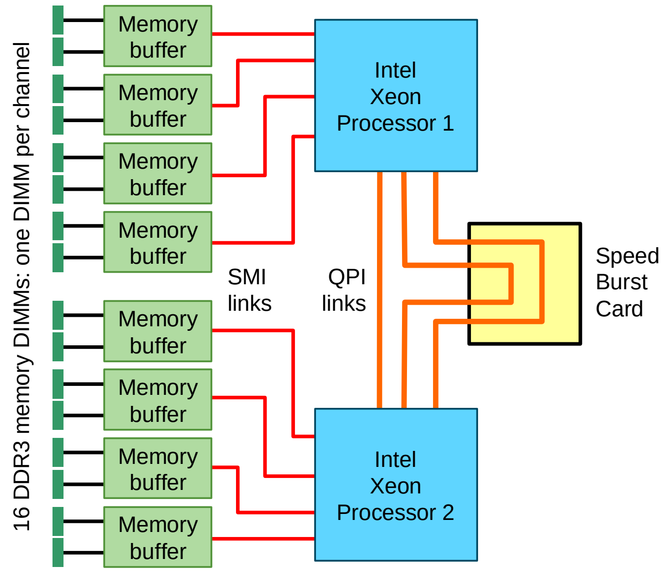


Figure 8-29 HX5 1-Node Speed Burst Card block diagram

Figure 8-30 shows where the Speed Burst Card is installed on the HX5.

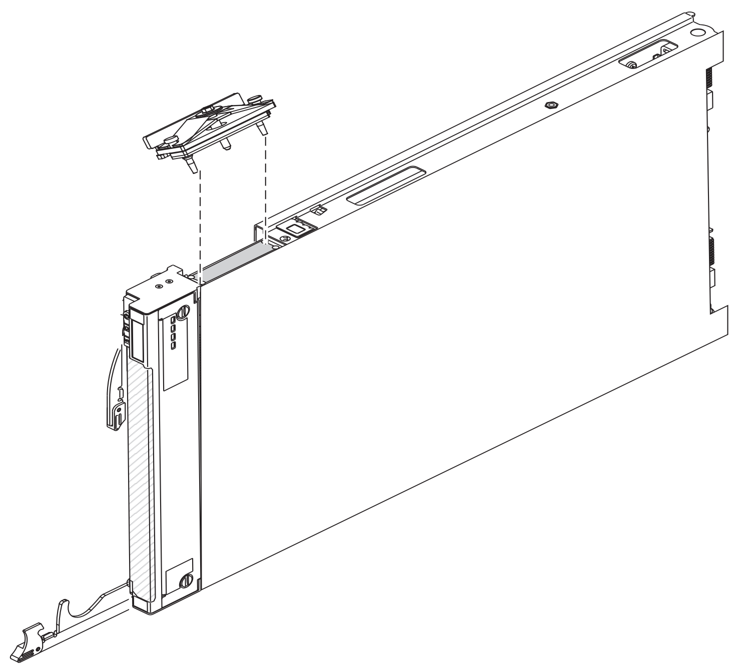


Figure 8-30 Installing the Speed Burst Card

8.5 UEFI settings

Unified Extensible Firmware Interface (UEFI) is an interface that takes care of handing over the pre-boot environment to the operating system. UEFI is effectively the replacement for BIOS. BIOS has been around for many years but it was not designed to handle the amount of hardware that can be added to a server today.

New IBM System x models, including the HX5, implement UEFI to take advantage of its advanced features. You access UEFI by pressing F1 during the system initialization process, as shown on Figure 8-31.

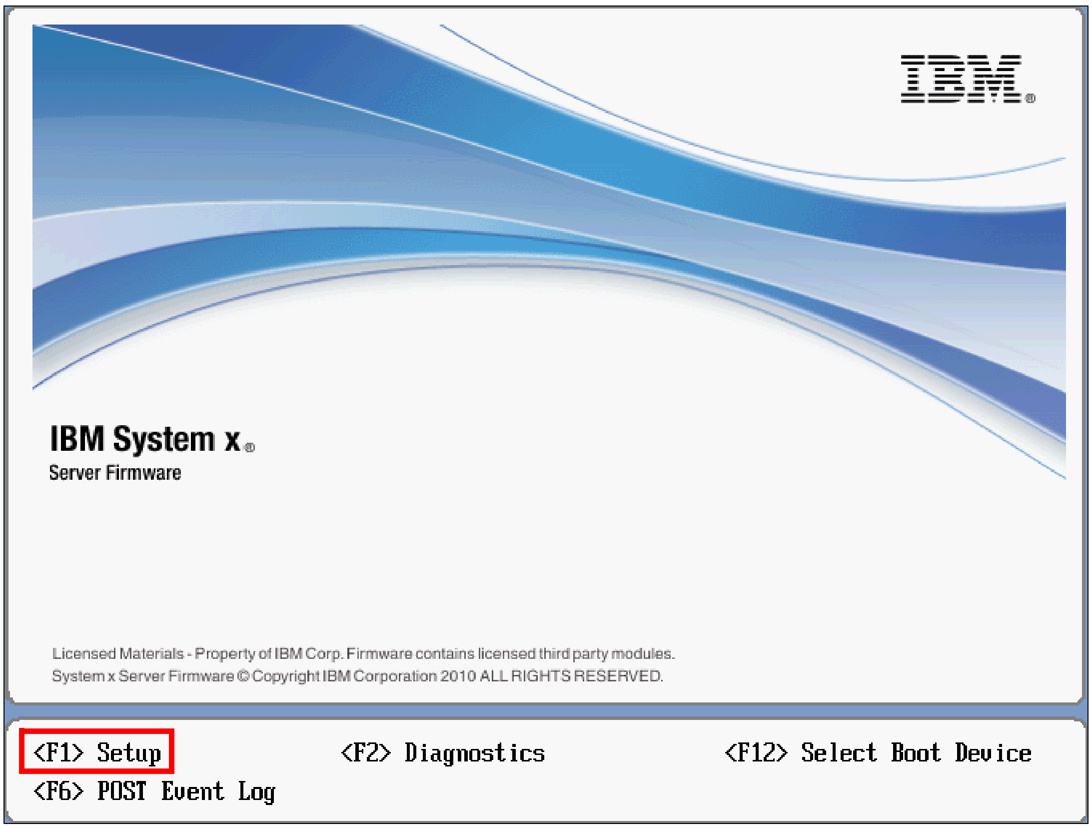


Figure 8-31 UEFI window on system start-up

Figure 8-32 on page 397 shows the main UEFI System Configuration and Boot Management panel.

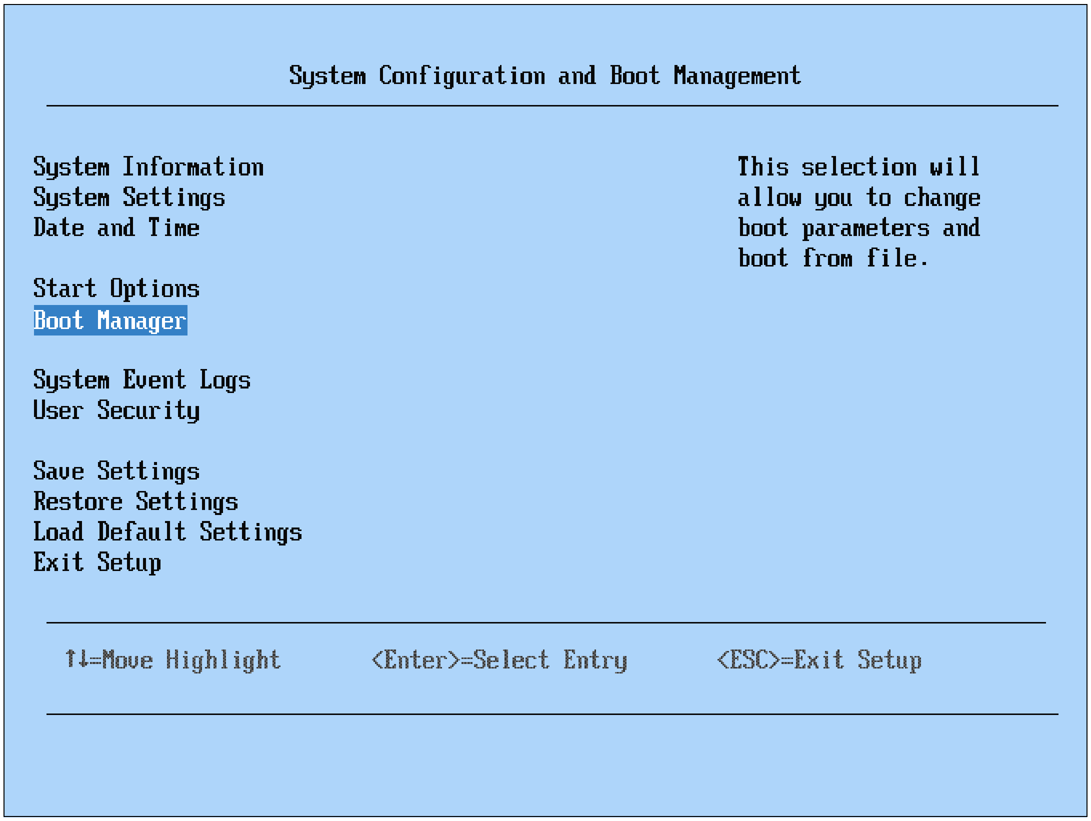


Figure 8-32 System Configuration and Boot Management UEFI menu

You can obtain more general information about UEFI at the following website:

http://www.uefi.org/home/

For an explanation of UEFI settings, see 2.7, "UEFI system settings" on page 36.

Generally, the HX5 UEFI requires little configuration. The factory default UEFI settings allow the HX5 to work correctly in a 2-socket, 4-socket, and MAX5-attached configuration. There are, however, a number of UEFI settings that can be adjusted to meet specific needs. You can adjust the UEFI settings, for example, for better performance.

This section contains the following topics:

- 8.5.1, "UEFI performance tuning" on page 397
- 8.5.2, "Start-up parameters" on page 398
- 8.5.3, "HX5 single-node UEFI settings" on page 400
- 8.5.4, "HX5 2-node UEFI settings" on page 401
- 8.5.5, "HX5 with MAX5 attached" on page 401
- 8.5.6, "Operating system-specific settings in UEFI" on page 402

8.5.1 UEFI performance tuning

The question of how to tune a server or blade for performance is never an easy one to answer. Many factors influence how you configure the server, such as the application installed. For example, a database server will generate a separate load on the hardware than a file and print server.

A server is only as good as its weakest link, so you can have a well-tuned server, but a poorly configured SAN to which it connects. Generally, you need to look at an entire solution, including the SAN and network infrastructure, to fully achieve the best performance from the server.

Table 8-4 provides an overview of UEFI settings for the HX5 that address a number of performance scenarios.

Table 8-4 Overview of settings

Setting	MaxImum performance	Virtualization	Low latency	Performance per watt	HPC
TurboMode	Enabled	Enabled	Disabled	Disabled	Disabled
TurboBoost	Traditional	Power Optimized	Automatically disabled	Automatically disabled	Automatically disabled
Processor Performance states	Enabled	Enabled	Disabled	Enabled	Disabled
C states	Disabled	Enabled	Disabled	Enabled	Enabled
C1E state	Disabled	Enabled	Disabled	Enabled	Enabled
Prefetcher	Enabled	Enabled	Enabled	Enabled	Enabled
Hyper-Threading	Enabled	Enabled	Disabled	Enabled	Enabled
Execute Disable	Disabled	Enabled	Disabled	Enabled	Disabled
Virtualization Extensions	Disabled	Enabled	Disabled	Enabled	Disabled
QPI Link Speed	Max Performance	Max Performance	Max Performance	Power Efficiency	Max Performance
IMM Thermal Mode	Performance	Performance	Performance	Normal	Performance
CKE Policy	Disabled	Disabled	Disabled	Disabled	Disabled
DDR Speed	Max Performance	Max Performance	Max Performance	Max Performance	Max Performance
Page Policy	Closed	Closed	Closed	Closed	Closed
Mapper Policy	Closed	Closed	Closed	Closed	Closed
Patrol Scrub	Disabled	Disabled	Disabled	Disabled	Disabled
Demand Scrub	Enabled	Enabled	Disabled	Enabled	Disabled

8.5.2 Start-up parameters

UEFI systems, especially when scaled, can take time to start up. The number of adapters installed in each node, for example, has a direct influence on the time that it takes for UEFI to initialize.

A simple step to reduce the overall system start-up time is to remove any unnecessary boot devices from the boot order list. In particular, ensure that PXE network boot is removed from the boot order list if it is not used. PXE network boot can increase start-up time substantially, especially if it is also placed in the incorrect order in the boot order list.

Use the following steps to remove entries from the Boot Order menu and configure the correct boot order:

1. Log in to the Advanced Management Module and select **Blade Tasks** → **Remote Control**. Click **Start Remote Control** and select the blade that you want to control from the blade selection pull-down box.
1. Power on the system using the power control feature on the remote control, and press F1 when the UEFI splash panel displays, as shown in Figure 8-31 on page 396.
2. Navigate to the **Boot Manager** menu by selecting it from the System Configuration and Boot Management main menu. The Boot Manager displays, as shown in Figure 8-33.

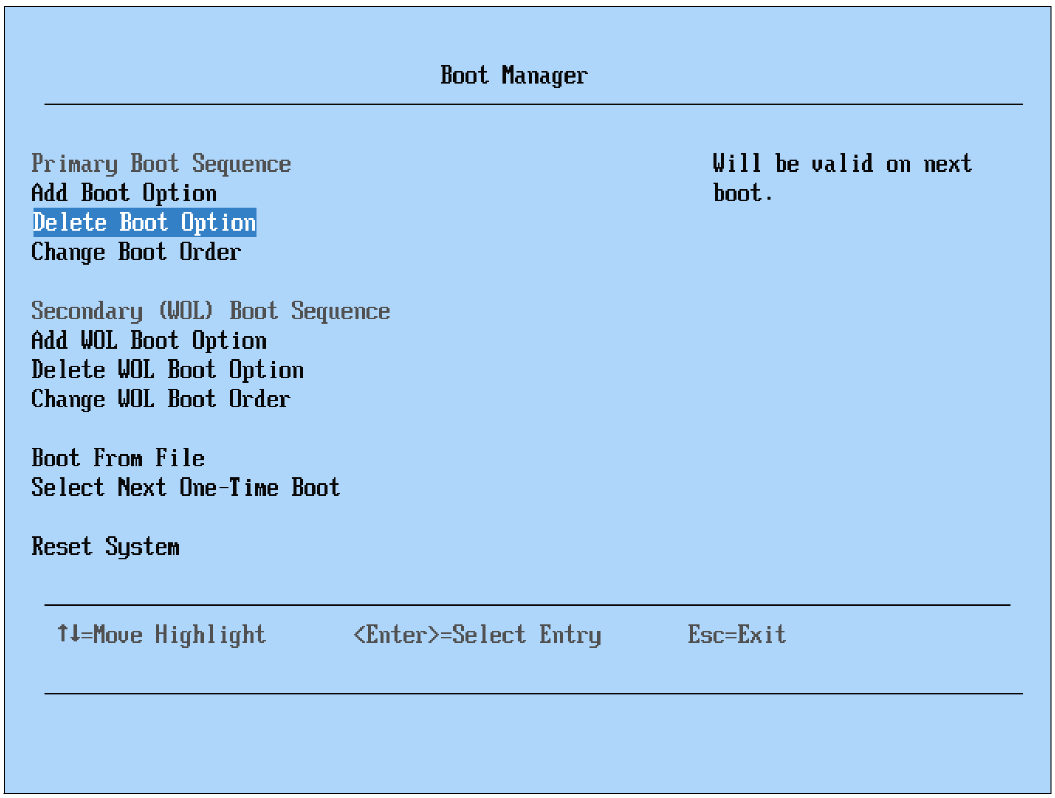


Figure 8-33 Selecting Boot Manager menu

3. Select **Delete Boot Option** from the Boot Manager menu.
4. Select all of the items from which you do not want to boot by using the Spacebar to select them. When you have selected all of the items, scroll down to the end of the page using the down arrow key and select **Commit Changes**, as shown in Figure 8-34 on page 400.

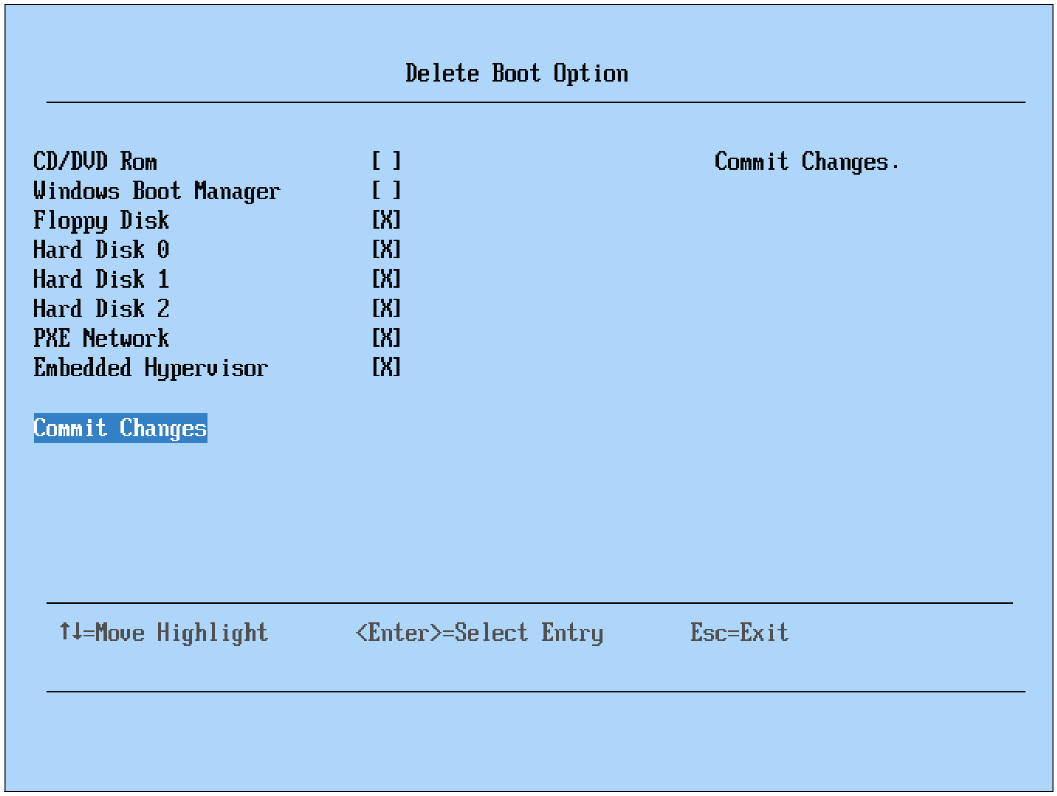


Figure 8-34 Deleting boot options

Note that, in our example in Figure 8-34, we also have selected **Hard Disk 0** for removal. Our system is running Windows 2008 with GUID Partition Table (GPT) disk, and therefore, Windows Boot Manager is being used to boot the operating system.

5. Press Esc to return to the Boot Manager menu.
6. Select **Change Boot Order** from the Boot Manager menu.
7. Press Enter to make the device list active.
8. Use the up and down arrow keys to navigate to the device for which you want to change the order. After you highlight the device, use the - or Shift and + keys to move the device up or down the list. You can then perform the same actions to move other devices up or down the list. Press Enter when finished.
9. Use the down arrow key to highlight **Commit Changes** and press Enter to commit the changes that you have made.
10. Press Esc to return to the Boot Manager main menu.
11. Press Esc again to exit to the System Configuration and Boot Management main menu.
12. Press Esc again to exit the UEFI and press the Y key to exit and save any changes that you have made. The HX5 then proceeds to boot normally.

8.5.3 HX5 single-node UEFI settings

No specific UEFI settings are required for the HX5 to operate in a single-node configuration. The settings used are determined by the operating system that is installed.

8.5.4 HX5 2-node UEFI settings

No specific UEFI settings are required for the HX5 to operate as a 2-node complex. The configuration of the complex is provided by the Advanced Management Module (AMM). See 8.6, “Creating an HX5 scalable complex” on page 402 for instructions to perform this task.

8.5.5 HX5 with MAX5 attached

When a MAX5 is attached to an HX5, the UEFI adds an additional setting to the Memory configuration page within the UEFI. Specifically, it adds the MAX5 Memory Scaling option, which you can see by entering the UEFI at start-up and navigating to **System Settings** → **Memory** from the System Configuration and Boot Management main menu. All other settings behave similarly in the *single-node*, *dual-node*, and *memory-expanded* configurations. Figure 8-35 shows this additional option.

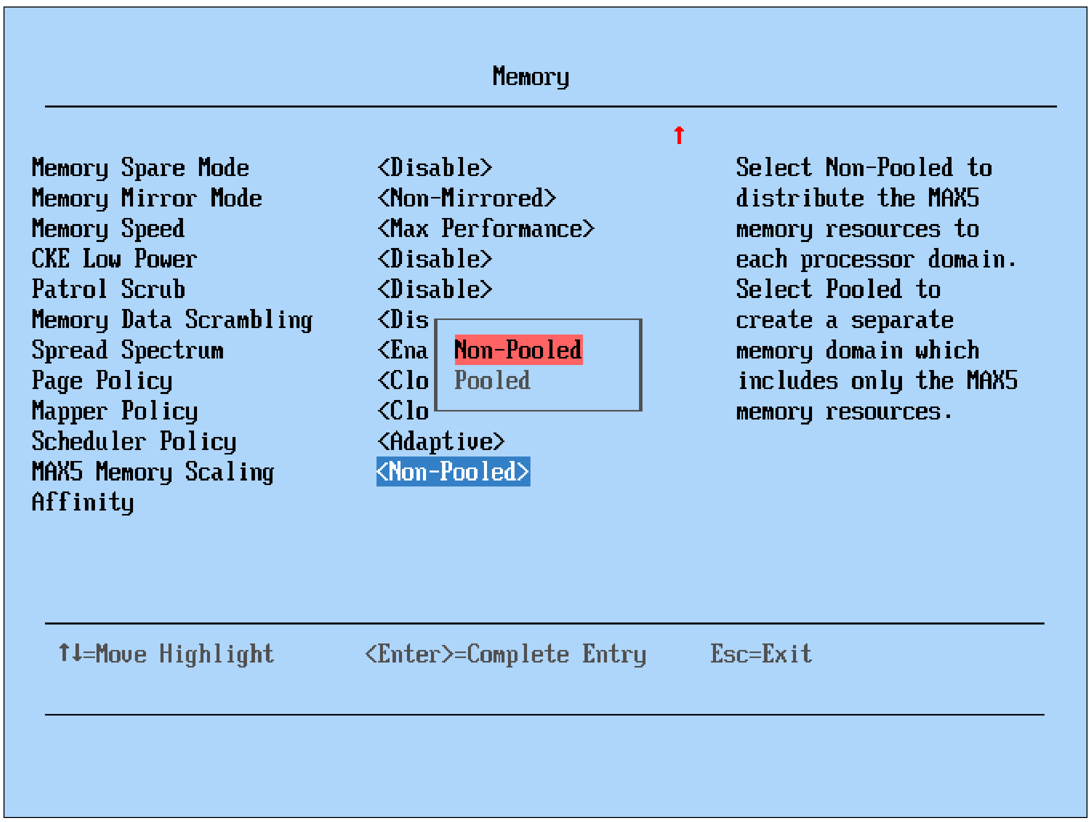


Figure 8-35 MAX5 Memory Scaling option

The MAX5 Memory Scaling setting provides two options to determine how the system will present the memory capacity in the MAX5 unit to the running operating system:

- Non-Pooled

  The default option divides and assigns the memory in the MAX5 between the two installed processors.

- Pooled

  This option presents the additional memory in the MAX5 as a pool of memory without being assigned to any particular processor.

Use these general settings for each operating system type:

- Pooled: Use this setting with Linux operating systems.
- Non-Pooled: VMware and Microsoft operating systems must be configured to use the Non-Pooled setting.

8.5.6 Operating system-specific settings in UEFI

When installing an operating system onto an HX5, there is little that you are required to configure in the UEFI. When using MAX5, ensure that you follow the MAX5 memory scaling settings that are described in 8.7, “Operating system installation” on page 407.

8.6 Creating an HX5 scalable complex

The HX5 provides the flexibility to scale from a single-node 2-socket system to a dual-node 4-socket system, and back again when the business demand requires it. You can achieve all of this scaling without having to make any physical changes to the hardware after the scalability kit has been installed. The configuration of a multinode HX5 is controlled through the AMM web interface. The AMM communicates with the onboard integrated management module (IMM) to enable and disable the scaling capabilities of the system, allowing for great flexibility.

In this section, we demonstrate how to scale two single-node HX5s into a dual-node complex.

> **Firmware requirements:** Before you attempt to scale the HX5, see Table 8-3 on page 380 for the minimum firmware requirements for the HX5 and the AMM.

Use the following steps to create a complex:

1. Log in to the AMM.
2. Navigate to **Scalable Complex** → **Configuration**. The scalable systems and the slots that they occupy are shown as tabs, as shown in Figure 8-36 on page 403. For our example, we use the HX5s installed in blade bays 1 and 2 to create a partition.

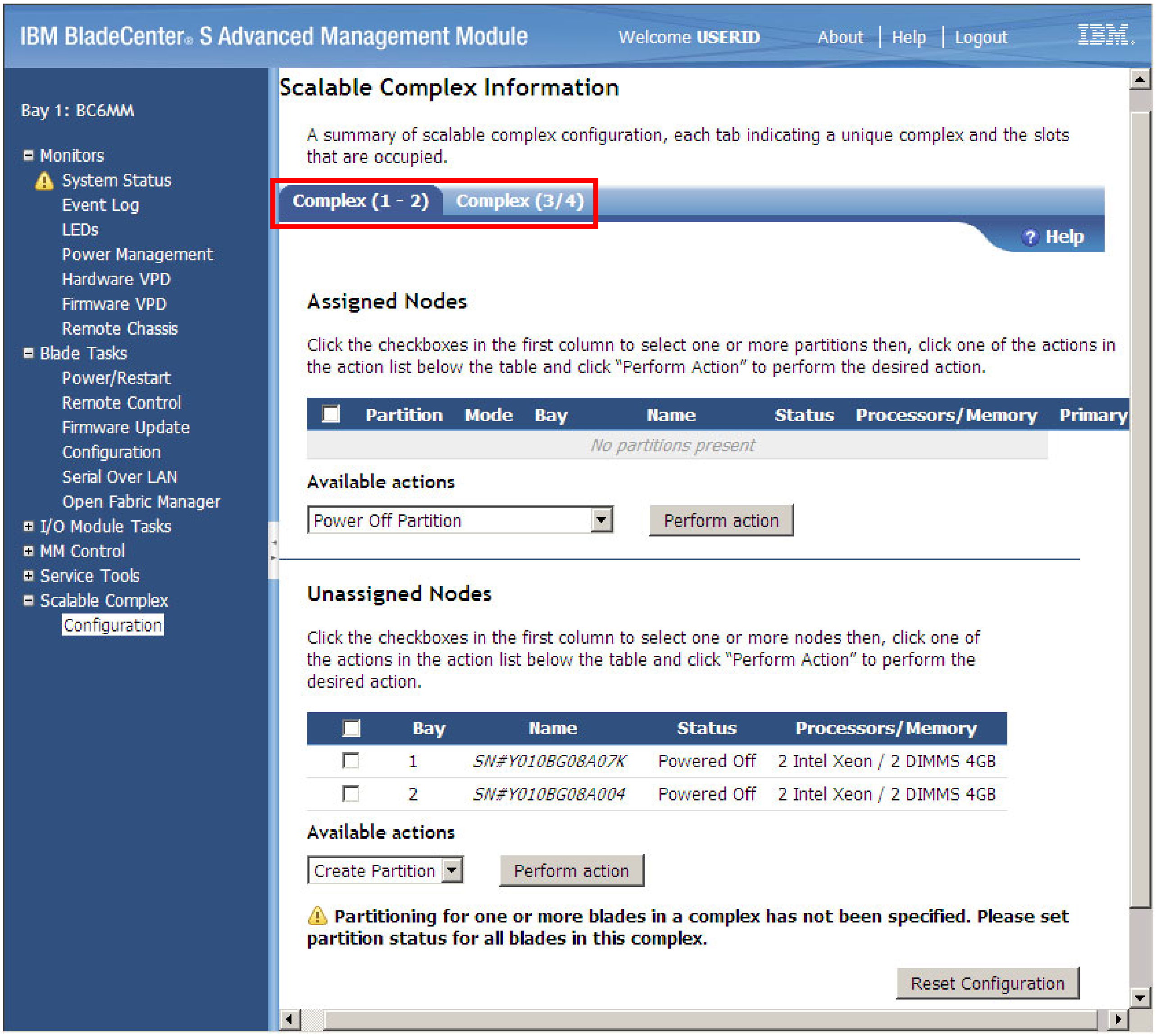


Figure 8-36 Scalable complex configuration panel

3. Ensure that both HX5s are powered off by looking at the Status column in the Unassigned Nodes section. If the systems are *not* in a `Powered Off` state, the following warning message (Figure 8-37) displays when you attempt to create a partition.

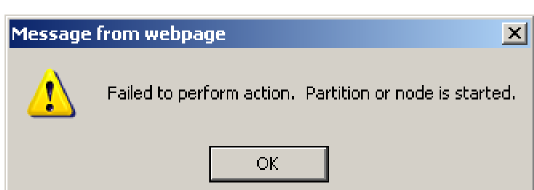


Figure 8-37 Partition creation failure message due to blades being in a powered on state

4. If the systems are powered on, you can shut them down from this page by clicking the check boxes next to them and by selecting the **Power Off Node** action from the Available actions pull-down, as shown in Figure 8-38 on page 404. The Status will change to `Powered Off` momentarily after the `Power Off Node` action has been applied.

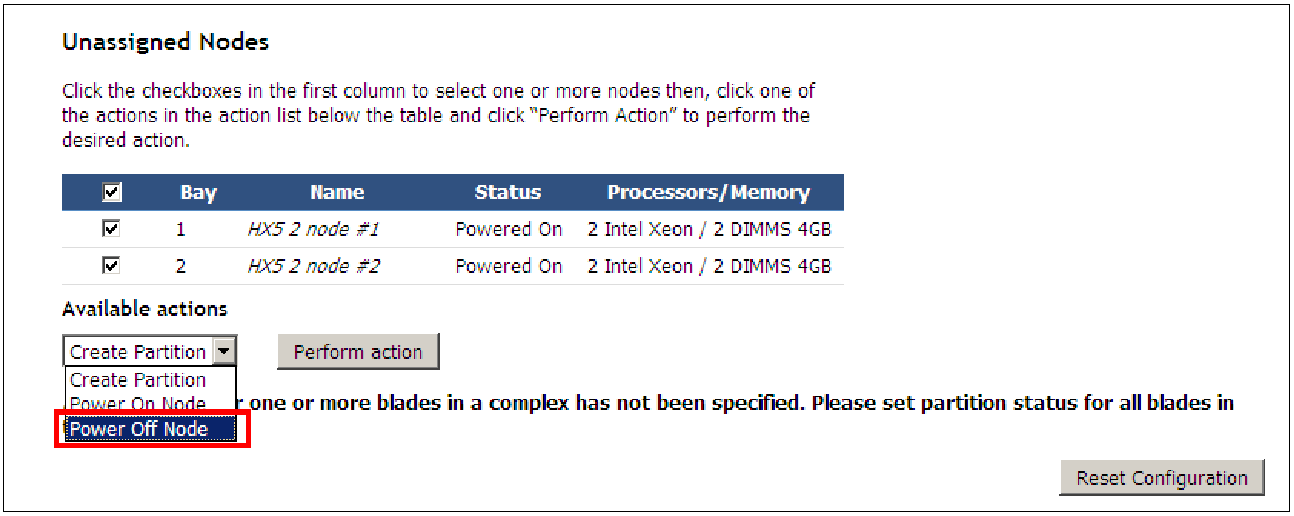


Figure 8-38 Powering off an HX5 from the scalable complex configuration page

5. Select the two HX5 systems in the Unassigned Nodes area by clicking the check boxes next to them. The **Create Partition** action in the Available actions pull-down list box is preselected for you, as shown in Figure 8-39.

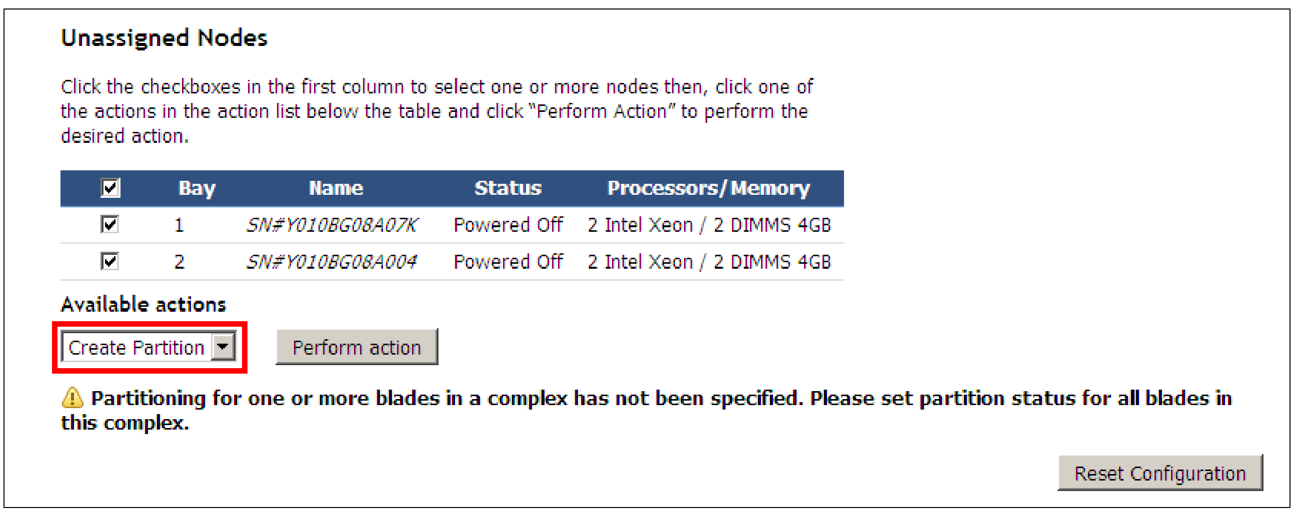


Figure 8-39 Creating a partition

6. Click **Perform action**. The complex will form and the task of creating a partition is now complete.

Figure 8-40 on page 405 shows the actions that you will be able to perform against the complex.

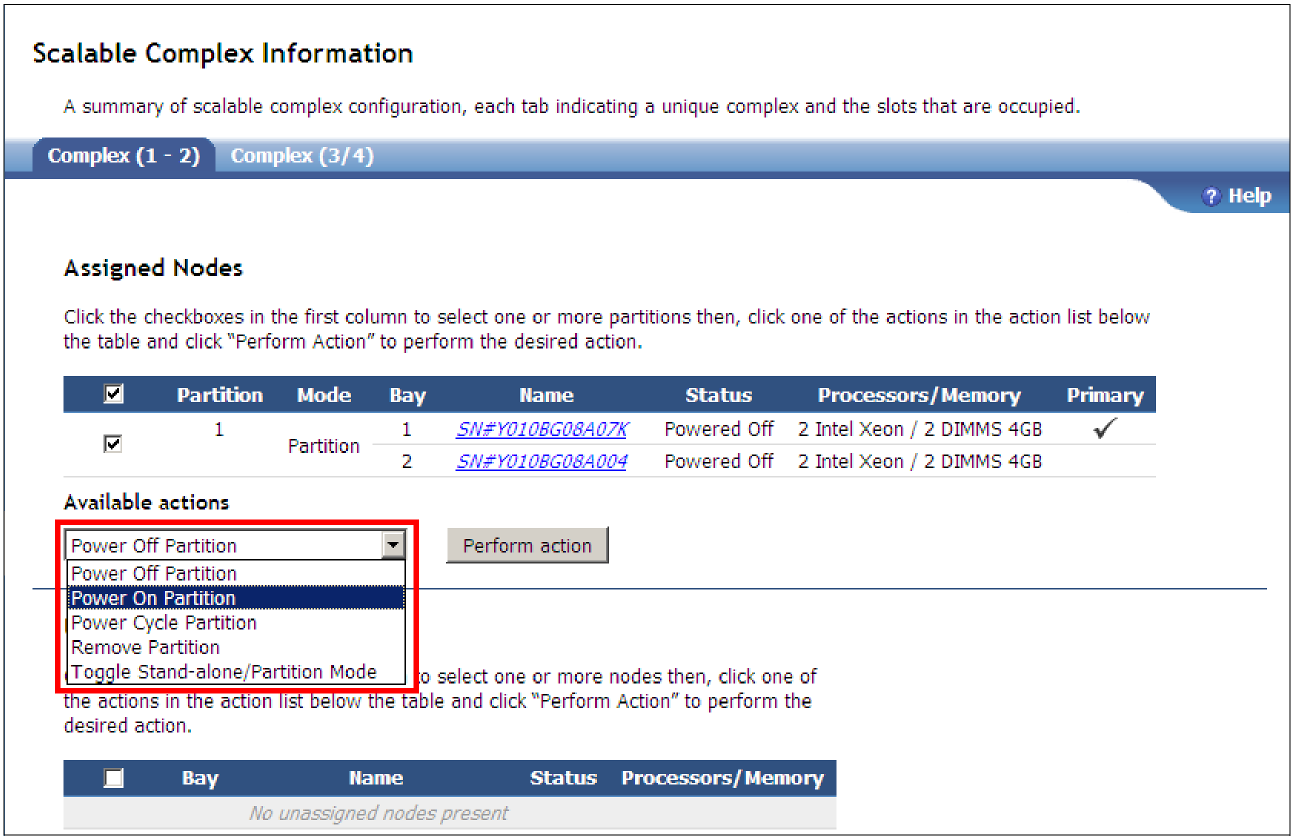


Figure 8-40 Available actions to perform against the complex

The following options are available:

- Power Off Partition

  This option powers off all blades in the partition (in this example, it powers off blade 1 and blade 2).

- Power On Partition

  This option powers on all blades in the partition.

- Remove Partition

  This option removes the partitioning of the selected partition. The nodes that formed the partition are then removed from the Assigned Nodes section and become available again in the Unassigned Nodes section.

- Toggle Stand-alone/Partition Mode

  This mode allows you to toggle between single partition mode and stand-alone mode without having to modify the physical setup of the blade servers, for example:

  – You can toggle the scalable blade complex to stand-alone mode and install a separate operating system on each blade server and run separate applications on each blade server.

  – You can then toggle the blade server complex back to a single partition and run applications that take advantage of the four processors and 32 DIMMs. The operating system that is in use in a single partition is the operating system that is installed on the primary blade server.

  – Later, you can toggle the complex back to stand-alone mode again to gain access to the operating system on the secondary blade server.

8.6.1 Troubleshooting HX5 problems

For general troubleshooting advice for the HX5, see the *HX5 Problem Determination and Service Guide*, which is available at this website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084529

Scalability card issues

The scalability card enables the HX5 to form a complex. Consider the following points when using the scalability card:

- If the scalability card is not attached correctly, the HX5 blades appear as individual complex systems. In our example, as shown in Figure 8-41, you can see that the HX5 blades in slots 1 and 2 appear as two separate complex systems.
- If the scalability card is removed from an HX5 complex with an existing partition, you must ensure that the scalability card is reattached correctly. You need to remove the scalability card, for example, if replacing a memory DIMM in the primary node. Failing to reattach the scalability card correctly removes the partition information and both blades that formed the complex appear as individual complex systems again, as shown in Figure 8-41.
- When a scalability card is attached, allow approximately 5 - 10 minutes for the IMM to initialize and the complex information to update in the Scalable Complex Information configuration page. Failing to allow this update will prevent you from creating a partition.

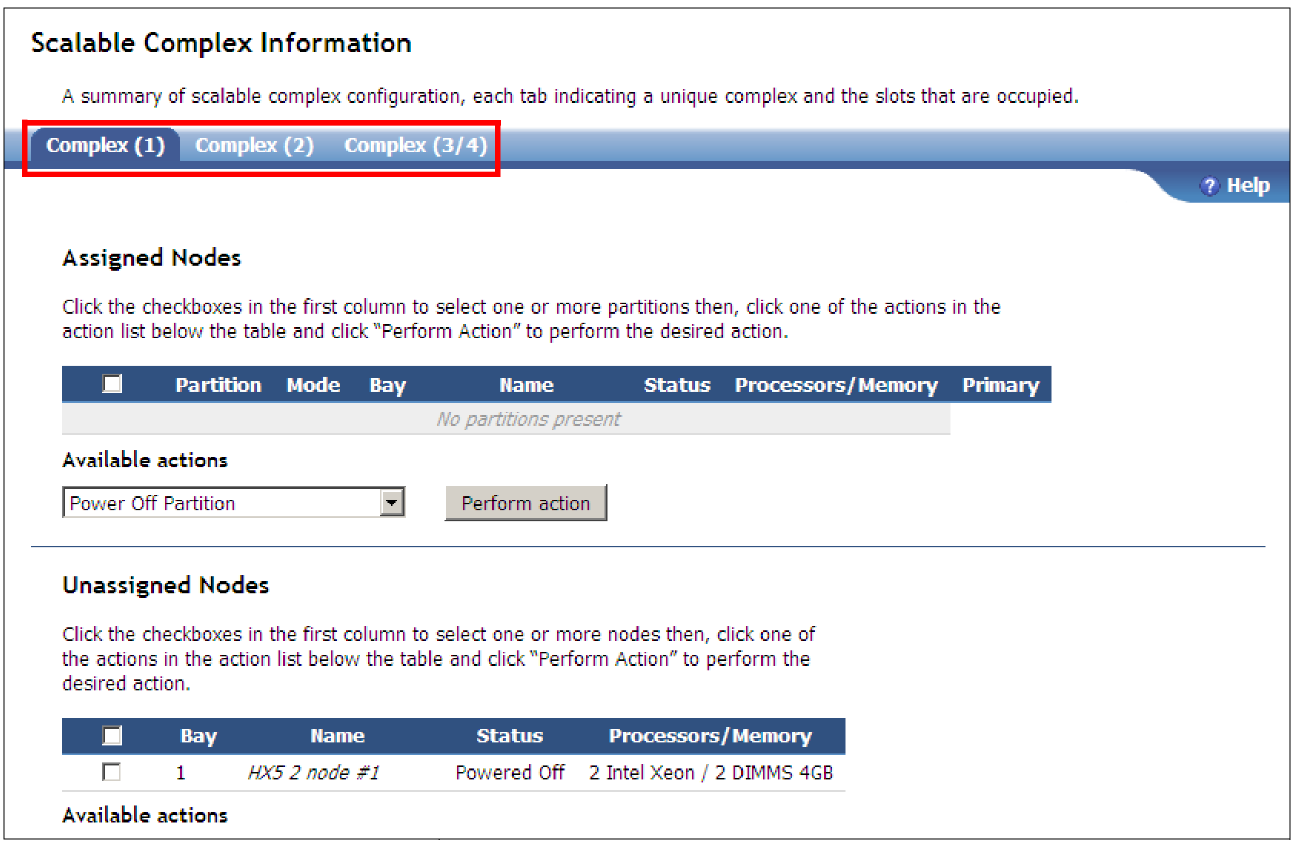


Figure 8-41 An example of a scalability card that is not installed correctly

8.7 Operating system installation

Installing an operating system on an HX5 is practically identical to installing an operating system on any other eX5 system. The only difference, when installing from a remote location, is that the AMM is used for remote control instead of the IMM web user interface.

In this section, we describe the various installation media that is available to deploy an operating system to the HX5. We also provide specific configuration information that must be used to successfully deploy the latest versions of ESX onto HX5. We also cover installation hints and tips.

This section includes the following topics:

8.7.1 Operating system installation media

The media that is typically used to install an operating system is a CD or DVD. However, in cases where it is not possible to use physical installation media, the HX5 offers a number of alternatives.

Preboot eXecution Environment (PXE) network boot

The onboard Broadcom 5709 Gb Ethernet supports PXE network boot so that it is possible to boot into the PXE and access installation files from a remote location.

Image load via remote control

The AMM Remote Control feature provides the HX5 with the ability to mount installation media to it remotely. You access the Remote Control feature by logging into the AMM and navigating to **Blade Tasks** → **Remote Control** and clicking **Start Remote Control**, as shown in Figure 8-42 on page 408.

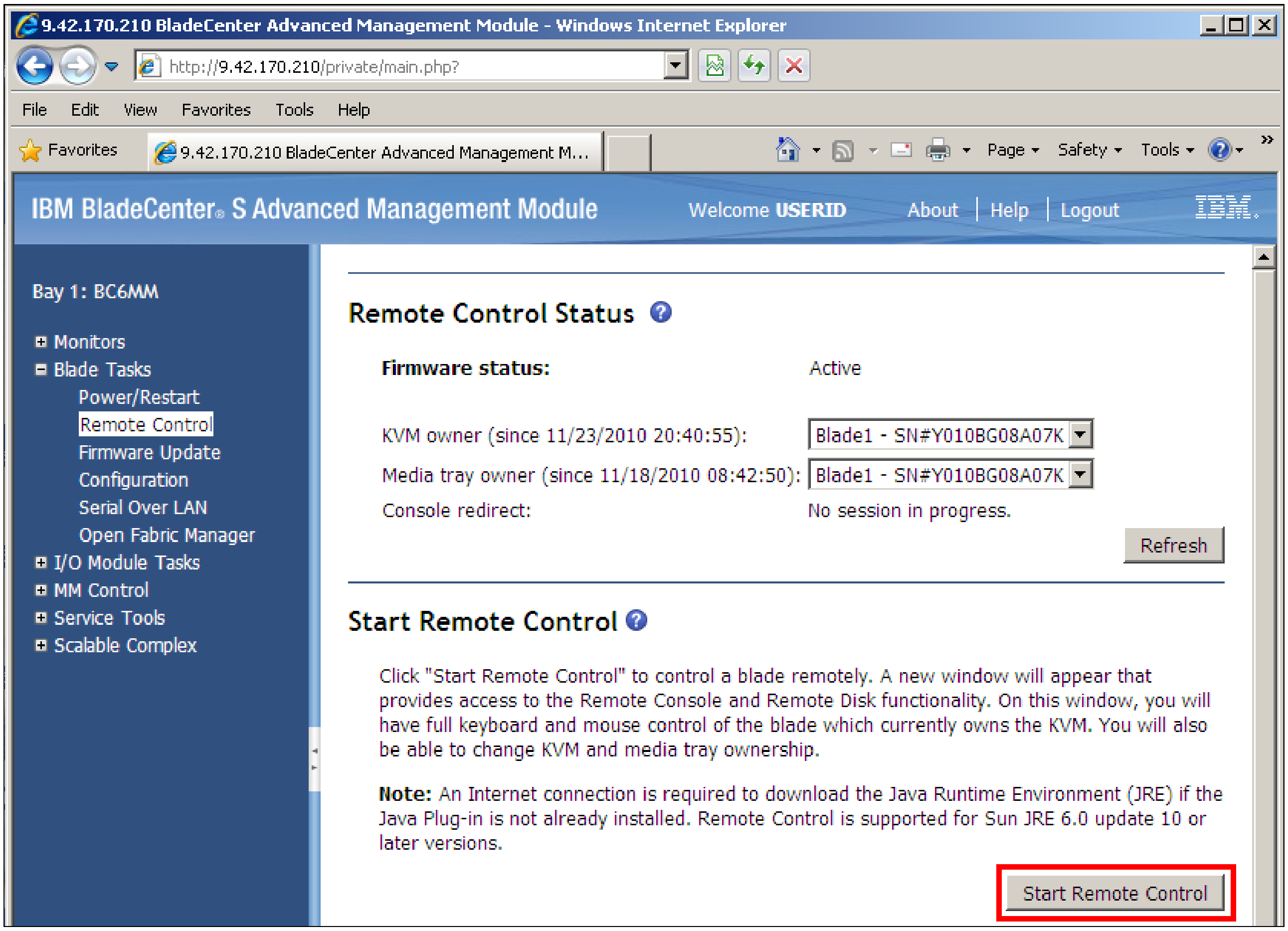


Figure 8-42 Remote Control function in the AMM

By clicking the **Remote Drive** icon  (as highlighted in Figure 8-43 on page 409), you can mount any local CD\DVD Rom drive from the management workstation from which you are running the remote control session. You can also mount a supported `.ISO` or `.IMG` type file, such as an operating system image file.

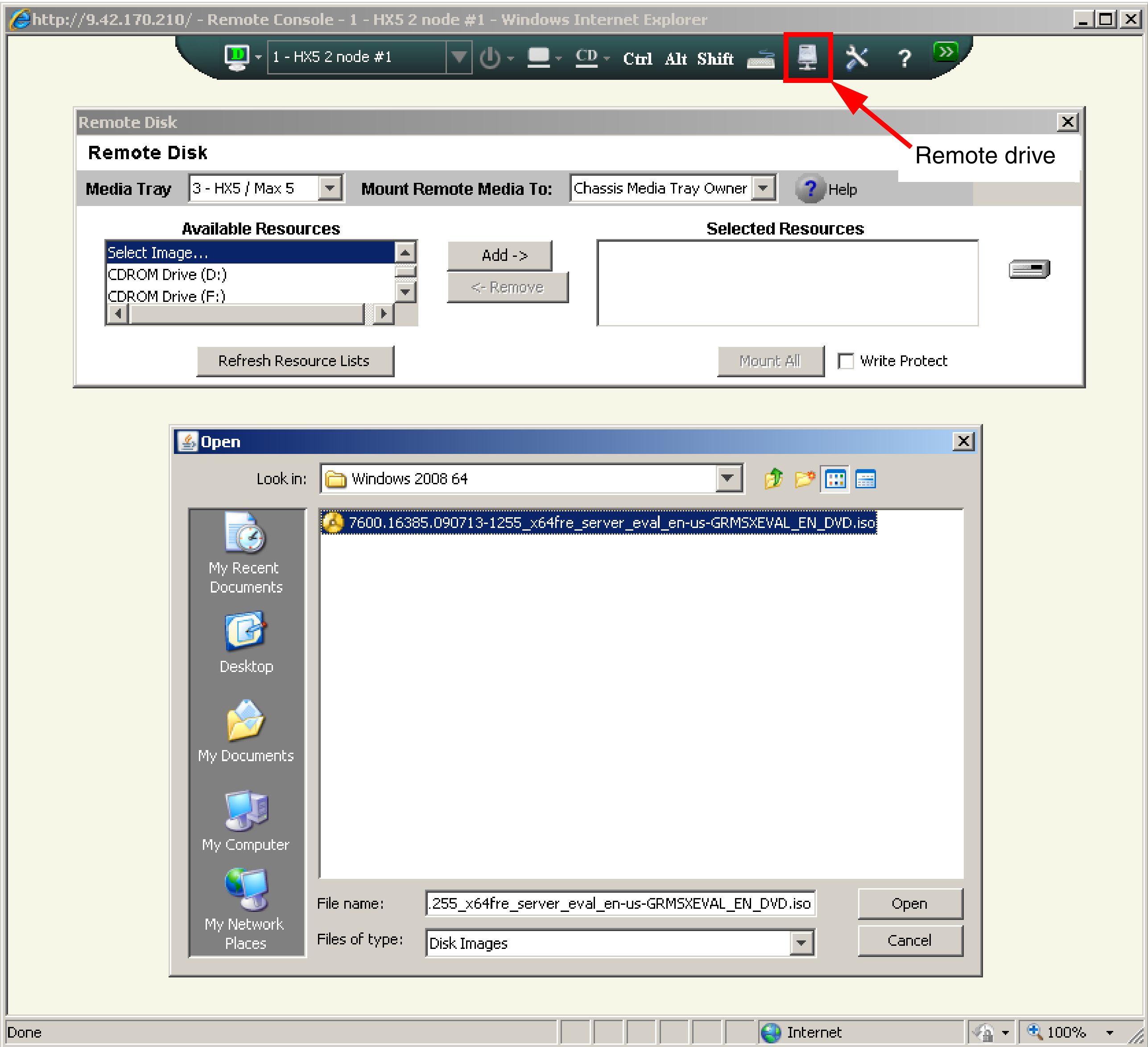


Figure 8-43 Mounting physical media or disk images

Use the following steps to mount an image as virtual CD/DVD media at the controlled system:

1. Click **Select Image** under the Available Resources area of the Remote Disk panel, as shown in Figure 8-43.
1. Click **Add Image**.
2. Browse to an image that is in a `.iso` or `.img` format and click **Open** when finished.
3. Click **Mount All** when finished to mount the image to the blade.

Local USB port

You can use the local USB port to attach a USB flash drive that contains the operating system installation files. There are several methods to create a bootable flash drive.

For VMware, you can use the embedded hypervisor key, which is preinstalled with ESXi, and you do not need to install VMware. For more information about the embedded hypervisor key, see 2.9.1, “VMware ESXi” on page 50.

For Linux, look on the vendor websites. They contain information about installation with a USB flash drive. For example, the following websites provide details for using a USB key as an installation medium:

- Installing Red Hat Linux from a USB flash drive:

  http://ibm.com/support/techdocs/atsmastr.nsf/WebIndex/WP101131

- How to create a bootable USB drive to install SLES:

  http://www.novell.com/support/php/search.do?cmd=displayKC&docType=kc&externalId=3499891

You can also use the IBM ServerGuide Scripting Toolkit to create a bootable USB flash drive, as explained in the next section.

ServerGuide Scripting Toolkit

As described in 9.9, "IBM ServerGuide Scripting Toolkit" on page 507, you can use the ServerGuide Scripting Toolkit to customize your operating system deployment. You can use the ServerGuide Scripting Toolkit for Windows, Linux, and VMware. This section contains information about deployment to allow you to begin using the Toolkit as quickly as possible.

For more information, see the *IBM ServerGuide Scripting Toolkit, Windows Edition User's Reference* and *IBM ServerGuide Scripting Toolkit, Linux Edition User's Reference* at the following website:

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

Windows installation

This section describes the process to install the ServerGuide Scripting Toolkit, to create a deployment image for Windows 2008 R2 Enterprise Edition, and to copy this image to a USB key for deployment. To configure a USB key for deployment, you need the following devices:

- A system running Windows Vista, Windows Server 2008, Windows 7, Windows Server 2008 R2, Windows 2.1 PE, or a Windows 3.0 PE session
- A USB key with a storage capacity at least 64 MB larger than your Windows PE image, but not less than 4 GB

Use this procedure:

1. Install the ServerGuide Scripting Toolkit

   You must install the English language version of the Windows Automated Installation Kit (AIK) for the Windows 7 family, Windows Server 2008 family, and Windows Server 2008 R2 family, which is available at the following website:

   http://www.microsoft.com/downloads/en/details.aspx?familyid=696DD665-9F76-4177-A811-39C26D3B3B34&displaylang=en

   Follow these steps to install the ServerGuide Scripting Toolkit, Windows Edition:

   a. Download the latest version from the following website:

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

   b. Create a directory, for example, `C:\sgshare`.

   c. Decompress the `ibm_utl_sgtkwin_X.XX_windows_32-64.zip` file to the directory that you have created, for example, `C:\sgshare\sgdeploy`.

2. Create a deployment image

   Follow these steps to create a Windows installation image:

   a. Start the Toolkit Configuration Utility in the `C:\sgshare\sgdeploy` directory.

b. Select **Add Operating System Installation Files**, as shown in Figure 8-44.

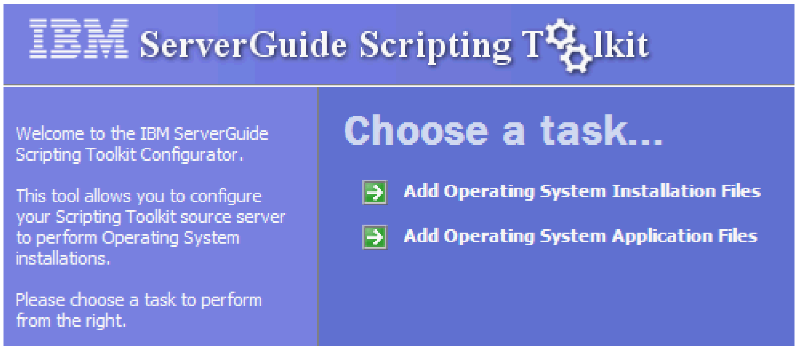


Figure 8-44 IBM ServerGuide Scripting Tool window

c. Choose the operating system type that you want and click **Next**, as shown in Figure 8-45.

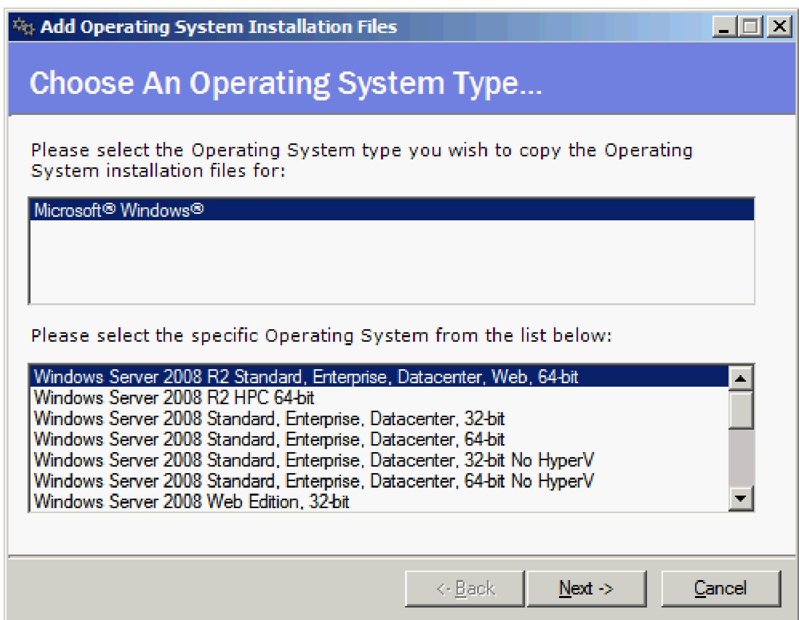


Figure 8-45 Select the type of the operating system

d. Insert the correct OS installation media or select the folder that contains the installation files for the source, as shown in Figure 8-46 on page 412. If necessary, modify the target and click **Next**.

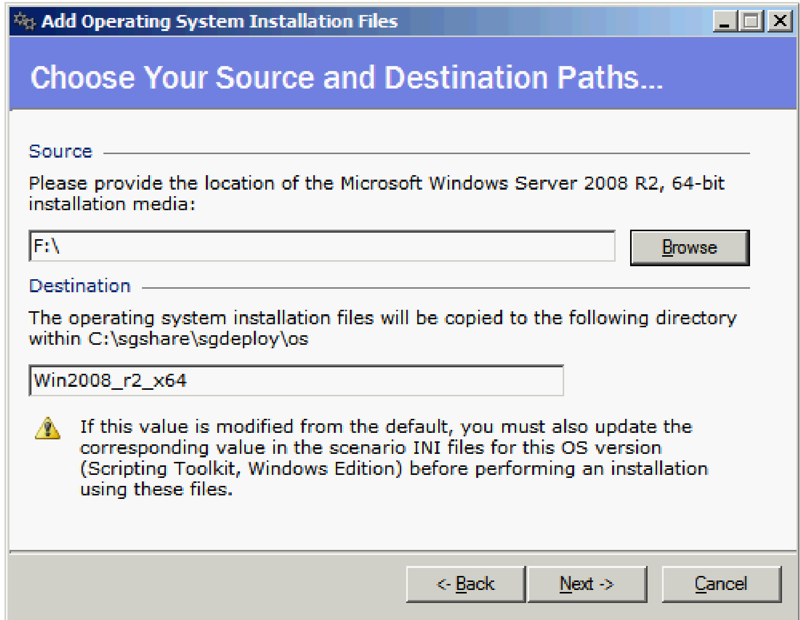


Figure 8-46 Define the source and target

e. When the file copy process completes, as shown in Figure 8-47, click **Finish**.

Figure 8-47 The file copy process completed successfully

f. Open a command prompt and change to the `C:\sgshare\sgdeploy\SGTKWinPE` directory. Use the following command to create the Windows installation image:

```
SGTKWinPE.cmd ScenarioINIs\Local\Win2008_R2_x64_EE.ini
```

g. When the process is finished, as shown in Figure 8-48 on page 413, your media creation software is started to create a bootable media from the image. Cancel this task.

```
18:26:21 - Creating the WinPE x64 ISO...

18:27:07 - The WinPE x64 ISO was created successfully.

*** WinPE x64 ISO:  c:\sgshare\sgdeploy\WinPE_ScenarioOutput\Local_Win2008_R2_x6
4_EE\WinPE_x64.iso

18:27:07 - Launching the registered software associated with ISO files...

*** Using ISO File:  c:\sgshare\sgdeploy\WinPE_ScenarioOutput\Local_Win2008_R2_x
64_EE\WinPE_x64.iso

18:27:08 - The WinPE x64 build process finished successfully.

SGTKWinPE complete.
```

Figure 8-48 Build process is finished

3. Prepare the USB key

   Follow these steps to create a bootable USB key with the Windows installation image that was created in step 2 on page 410:

   a. Insert your USB key.
   b. Use **diskpart** to format the USB key using FAT32. All files on the USB key will be deleted. At the command prompt, type the commands that are listed in Figure 8-49 on page 414.

```
C:\>diskpart
Microsoft DiskPart version 6.1.7600
Copyright (C) 1999-2008 Microsoft Corporation.
On computer:

DISKPART> list disk

  Disk ###  Status         Size     Free     Dyn  Gpt
  --------  -------------  -------  -------  ---  ---
  Disk 0    Online          271 GB      0 B        *
  Disk 1    Online          135 GB      0 B
  Disk 2    Online         7839 MB      0 B

DISKPART> select disk 2
Disk 2 is now the selected disk.

DISKPART> clean
DiskPart succeeded in cleaning the disk.

DISKPART> create partition primary
DiskPart succeeded in creating the specified partition.

DISKPART> select partition 1
Partition 1 is now the selected partition.

DISKPART> active
DiskPart marked the current partition as active.

DISKPART> format fs=fat32
  100 percent completed
DiskPart successfully formatted the volume.

DISKPART> assign
DiskPart successfully assigned the drive letter or mount point.

DISKPART> exit
```

Figure 8-49 Using diskpart to format the USB memory key

c. Copy the contents from `C:\sgshare\sgdeploy\WinPE_ScenarioOutput\Local_Win2008_R2_x64_EE\ISO` to the USB key. The USB key includes the folders and files that are shown in Figure 8-50.

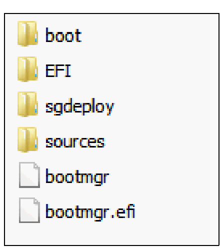


Figure 8-50 Contents of the USB key

d. Boot the target system from the USB key. The deployment executes automatically.

RAID controller: If the target system contains a RAID controller, RAID is configured as part of the installation.

Linux and VMware installation

The procedure for Linux and VMware is similar to the Windows procedure:

1. Install the ServerGuide Scripting Toolkit.
2. Create a deployment image.
3. Prepare a USB key.

For more information, see the *IBM ServerGuide Scripting Toolkit, Linux Edition User's Reference* at the following website:

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

Preboot eXecution Environment (PXE)

The Preboot eXecution Environment (PXE) is an environment to boot computers using a network interface for the operating system deployment. All eX5 systems support PXE.

For example, you can use the ServerGuide Scripting Toolkit. For more information, see the *IBM ServerGuide Scripting Toolkit User's Reference* at the following website:

http://www.ibm.com/support/docview.wss?uid=psg1SERV-TOOLKIT

Tivoli Provisioning Manager for OS Deployment

IBM Software has an offering for users needing advanced features in automating and managing remote deployment of operating systems and virtual images, in the form of Tivoli Provisioning Manager for OS Deployment. It is available as a stand-alone package and as an extension to IBM Systems Director.

You can obtain more information about these offerings at the following websites:

- http://ibm.com/software/tivoli/products/prov-mgr-os-deploy/
- http://ibm.com/software/tivoli/products/prov-mgr-osd-isd/

8.7.2 VMware ESXi on a USB key

VMware ESXi is an embedded version of VMware ESX. The footprint of ESXi is small (approximately 32 MB) because it does not use the Linux-based Service Console. Instead, it uses management tools, such as vCenter, the Remote Command-Line Interface, and Common Information Model (CIM) for standards-based and agent-less hardware monitoring. VMware ESXi includes full VMware File System (VMFS) support across Fibre Channel and iSCSI SAN, and network-attached storage (NAS). It supports 4-way virtual symmetric multiprocessing (SMP) (VSMP). ESXi 4.0 supports 64 CPU threads (for example, eight x 8-core CPUs) and can address 1 TB of RAM.

You can order the VMware ESXi 4.0 and 4.1 embedded virtualization keys from IBM. See Table 5-29 on page 214 for part number information.

Installing VMware ESXi

To successfully complete the installation of a supported version of ESXi onto the HX5, complete the following tasks:

1. Install the system memory in a balanced configuration

   When installing the ESXi operating system on the HX5, the memory must be balanced across all processors in the system. This rule applies to 2-socket, 4-socket, and HX5+MAX5 configurations. Failure to follow this rule prevents the operating system from starting correctly.

2. Set the boot order

   To ensure that you can boot ESXi successfully, you must change the boot order options in the UEFI. ESXi is not UEFI aware at present, and therefore, the *Legacy Only* option must be used for the first boot entry. The second boot entry must be the *Embedded Hypervisor*.

   Use the following steps to set the boot options and boot order in UEFI:

   a. Power on the system and press F1 when the UEFI splash panel displays.

   b. Select **Boot Manager** → **Add Boot Option**.

   c. Select **Legacy Only** and **Embedded Hypervisor**. If you cannot find these options, the options are already in the boot list. When finished, press Esc to go back one panel.

Figure 8-51 Add these boot options

   d. Select **Change Boot Order**. Change the boot order so that Legacy Only appears at the top of the list and the Embedded Hypervisor option appears beneath it, as shown in Figure 8-52.

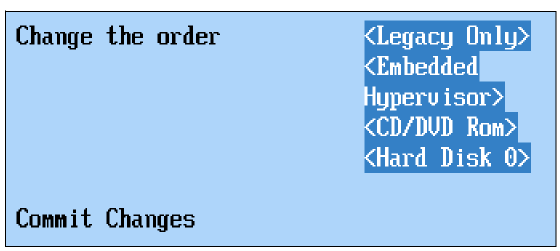


Figure 8-52 Example of a boot order

   e. Select **Commit Changes** and press Enter to save the changes.

3. Configure UEFI for embedded ESXi 4.1 if the MAX5 is attached

   Systems running VMware ESXi Server must use **Non-Pooled** mode in the MAX5 Memory Scaling option with the UEFI. See 8.5.5, "HX5 with MAX5 attached" on page 401 for instructions to configure the MAX5 Memory Scaling option.

4. Boot a new embedded ESXi 4.1 with MAX5 attached

   Use the following steps to successfully boot a new ESXi 4.1 embedded hypervisor on the HX5 with MAX5 attached:

   a. Ensure that you have the latest FPGA code installed on the HX5 by updating the FPGA code if required.

b. Physically attach the MAX5 by using the instructions that are provided in the *IBM BladeCenter HX5 Installation and User's Guide*, which is available at this website:

`http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084612`

c. Reflash the FPGA code to ensure that both the HX5 FPGA firmware and the MAX5 FPGA firmware are at the same level.

d. When the FPGA code has been flashed, enter UEFI by pressing F1 at the UEFI splash panel.

e. Boot the host.

f. In the Loading VMware Hypervisor panel, press Shift+O when the gray progress bar is displayed.

g. Enter the following command at the prompt:

```
esxcfg-advcfg -k TRUE allowInterleavedNUMAnodes
```

h. After the system boots, connect to the system using the vSphere Client.

i. Select the **Configuration** tab of the host and click **Advanced Settings** under the Software panel, as shown in Figure 8-53.

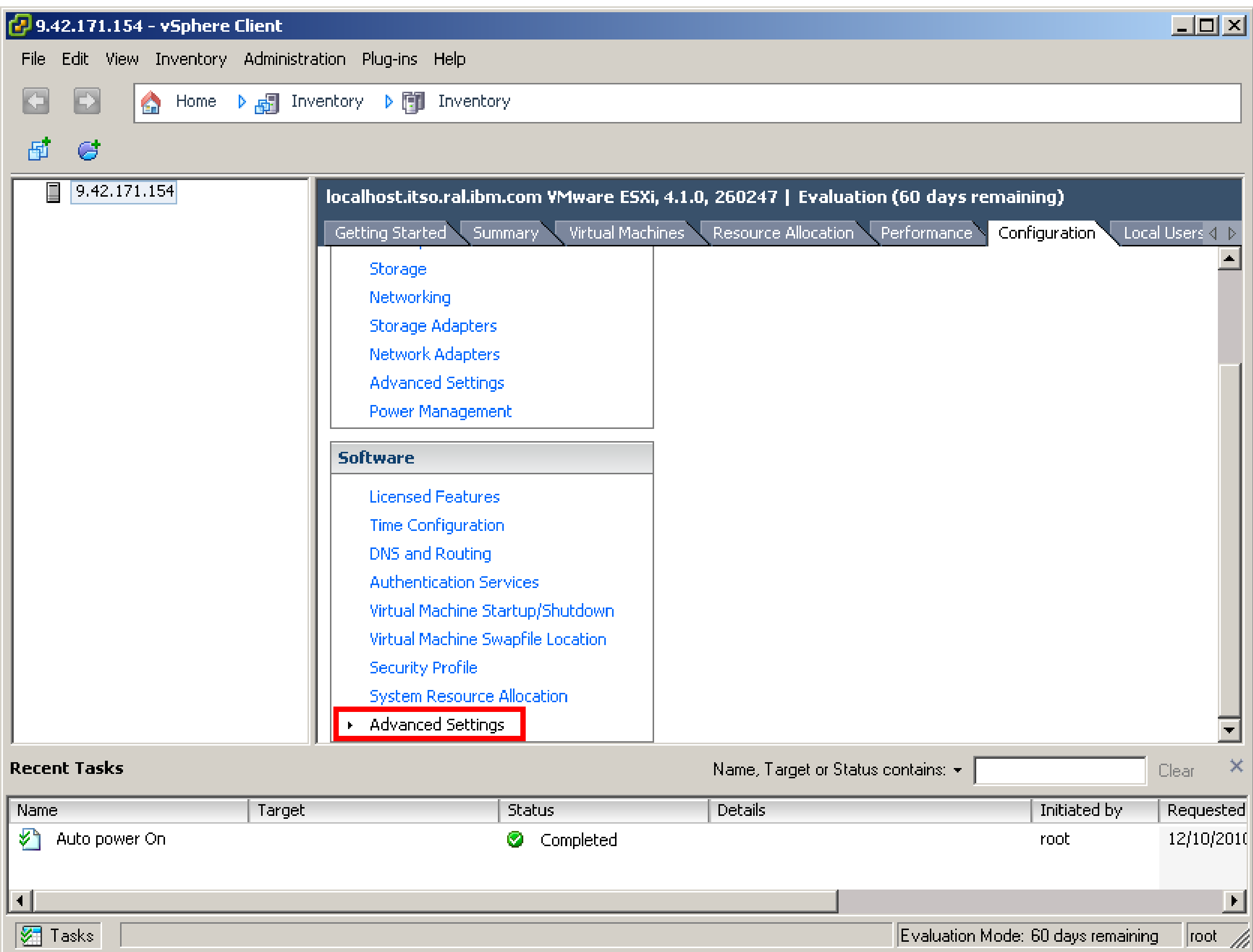


Figure 8-53 Editing the advanced settings in the vSphere Client

j. Click **VMkernel** in the left pane and select the check box next to **VMkernel.Boot.allowInterleavedNUMAnodes**, as shown in Figure 8-54 on page 418.

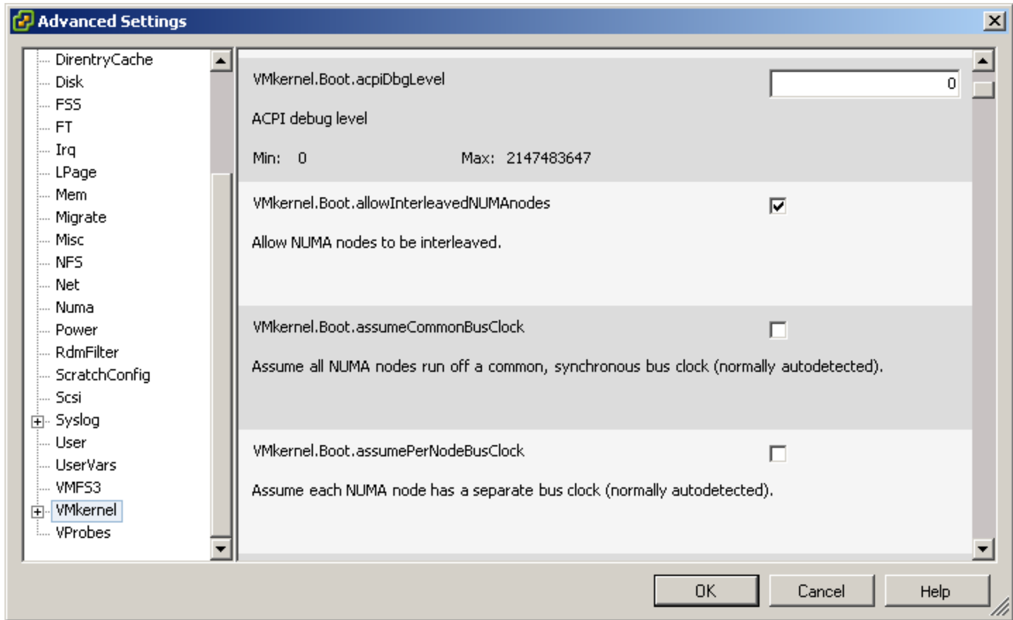


Figure 8-54 Editing the WMkernel settings in the vSphere Client

k. Click **OK**.

At this point, you have completed the process of configuring the embedded ESXi 4.1 on the HX5 with MAX5.

Attaching MAX5 to an existing embedded ESXi 4.1 configuration

The easiest way to add the MAX5 to an existing installation of ESX 4.1 or an ESXi 4.1 Installable running on an HX5 is to use the following steps:

1. Connect to the relevant ESX 4.1 or ESXi 4.1 Installable server by logging in to it using the VMware vSphere Client.
2. Select the **Configuration** tab of the host and click **Advanced Settings** in the Software panel, as shown in Figure 8-55 on page 419.

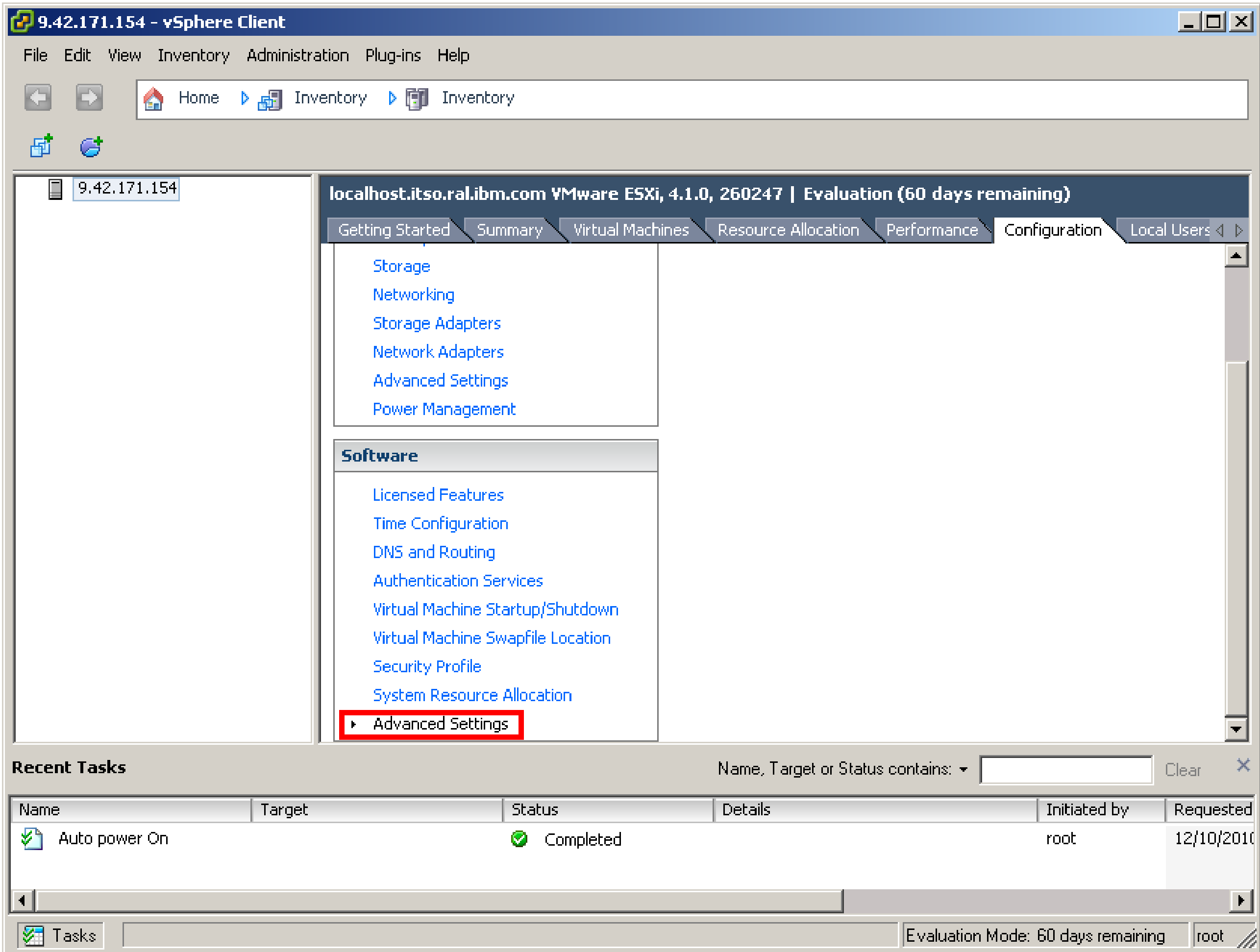


Figure 8-55 Configuration tab Software Advanced Settings

3. Click **VMkernel** in the left pane and select the check box next to **VMkernel.Boot.allowInterleavedNUMAnodes**, as shown in Figure 8-56 on page 420.

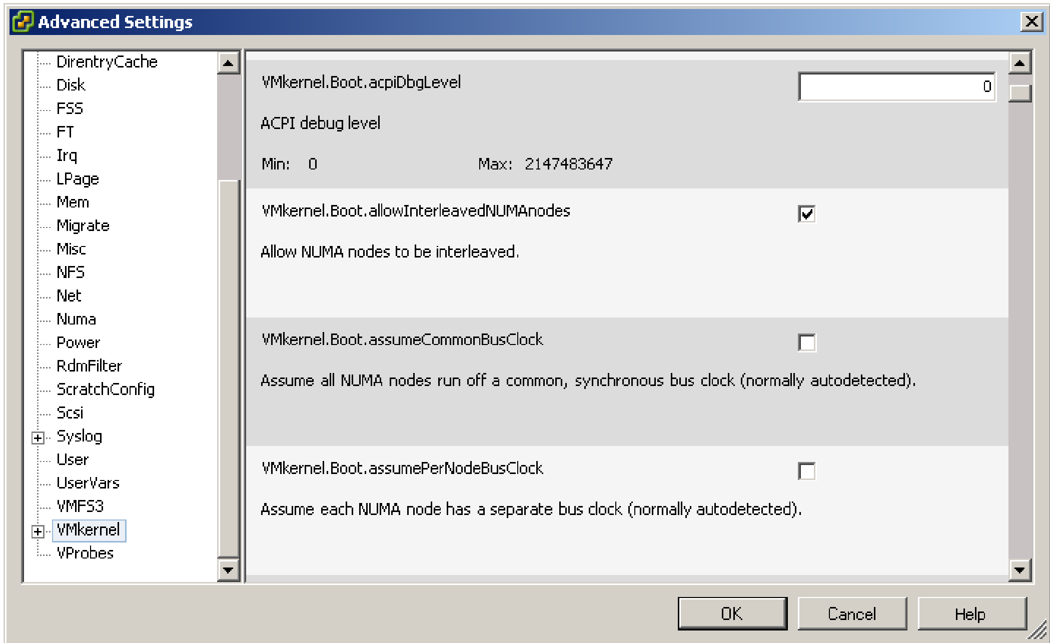


Figure 8-56 Editing the WMkernel settings in vSphere Client

4. Click **OK**.
5. Shut down the HX5.
6. Ensure that you have the latest FPGA code installed on the HX5 by updating it, if required.
7. Physically attach the MAX5 by using the instructions that are provided in the Installation Guide.
8. Reflash the FPGA code to ensure that both the HX5 FPGA code and MAX5 FPGA code are at the same level.
9. After the FPGA code has been flashed, log in to the UEFI by pressing F1 at the UEFI splash panel.
10. Select **System Settings** → **Memory** from the System Configuration and Boot Management main menu.
11. Ensure that the MAX5 Memory Scaling option is set to **Non-Pooled**.
12. Continue to boot the server normally. Log in to the vSphere Client and check that the additional memory shows on the Systems Summary tab.

Restoring ESXi to the factory defaults

You can use the IBM recovery CD to recover the IBM USB Memory Key to a factory-installed state. Table 8-5 shows the available CDs.

Table 8-5 VMware ESXi recovery CD

Part number	Description
68Y9634	VMware ESXi 4.0 U1
49Y8747	VMware ESXi 4
68Y9633	VMware ESX Server 3i v 3.5 Update 5
46M9238	VMware ESX Server 3i v 3.5 Update 4

Part number	Description
46M9237	VMware ESX Server 3i v 3.5 Update 3
46M9236	VMware ESX Server 3i v 3.5 Update 2
46D0762	VMware ESX Server 3i version 3.5

To order a recovery CD, contact your local support center at the following website:

http://www.ibm.com/planetwide/region.html

Updating ESXi

You can install the latest version of ESXi 4 on IBM Hypervisor keys, and it is supported by IBM. Use the following VMware upgrade mechanisms for the update:

- VMware Upgrade Manager
- Host Update Utility

For more information, see the VMware Documentation website:

http://www.vmware.com/support/pubs

8.7.3 Installing ESX 4.1 or ESXi 4.1 Installable onto HX5

Before installing any VMware operating system, see the latest operating system support information that is contained on the IBM ServerProven website. You can access the IBM ServerProven information at the following website:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/vmwaree.html

The IBM ServerProven website provides general operating system support information for the HX5. The information that is contained in Table 8-6 provides VMware version-specific support for the various HX5 hardware configurations.

Table 8-6 Supported VMware operating system versions based on HX5 hardware configuration

WMware operating system	Two-socket HX5	Four-socket HX5	HX5 with MAX5
VMware ESX Server 4.0 Update 1	Yes	Yes	No
VMware ESXi Server 4.0 Update 1	Yes	No	No
VMware ESX Server 4.1	Yes	Yes	Yes
VMware ESXi Server 4.1	Yes	Yes	Yes

When installing any supported version of the ESX server operating system onto the HX5, the memory must be balanced across all processors in the system. This rule applies to 2-socket and 4-socket HX5 configurations, as well as the HX5 with MAX5 attached. Failure to follow this rule prevents the operating system from installing correctly.

Installing on an HX5 with MAX5 attached

To correctly configure and install the ESX 4.1 or ESXi 4.1 Installable editions of the ESX server, follow these instructions.

Common steps for both ESX 4.1 and ESXi 4.1 Installable editions

Perform the following steps for both operating system types:

1. Ensure that you have the latest FPGA code installed on the HX5 by updating it, if required.
2. Physically attach the MAX5 using the instructions that are provided in the *IBM BladeCenter HX5 Installation and User's Guide*, which is available at this website:

   `http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084612`
3. Reflash the FPGA code to ensure that both the HX5 FPGA code and MAX5 FPGA code are at the same level.
4. When the FPGA code has been flashed, log in to UEFI by pressing F1 at the UEFI splash panel.
5. Select **System Settings** → **Memory** from the System Configuration and Boot Management main menu.
6. Ensure that the MAX5 Memory Scaling option is set to **Non-Pooled**.

Exit the UEFI when finished and proceed to the installation of the respective version of ESX server.

For ESX 4.1 installations

To correctly install ESX 4.1 onto an HX5 with a MAX5 attached, complete the following steps:

1. Set the boot order

   > **Configure RAID before proceeding:** You must have already configured an existing RAID array before setting the boot order for ESX 4.1.

   To ensure that you can install and boot ESX 4.1 successfully, you must change the boot order options in the UEFI. ESX server is not UEFI aware at present, and therefore, you must use the **Legacy Only** option for the first bootable device entry after the CD/DVD Rom. The next boot entry must be the hard disk to which you are installing.

   Use the following steps to set the boot options and boot order in UEFI:

   a. Power on the system and press F1 when the UEFI splash panel displays.

   b. Select **Boot Manager** → **Add Boot Option**.

   c. Select **CD/DVD Rom**, **Legacy Only**, and **Hard Disk 0**, as shown in Figure 8-57 on page 423, for example, if you use internal drives that have been configured as an integrated mirror with the onboard RAID controller. If you cannot find these boot options, the options are already in the boot list.

   d. Press Esc when finished to go back one panel.

   e. Select **Change Boot Order**. Change the boot order to the boot order that is shown in Figure 8-52 on page 416.

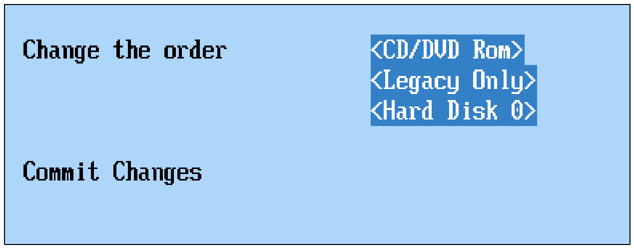


Figure 8-57 Example of a boot order

   f. Select **Commit Changes** and press Enter to save the changes.

2. Boot the host from the ESX installation media.
3. Press F2 when you receive the installation options panel, as shown in Figure 8-58.

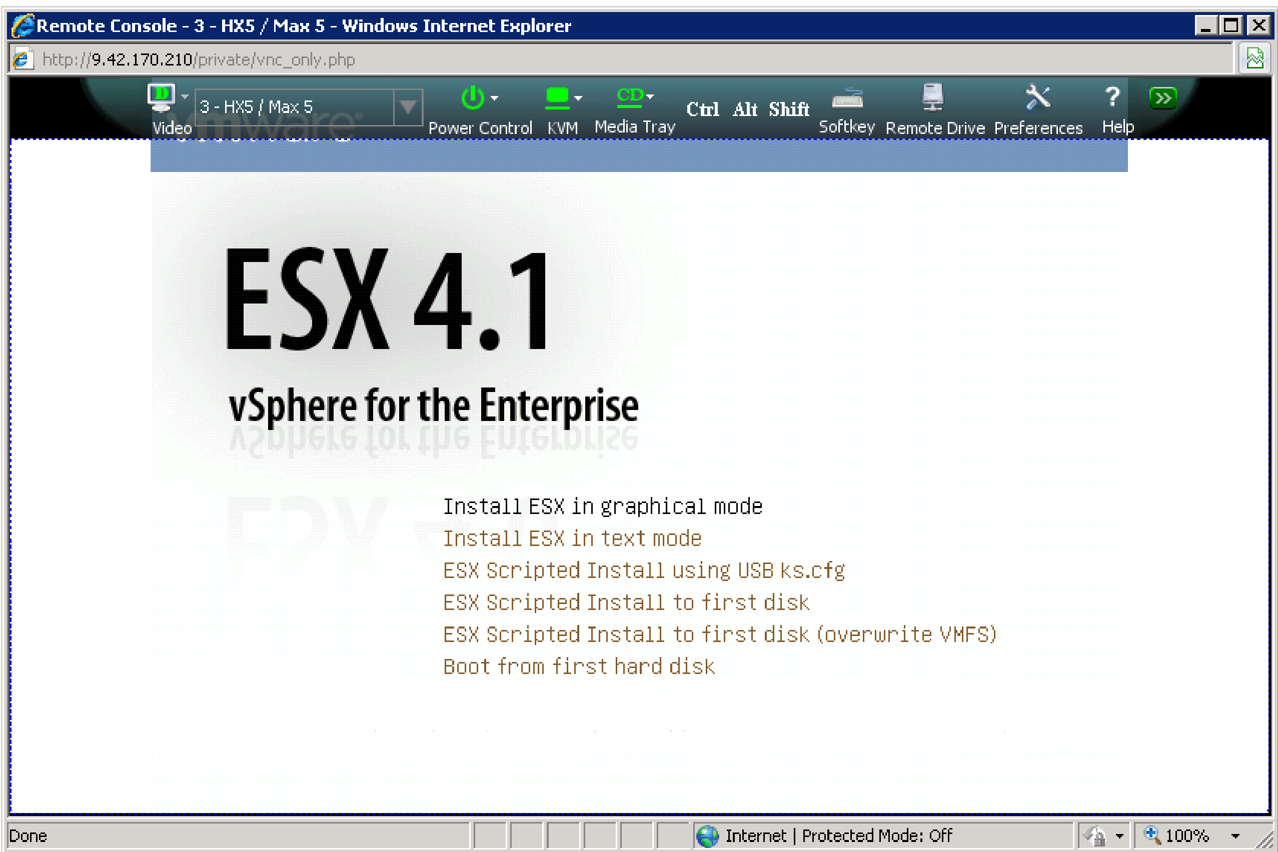


Figure 8-58 ESX installation options panel

4. The Boot Options line appears on the panel. Type the following parameter at the end of the Boot Options line:

```
allowInterleavedNUMAnodes=TRUE
```

The edited result looks like Figure 8-59 on page 424. Press Enter to proceed.

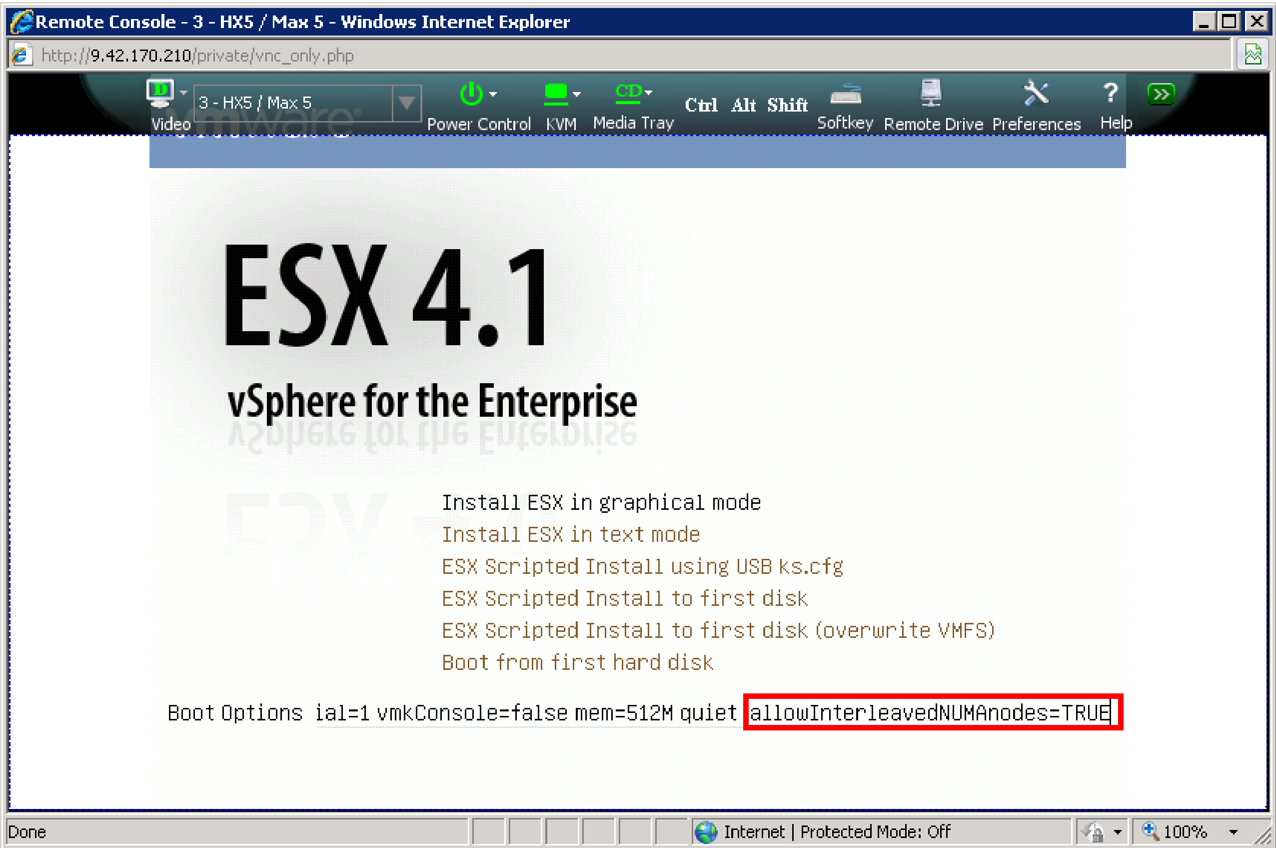


Figure 8-59 Editing the boot options

5. Proceed through the installer until you reach the Setup Type page. Click **Advanced setup** and clear **Configure boot loader automatically (leave checked if unsure)**, as shown in Figure 8-60 on page 425.

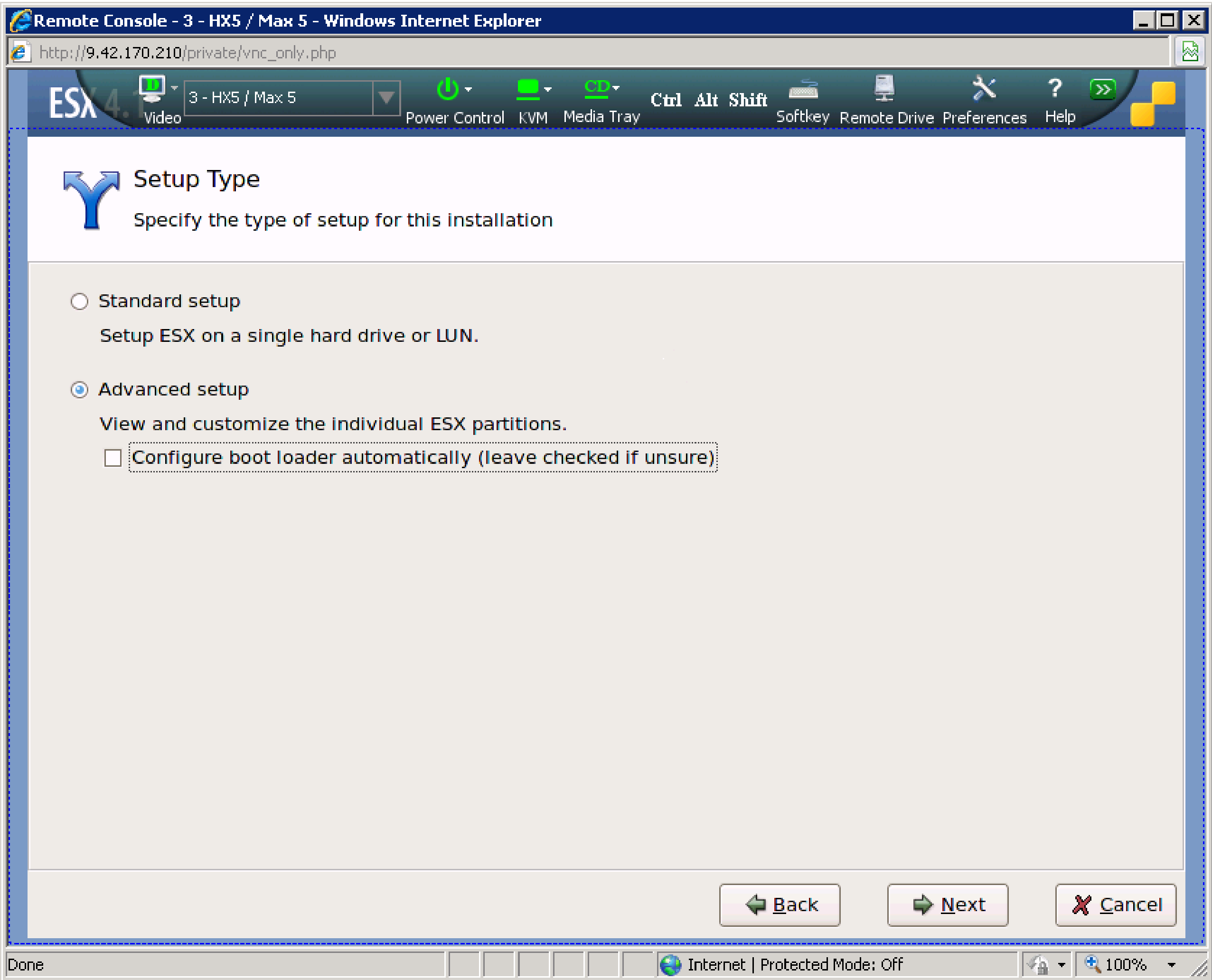


Figure 8-60 Modifying the ESX installation

6. Proceed through the installer to the Set Boot loader Options page and type the following parameter in the Kernel Arguments text box, as shown in Figure 8-61 on page 426:

```
allowInterleavedNUMAnodes=TRUE
```

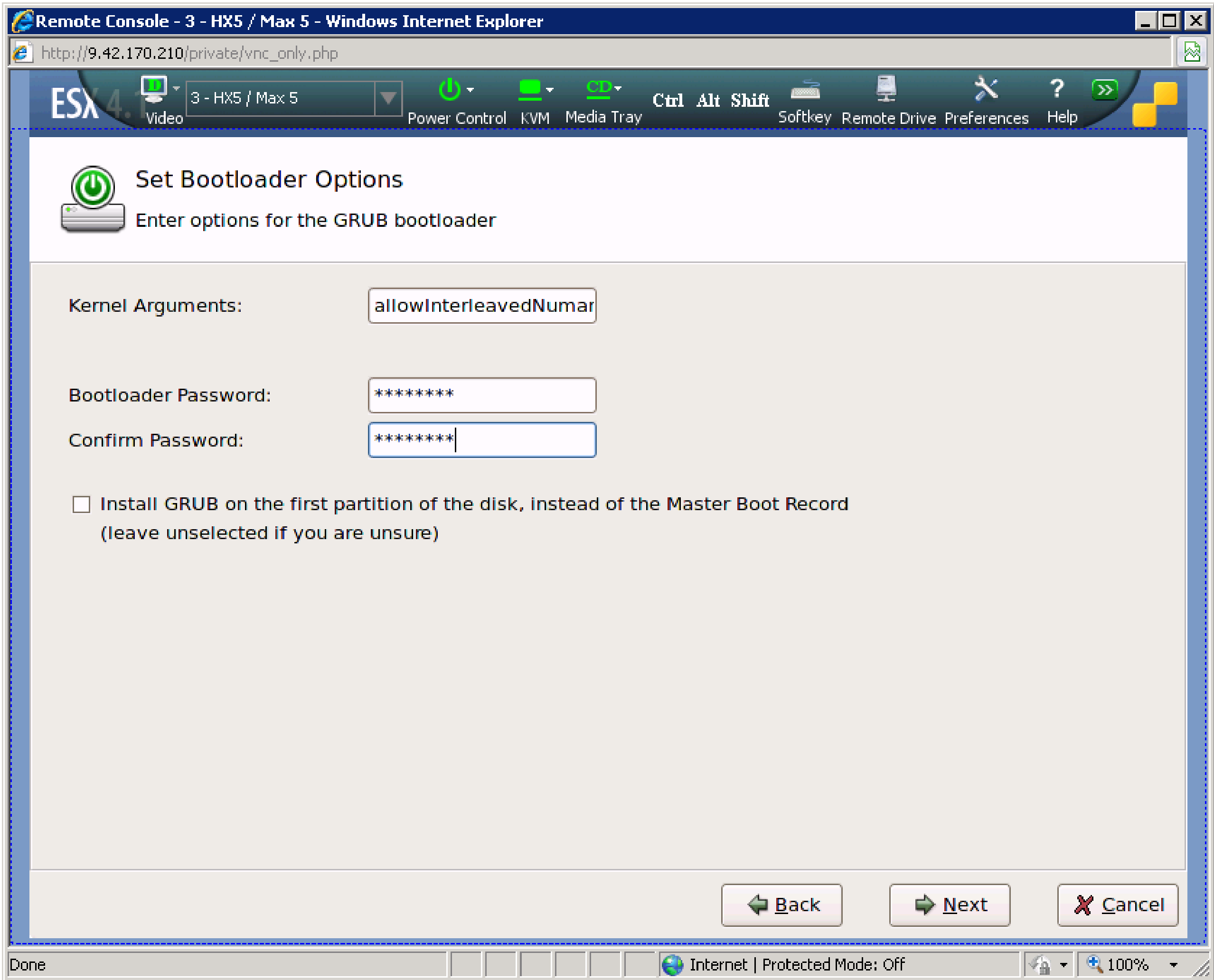


Figure 8-61 Editing the boot loader options

7. Complete the remainder of the ESX installation and reboot the host.

For ESXi 4.1 Installable edition hosts

Use the following steps to install the ESXi 4.1 Installable onto the HX5 with MAX5 attached.

To ensure that you can install and boot ESXi 4.1 successfully, you must change the boot order options in the UEFI. The ESXi server is not UEFI aware at present, and therefore, you must use the **Legacy Only** option for the first bootable device entry after the CD/DVD Rom. The next boot entry must be the hard disk to which you are installing. Perform these steps:

1. Set the boot order

> **Configure RAID before proceeding:** You must have already configured an existing RAID array before commencing the boot order steps for the ESXi 4.1 Installable edition.

Use the following steps to set the boot options and boot order in UEFI:

a. Power on the system and press F1 when the UEFI splash panel is shown.

b. Select **Boot Manager** → **Add Boot Option**.

c. Select **CD/DVD Rom**, **Legacy Only**, and **Hard Disk 0**, as shown in Figure 8-62, for example, if you are using internal drives that have been configured as an integrated mirror with the onboard RAID controller. If you cannot find these options, the options are already in the boot list.

d. Press Esc when finished to go back one panel.

e. Select **Change Boot Order**. Change the boot order to look like the boot order that is shown in Figure 8-62.

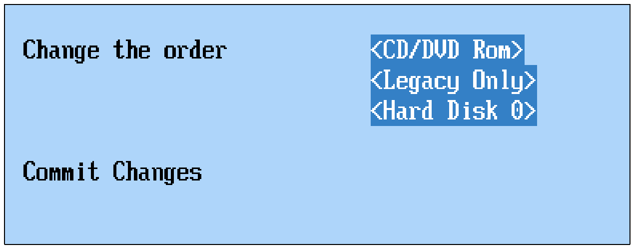


Figure 8-62 Example of a boot order

f. Select **Commit Changes** and press Enter to save the changes.

2. Boot from the ESXi Installable installation media.
3. Press the Tab key when the blue boot panel appears, as shown in Figure 8-63.

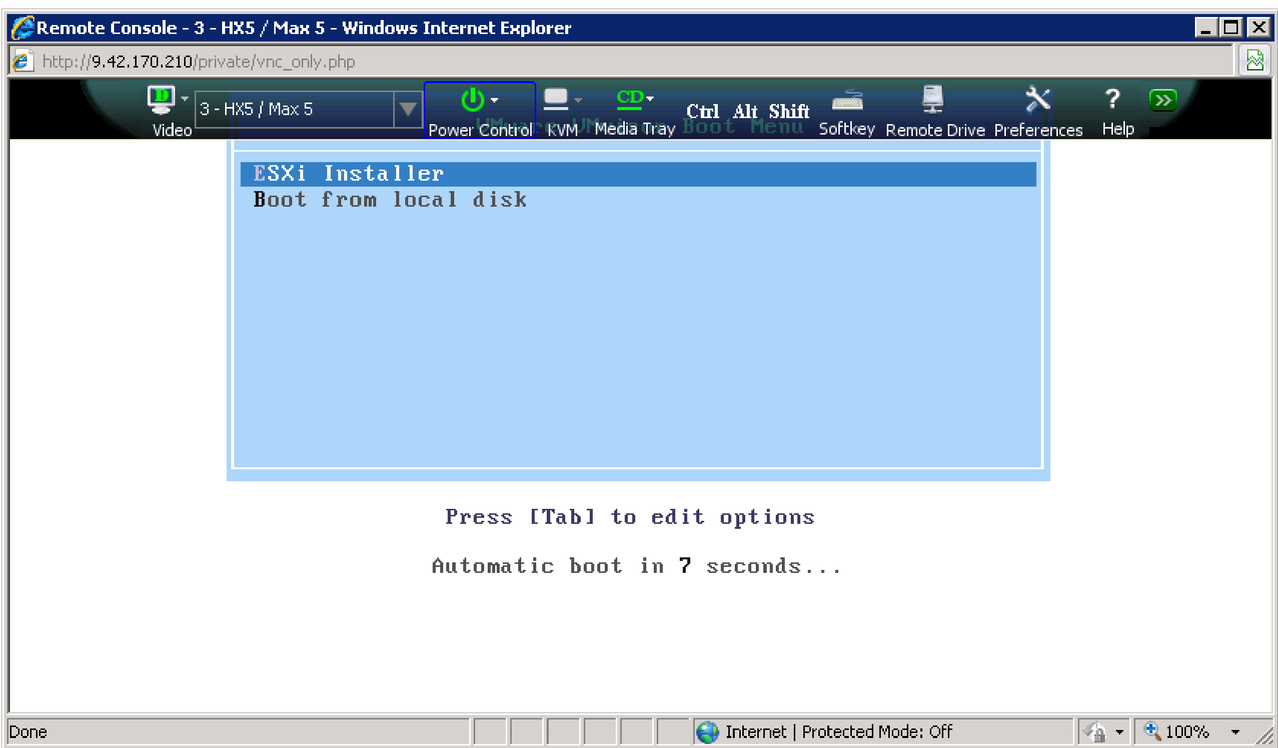


Figure 8-63 Installing ESXi 4.1 Installable edition

4. Add the following line after `vmkboot.gz`:

   `allowInterleavedNUMAnodes=TRUE`

   Ensure that you leave a space at the beginning and the end of the text that you enter, as shown in Figure 8-64 on page 428. Otherwise, the command will fail to execute at a later stage during the installation. Press Enter to proceed.

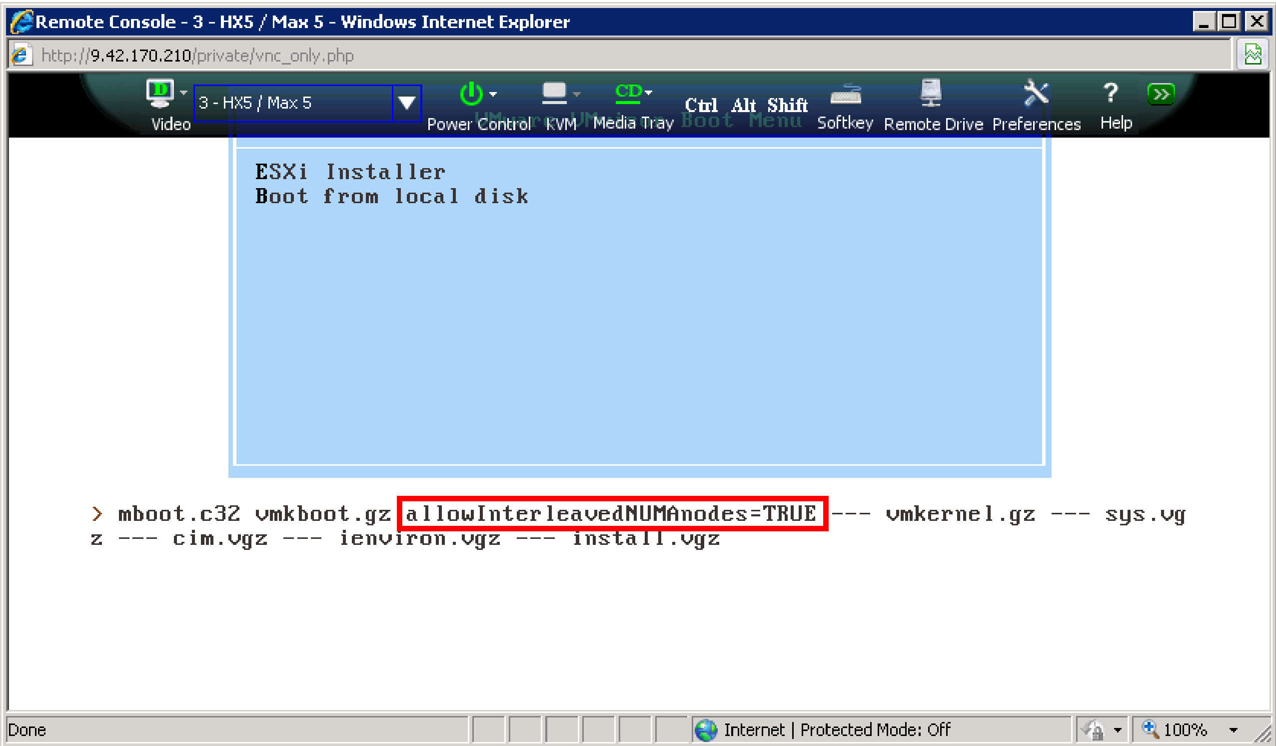


Figure 8-64 Editing the boot load command

5. Complete the ESXi installation and reboot when prompted. Ensure that you remove the media or unmount the installation image before the system restarts.
6. In the Loading VMware Hypervisor panel, when the gray progress bar is displayed, press Shift+O.

> **Note:** If you do not press Shift+O during the Loading VMware Hypervisor panel, you receive this error:
>
> `“The system has found a problem on your machine and cannot continue. Interleave NUMA nodes are not supported.”`

7. Enter the following command at the prompt after you have pressed Shift+O:

```
esxcfg-advcfg -k TRUE allowInterleavedNUMAnodes
```

   Your output looks like Figure 8-65 on page 429.

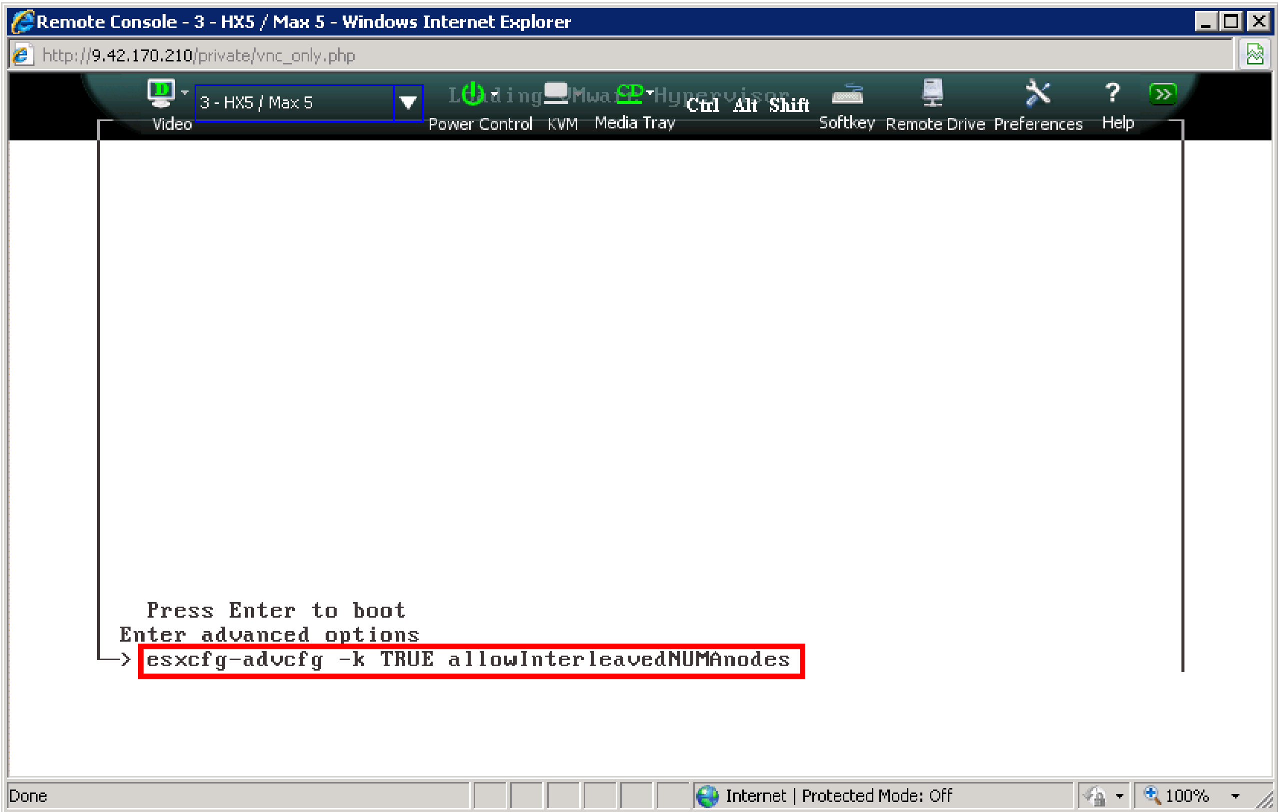


Figure 8-65 Loading VMware Hypervisor boot command

8. Press Enter after the command has been entered. Press Enter again to continue to boot.
9. After the system boots, connect to it using the vSphere Client.
10. Select the **Configuration** tab of the host and click **Advanced Settings** under Software, as shown in Figure 8-66 on page 430.

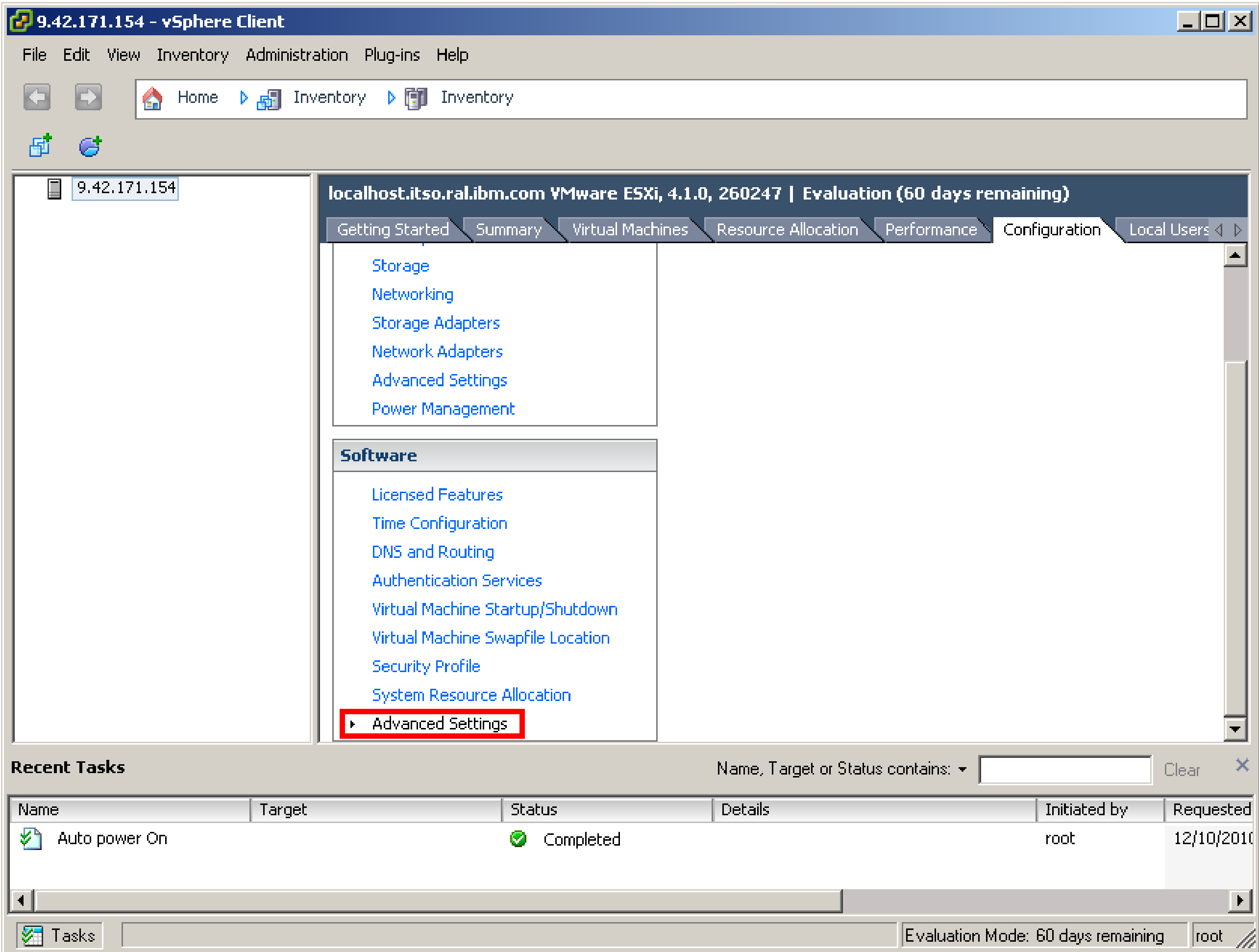


Figure 8-66 Configuration tab Software Advanced Settings

11. Click **VMkernel** in the left pane and select the check box next to **VMkernel.Boot.allowInterleavedNUMAnodes**, as shown in Figure 8-67 on page 431.

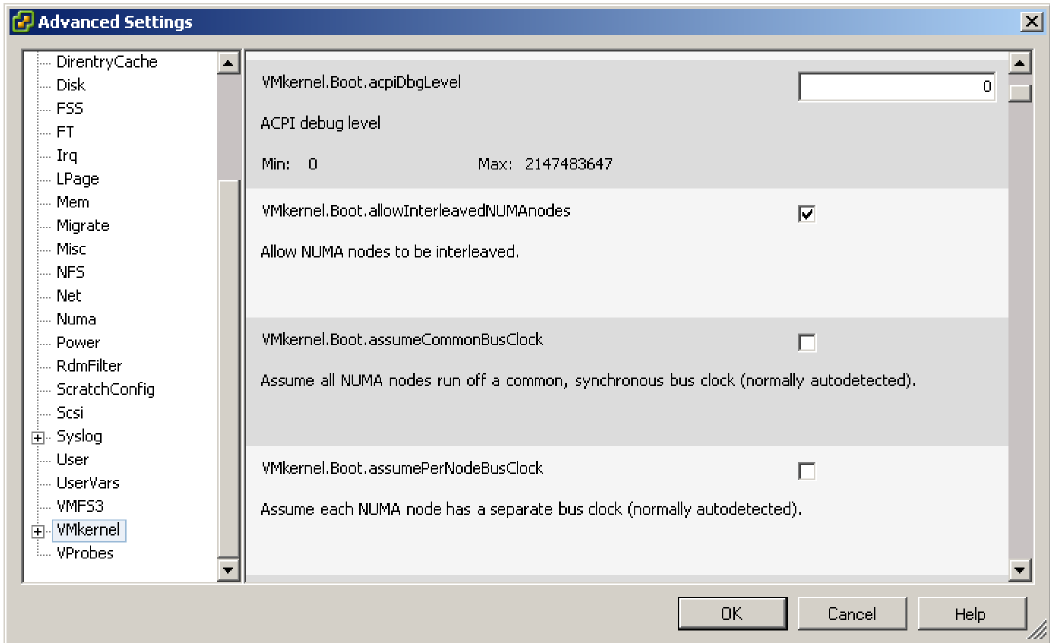


Figure 8-67 Editing the VMkernel settings in vSphere Client

12. Click **OK**. At this point, you have completed the process of installing ESXi 4.1 Installable on the HX5 with MAX5.

Attaching MAX5 to HX5 with an existing installation of ESX 4.1 or ESXi 4.1 Installable edition

The easiest way to add the MAX5 to an existing installation of the ESX 4.1 or ESXi 4.1 Installable running on an HX5 is by using the following steps:

1. Connect to the relevant ESX 4.1 or ESXi 4.1 Installable server by logging in to it using the VMware vSphere Client.
2. Select the **Configuration** tab of the host and click **Advanced Settings** under Software.
3. Click **VMkernel** in the left pane and select the check box next to **VMkernel.Boot.allowInterleavedNUMAnodes**, as shown in Figure 8-68 on page 432. Click **OK** when finished.

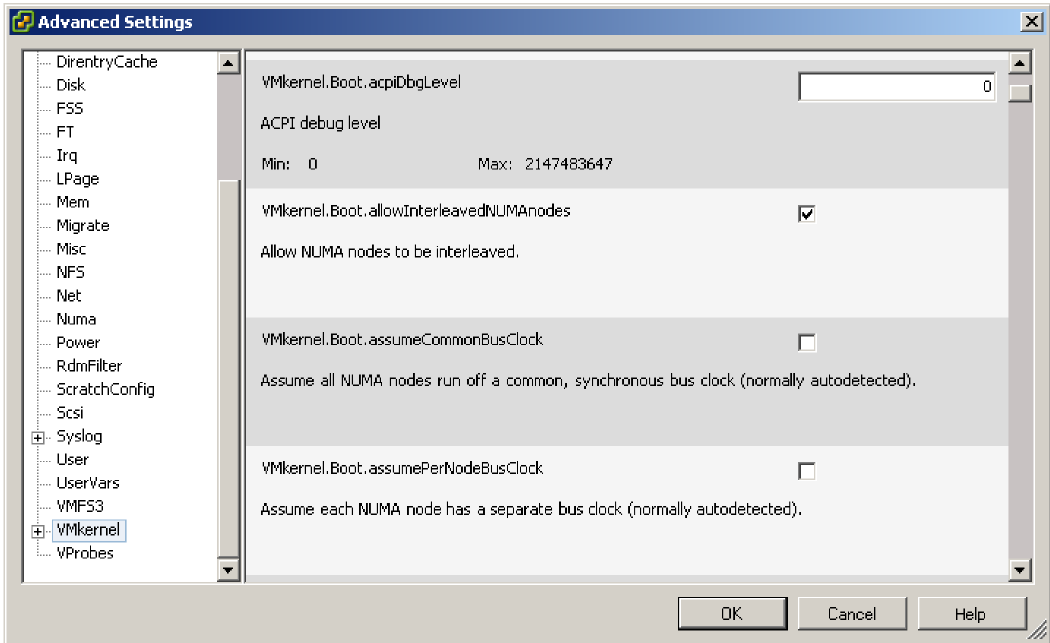


Figure 8-68 Editing the WMkernel settings in vSphere Client

4. Shut down the HX5.
5. Ensure that you have the latest FPGA code installed on the HX5 by updating it, if required.
6. Physically attach the MAX5 using the instructions that are provided in the Installation Guide.
7. Reflash the FPGA code to ensure that both the HX5FPGA code and the MAX5 FPGA code are at the same level.
8. After the FPGA code has been flashed, log in to UEFI by pressing F1 at the UEFI splash panel.
9. Select **System Settings** → **Memory** from the System Configuration and Boot Management main menu.
10. Ensure that the MAX5 Memory Scaling option is set to **Non-Pooled**.
11. Continue to boot the server normally. Log in to the vSphere Client and check that the additional memory shows on the Systems Summary tab.

Adding MAX5 to HX5 with an existing installation of ESX 4.1 without first editing the VMkernel before attaching

Use the following steps if you have already shut down the existing ESX 4.1 server without having made the changes that are detailed in “Attaching MAX5 to HX5 with an existing installation of ESX 4.1 or ESXi 4.1 Installable edition” on page 431:

1. Power on the HX5.
2. At the VMware boot loader panel, as shown in Figure 8-69 on page 433, ensure that **VMware ESX 4.1** is highlighted and press the a key to modify the kernel arguments.

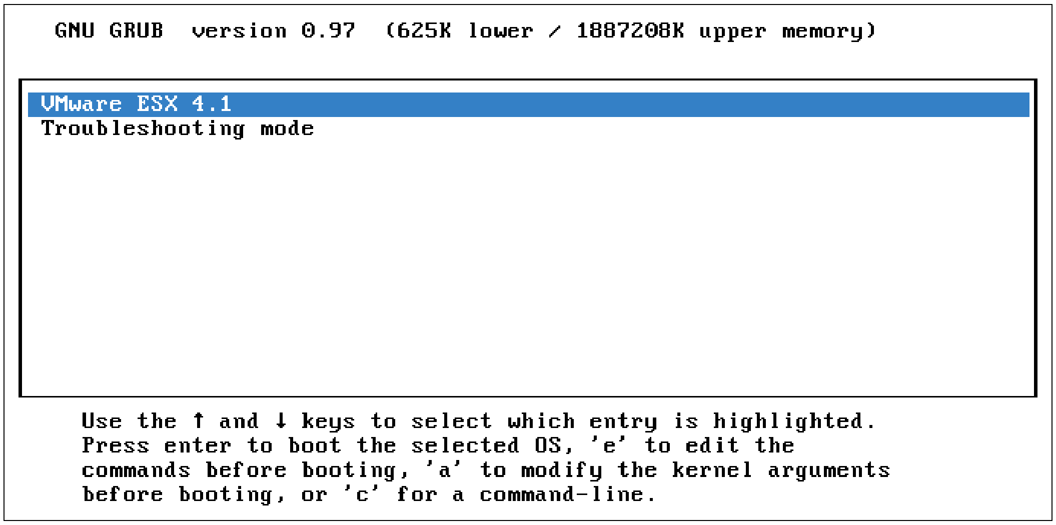


Figure 8-69 VMware ESX GRUB

3. Add the following line at the beginning of the boot load command:

   `allowInterleavedNUMAnodes=TRUE`

   Ensure that you leave a space at the beginning and at the end of the text that you enter, as shown in Figure 8-70. Otherwise, the command will fail to execute at a later stage during the boot process. Press Enter to proceed after you have correctly edited the line to continue the boot process.

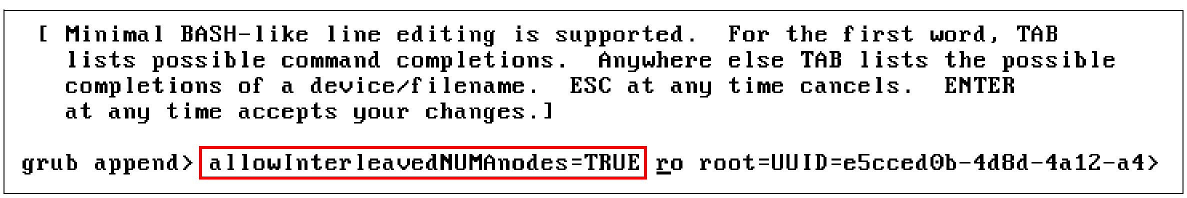


Figure 8-70 Editing the boot load command

4. When the ESX 4.1 operating system has completely loaded, proceed with step 1 on page 431 through step 3 on page 431 to complete the process.

Adding MAX5 to HX5 with an existing installation of ESXi 4.1 Installable without first editing the VMkernel before attaching

If you have already shut down the existing ESXi 4.1 server without having made the changes that are detailed in “Attaching MAX5 to HX5 with an existing installation of ESX 4.1 or ESXi 4.1 Installable edition” on page 431, complete step 6 on page 428 through step 11 on page 430.

More useful VMware links and tips

We provide the following links to further assist you with configuring and troubleshooting VMware with HX5:

- For configuration maximums on VMware vSphere 4.1, see the following website:

  http://www.vmware.com/pdf/vsphere4/r41/vsp_41_config_max.pdf

- We recommend that you review all of the RETAIN tips, which relate to this system and the operating system that you are installing, which are available at the following website:

  http://ibm.com/support/entry/portal/Troubleshooting/Hardware/Systems/BladeCenter/BladeCenter_HX5/7872

- For further general VMware tuning information, see the paper *VMware vCenter Server Performance and Best Practices for vSphere 4.1*, which is available at this website:

  http://www.vmware.com/resources/techresources/10145

8.7.4 Windows installation tips and settings

In this section, we provide useful information to assist with the installation of Windows Server 2008. We have not described the installation process in this chapter, because there are no particular deviations from a standard Windows 2008 installation. However, we provide links to useful installation instructions and troubleshooting information for certain configurations.

The latest versions of Microsoft Windows 2008 are all UEFI compliant and can therefore take advantage of the full functionality that is provided by UEFI. The HX5 only supports x64-bit versions of the Windows 2008 operating system. You can obtain a complete list of supported Windows Server operating systems for HX5 at the following website:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/microsofte.html

Windows Server 2008 and Windows 2008 R2 with MAX5

Consider these points when installing Windows Server 2008:

- When running Windows Server on a memory-expanded HX5, set the MAX5 Memory Scaling option to **Non-Pooled**. See 8.5.5, “HX5 with MAX5 attached” on page 401 for instructions.
- Use at least the Enterprise Edition of Windows 2008 or Windows 2008 R2 to take full advantage of the MAX5 expansion memory capability. You can obtain information about Windows Server memory limits at the following website:

  http://msdn.microsoft.com/en-us/library/aa366778.aspx

Installing Windows 2008 with IBM ServerGuide

Consider these points when installing with ServerGuide:

- We recommend using IBM ServerGuide to simplify the installation of Windows 2008 Server. It integrates RAID-array configuration and driver installation into the operating system installation process. You can download the IBM ServerGuide image from the following website:

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

- IBM ServerGuide automatically detects whether to install in Legacy mode or native UEFI mode, depending on the boot settings that are defined in the UEFI. The operating system will install in UEFI mode unless Legacy Only is the first entry in the boot order.
- You can use IBM ServerGuide v8.31 x64 bit or later to install Windows 2008 on the HX5.
- When using ServerGuide, you set the RAID array configuration in the Configure RAID Adapter panel. Choose **Keep Current Adapter Configuration** if the RAID array has been set previously and the next installation will use that current configuration.
- In the IBM ServerGuide partition setup panel, selecting the **Clear All Disks** check box proceeds with clearing the contents of all of the disks that are attached to the system. This deletion *includes SAN storage* logical drives that are mapped to the system. Ensure that

the HX5 has not been zoned to external SAN storage at the time of installation unless you are performing boot from SAN. *Failure to follow this rule potentially will result in unwanted data loss.*

Installing Microsoft Windows Server 2008 R2 on HX5 without the use of ServerGuide

The following website provides specific information about installing Microsoft Windows Server 2008 R2 without using the ServerGuide CD:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5085004

You can access more information about the Windows Server 2008 R2 installation and configuration at the following website:

http://www.microsoft.com/windowsserver2008/en/us/product-documentation.aspx

Troubleshooting information for Windows 2008: All versions

You must review all RETAIN tips that relate to the HX5 and the operating system that you are installing. Go to the following website for troubleshooting information relating to HX5 and Windows 2008 versions:

http://ibm.com/support/entry/portal/Troubleshooting/Hardware/Systems/BladeCenter/BladeCenter_HX5/7872

Be aware of these scenarios when performing an installation of Windows 2008 R2 on an HX5:

- To correct an issue where Windows 2008 R2 fails to boot on UEFI systems with raw disks attached, see the following Microsoft Hotfix at the following website:

  http://support.microsoft.com/kb/975535

- If your system has greater than 128 GB of memory and you plan to enable Hyper-V after installing Windows 2008 R2, you must first apply the Microsoft Hotfix from the following website:

  http://support.microsoft.com/kb/979903/en-us

Performance tuning for Windows 2008

The following links provide specific information about performance tuning the Windows 2008 operating system:

- Performance tuning recommendations for Windows 2008

  For general Windows 2008 performance configuration guidance, see the Performance Tuning Guidelines for Windows Server 2008 at the following website:

  http://www.microsoft.com/whdc/system/sysperf/Perf_tun_srv.mspx

- Performance tuning recommendations for Windows 2008 R2

  For general Windows 2008 R2 performance configuration guidance, see the Performance Tuning Guidelines for Windows Server 2008 R2 at the following website:

  http://www.microsoft.com/whdc/system/sysperf/Perf_tun_srv-R2.mspx

- Windows software, firmware, drivers, and fixes for HX5

  See the following website for all the latest information relating to software, firmware, drivers, and fixes for HX5 when used with Windows 2008:

  http://www-933.ibm.com/support/fixcentral/systemx/quickorder?parent=ibm~BladeCenterHX5&product=ibm/systemx/7872&&platform=Windows+2008+x64&function=all&source=fc

8.7.5 Red Hat Enterprise Linux installation tips and settings

This section provides useful information to assist you with the installation of Red Hat Enterprise Linux (RHEL). We do not describe the Red Hat installation process in this chapter, because there are no particular deviations from a standard Red Hat installation. However, we provide links to useful installation instructions and troubleshooting information for certain configurations.

The HX5 supports only the 64-bit versions of Red Hat Enterprise Linux. For a list of supported versions of Red Hat Linux on an HX5, see IBM ServerProven:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/redchate.html

Using Red Hat Linux with MAX5

When running Red Hat Linux on a memory-expanded HX5, set the MAX5 Memory Scaling option to **Pooled**. See 8.5.5, “HX5 with MAX5 attached” on page 401 for instructions.

Installing Red Hat Linux on HX5

The installation of Red Hat Linux is straightforward, because all drivers that are required for the HX5 are included in the kernel. We provide the following operating system links for your benefit:

- *Installing Red Hat Enterprise Linux Version 6 - IBM BladeCenter HX5 (Type 7872)*

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5086419

- *Installing Red Hat Enterprise Linux Version 5 Update 5 - IBM BladeCenter HX5 (Type 7872)*

  http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5085426

Note the following considerations when installing Red Hat Linux:

- RHEL6 is an UEFI-compliant operating system; therefore, it will add its boot file to the top of the system’s UEFI boot order.
- We highly recommend that you do *not* attach SAN storage to the HX5 during the installation process, because this action can lead to accidental data loss at the disk partitioning phase.

You can access more information about the installation and configuration of Red Hat Enterprise Linux at the following website:

https://access.redhat.com/knowledge/docs/manuals/Red_Hat_Enterprise_Linux/

You can download Red Hat Enterprise Linux 6 evaluation software from this website:

https://access.redhat.com/downloads/

Troubleshooting information for HX5 and Red Hat Linux

We recommend that you review all RETAIN tips that relate to HX5 and the operating system that you are installing. Refer the following website for troubleshooting information pertaining to HX5:

http://ibm.com/support/entry/portal/Downloads/Hardware/Systems/BladeCenter/BladeCenter_HX5

Red Hat Linux 5 software, firmware, drivers, and fixes for HX5

See the following website for all the latest information relating to software, firmware, drivers, and fixes for HX5 when used with Red Hat Linux 5:

http://www-933.ibm.com/support/fixcentral/systemx/quickorder?parent=ibm~BladeCenterHX5&product=ibm/systemx/7872&&platform=RHEL+5+x64&function=all&source=fc

8.7.6 SUSE Linux Enterprise Server installation tips and settings

SUSE Linux Enterprise Server

In this section, we provide useful information to assist you with the installation of SUSE Linux Enterprise Server. We do not describe the installation process in this chapter, because there are no particular deviations from a standard SUSE Linux Enterprise Server installation. However, we provide links to useful installation instructions and troubleshooting information for certain configurations.

The HX5 supports only the 64-bit versions of SUSE Linux Enterprise Server. For a list of supported versions of Red Hat Linux for HX5, see IBM ServerProven on the following website:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/suseclinuxe.html

Using SUSE Linux Enterprise Server with MAX5

When running SUSE Linux Enterprise Server on a memory-expanded HX5, set the MAX5 Memory Scaling option to **Pooled**. See 8.5.5, “HX5 with MAX5 attached” on page 401 for instructions.

Installing SUSE Linux Enterprise Server on HX5

The installation of SUSE Linux Enterprise Server is straightforward, because all drivers that are required for the HX5 are included in the kernel. The installation instructions for installing SUSE Linux Enterprise Server on the HX5 are identical to the instructions for installing the operating system on an x3850 X5. See *Installing SUSE Linux Enterprise Server 11 - IBM System x3850 X5, x3950 X5* at the following website for the installation instructions:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5083918

Go to the following website for information about installing SUSE Linux Enterprise Server 11 SP1 on an UEFI-aware system:

http://www.novell.com/support/documentLink.do?externalID=7003263

We highly recommend that you do *not* attach SAN storage to the HX5 during the installation process, because this action can lead to accidental data loss at the disk partitioning phase.

You can obtain additional installation and configuration information about SUSE Linux Enterprise Server at the following websites:

- *SUSE Linux Enterprise Server 10*

  http://www.novell.com/documentation/sles10/

- *SUSE Linux Enterprise Server 11*

  http://www.novell.com/documentation/sles11/

You can download SUSE Linux Enterprise Server evaluation software from this website:

http://www.novell.com/linux/download_linux.html

Troubleshooting information for HX5 and SUSE Linux Enterprise Server

We recommend that you review all Retain tips relating to the HX5 and the operating system that you are installing.

Go to the following website for troubleshooting information pertaining to HX5:

http://ibm.com/support/entry/portal/Troubleshooting/Hardware/Systems/BladeCenter/BladeCenter_HX5

SUSE Linux Enterprise Server software, firmware, drivers, and fixes for HX5

Go to the following websites for all the latest information relating to software, firmware, drivers, and fixes for HX5 when used with SUSE Linux Enterprise Server:

- For SUSE Linux Enterprise Server 11:

  http://www-933.ibm.com/support/fixcentral/systemx/quickorder?parent=ibm~BladeCenterHX5&product=ibm/systemx/7872&&platform=SLES+11+x64&function=all&source=fc

- SUSE Linux Enterprise Server 10:

  http://www-933.ibm.com/support/fixcentral/systemx/quickorder?parent=ibm~BladeCenterHX5&product=ibm/systemx/7872&&platform=SLES+10+x64&function=all&source=fc

8.7.7 Downloads and fixes for HX5 and MAX5

Typically, during the support lifetime of a product, IBM releases updates to provide you with enhanced capabilities, extended functions, and problem resolutions. Most of the updates are in the form of firmware, drivers, and operating system patches.

We recommend that you perform a scheduled review of the available updates to determine if they apply to the systems that are used in your environment.

Server firmware

Software that resides on flash memory and controls the lower-level function of server hardware is called the server *firmware*. An IBM System, such as the HX5, runs a number of firmware images that control various components of the blade.

The following list shows the primary firmware for the HX5:

- *Unified Extensible Firmware Interface (UEFI)*
- *Integrated Management Module (IMM)*
- *Field-Programmable Gate Array (FPGA)*
- *Preboot Dynamic System Analysis (DSA)*

Additional devices, such as network cards and RAID controllers, also contain their own firmware revisions. IBM provides firmware updates, including proven firmware from other manufacturers to be applied on IBM systems, that can be downloaded from the IBM website.

We describe several methods of performing firmware updates on IBM eX5 servers in Chapter 9, “Management” on page 447.

Important tip: Update all System x firmware to the latest levels prior to installing the operating system or application installation.

Tip: *IBM Bootable Media Creator (BoMC)* is a tool that simplifies the IBM System x firmware update process without an operating system running on the system. More information about this tool is available at the following website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-BOMC

Device drivers

Device drivers are software that controls the hardware server components on the operating system level. They are specific to the operating system version, and therefore, critical device drivers are included with the installation media.

IBM, operating system vendors, and component device vendors provide device driver updates. Often, you can download them from each company's support website. Whenever possible, we recommend acquiring tested and approved driver updates from IBM.

Tip: The Windows Server installation process using IBM ServerGuide will perform IBM driver updates after the operating system installation has completed. You can obtain this tool from the following website:

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

Tip: *IBM UpdateXpress* is a tool that allows the IBM System x firmware and drivers to be updated through the operating system. More information about this tool is available at the following website:

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

Operating system updates, fixes, and patches

The performance and reliability of an HX5 tightly relate to the operating system running on the HX5. IBM supports an assortment of modern and widely used operating systems that are capable of utilizing the system's potential. Each vendor supports its operating system by releasing updates, fixes, and patches that provide enhanced functionality and fixes to known problems. Many of these updates, fixes, and patches only apply to certain configurations, while other updates, fixes, and patches apply to all configurations. The operating system vendor's support website has extensive information about these updates, fixes, and patches.

System update resources

Table 8-7 provides useful links to IBM tools and vendor operating system support websites.

Table 8-7 Internet links to support and downloads

Vendor	Product	Address	Available support
IBM	Systems Support	http://ibm.com/systems/support/	Documentation, firmware, driver
IBM	ServerGuide	http://ibm.com/support/entry/portal/docdisplay?lndocid=SERV-GUIDE	Installation tool
IBM	UpdateXpress	http://ibm.com/support/entry/portal/docdisplay?lndocid=SERV-XPRESS	Firmware and driver update tool
IBM	Bootable Media Creator	http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-BOMC	Firmware update tool
Microsoft	Windows Server	http://support.microsoft.com/ph/14134	Documentation, driver, and OS update

Vendor	Product	Address	Available support
Red Hat	RHEL	https://www.redhat.com/support/	Documentation, driver, and OS update
Novell	SLES	http://www.novell.com/support/	Documentation, driver, and OS update
VMware	vSphere	http://downloads.vmware.com/d/	Documentation, driver, and OS update

8.7.8 SAN storage reference and considerations

BladeCenters are designed to provide blades with usable storage via the external storage attachment due to the blade's compact design. We highly recommend that you have an understanding of the terms and methods that relate to SAN attachment. SAN is out of the scope of this IBM Redbooks publication; however, we provide information related to the protocols that are used to attach to SAN and boot from SAN.

SAN storage attachment terms

The primary purpose of the storage area network (SAN) is to transfer data between computer systems and storage elements. The following list describes each of the most common SAN protocols and their characteristics:

- Fibre Channel (FC)

  The Fibre Channel Protocol (FCP) is the interface protocol of SCSI on FC. FCP is a transport protocol that predominantly transports SCSI commands over FC networks. FC is the prevalent technology standard in the SAN data center environment. Typical requirements for this configuration are an FC host bus adapter (HBA) and FC SAN infrastructure. Despite its name, FC signaling can run on both twisted-pair copper wire and fiber optic cables.

- Fibre Channel over Ethernet (FCoE)

  FCoE is the transport, or mapping, of encapsulated FC frames over the Ethernet. The Ethernet provides the physical interface, and FC provides the transport protocol. The system setup for FCoE requires the Converged Network Adapter (CNA) to pass both network and storage data that is connected to a 10 Gb converged network infrastructure.

- Internet SCSI (iSCSI)

  iSCSI is an Internet Protocol (IP)-based storage networking standard for linking data storage facilities. The protocol allows clients (called *initiators*) to send SCSI commands to SCSI storage devices (targets). A hardware initiator might improve the performance of the server. Often, iSCSI is seen as a low-cost alternative to FC.

- Serial-attached SCSI (SAS)

  SAS uses point-to-point connection. The typical SAS throughput is 6 Gbps full duplex. If a complex SAN configuration is not necessary, SAS is a good choice although performance and distance are limited compared to the other solutions.

Booting from SAN

This section provides useful guidelines about booting from SAN:

- Check to make sure that UEFI recognizes the HBA. Select **UEFI** → **System Settings** → **Adapter and UEFI Drivers**. When you press Enter, the respective HBA is visible, as shown in Figure 8-71. If it is not visible, you might need to reflash the UEFI, IMM, and firmware of the HBA and check again.

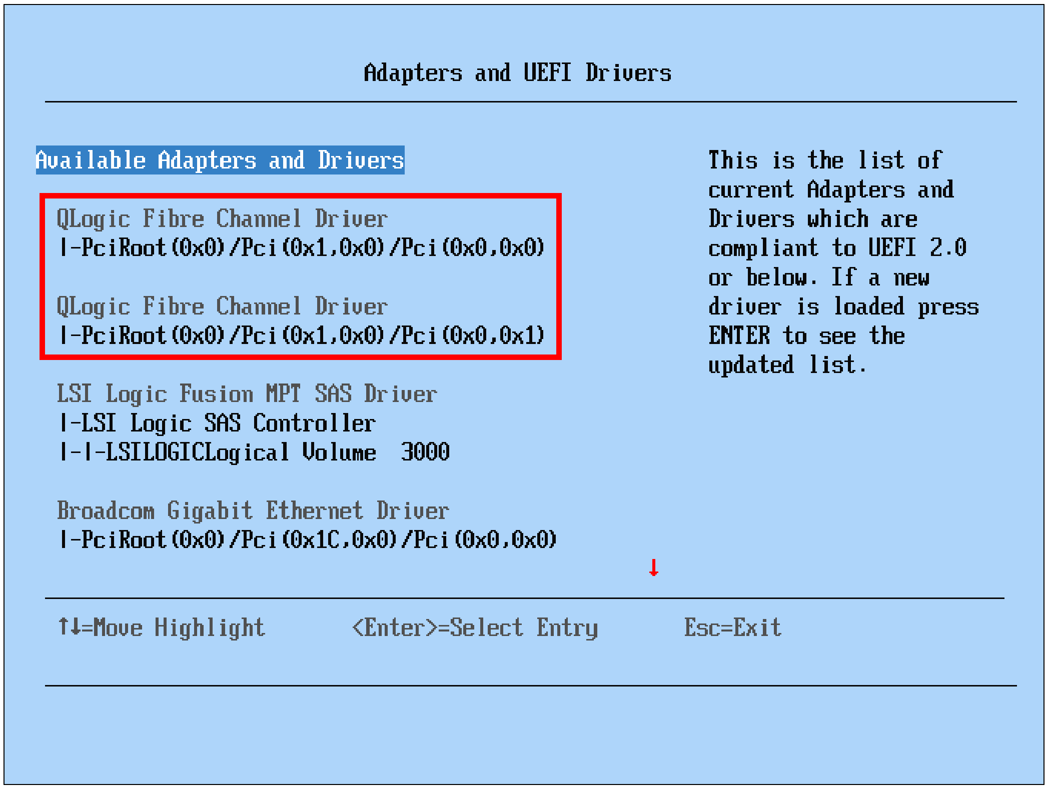


Figure 8-71 Adapters visible in UEFI

- If you do not have internal drives installed, disable the onboard SAS RAID Controller in UEFI by navigating to **System Settings** → **Devices and IO ports** → **Enable/Disable Onboard Devices** and disabling the SAS Controller or Planar SAS.
- Set the HBA as the first device in the Option ROM Execution Order by selecting **System Settings** → **Devices and IO Ports** → **Set Option ROM Execution Order.**
- All operating systems, except Windows 2008 and SLES 11 SP1, now have **Legacy Only** set as their first boot device.
- Remove all devices that might not host an operating system from the boot order. The optimal minimum configuration is CD/DVD and Hard Disk 0. For existing operating systems only, set **Legacy Only** as the first boot device.
- You must set the BIOS on the HBA to **Enabled**.
- Make sure that the logical unit number (LUN) that will host the operating system installation is accessible through only one path on the SAN at the time of the installation.
- Verify that your HBA can see a LUN from your storage.
- After installation, do not forget to install the multipath driver *before* you set more than one path if you have more than one path to the LUN.

IBM Redbooks references for SAN-related information

The following IBM Redbooks publications describe IBM System Storage products and their various implementations, including with the IBM System x product lines:

- *IBM System Storage Solutions Handbook*, SG24-5250

  This book provides overviews and pointers for information about the current IBM System Storage products.

  http://www.redbooks.ibm.com/abstracts/sg245250.html

- *Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*, SG24-6116

  This book consolidates critical information while also covering procedures and tasks that you are likely to encounter on a daily basis when implementing an IBM/Brocade SAN.

  http://www.redbooks.ibm.com/abstracts/sg246116.html

- *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363

  This book represents a compilation of best practices for deploying and configuring IBM Midrange System Storage servers, which include the DS4000 and DS5000 family of products.

  http://www.redbooks.ibm.com/abstracts/sg246363.html

- *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065

  This book introduces the IBM System Storage DS3000, providing an overview of its design and specifications and describing in detail how to set up, configure, and administer it.

  http://www.redbooks.ibm.com/abstracts/sg247065.html

- *Implementing an IBM/Cisco SAN*, SG24-7545

  This book consolidates critical information while describing procedures and tasks that are likely to be encountered on a daily basis when implementing and IBM/Cisco SAN.

  http://www.redbooks.ibm.com/abstracts/sg247545.html

- *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659

  This book describes the concepts, architecture, and implementation of the IBM XIV Storage System.

  http://www.redbooks.ibm.com/abstracts/sg247659.html

- *IBM Midrange System Storage Hardware Guide*, SG24-7676

  This book consolidates, in one document, detailed descriptions of the hardware configurations and options that are offered as part of the IBM Midrange System Storage servers, which include the IBM System Storage DS4000 and DS5000 families of products.

  http://www.redbooks.ibm.com/abstracts/sg247676.html

For more information regarding HBA storage-specific settings and zoning, contact your SAN vendor or storage vendor.

8.8 Failure detection and recovery

In this section, we provide an overview of the available tools to assist you with problem resolution for the HX5 in any given configuration. We also provide considerations with regard to blade placement and extended outages.

8.8.1 Tools to aid hardware troubleshooting for the HX5

Use the following tools when troubleshooting problems on the HX5 in any given configuration.

Advanced Management Module

The first place to start troubleshooting the HX5 is typically the Advanced Management Module (AMM). The AMM is extremely intuitive and, in general, quickly provides the required information to resolve most of the problems that are experienced with the HX5 in any configuration. At a high level, the AMM provides the following features for problem diagnosis and alerting:

- The AMM provides a quick health view of all systems on the System Summary page to determine a faulty system from a remote location quickly.
- It centralizes hardware alerts from all blades into its event log and provides filtered search features to diagnose problems quickly and easily.
- You can configure the AMM to send alerts via the following methods:
  - Simple Network Management Protocol (SNMP) traps
  - Email alerts
  - Call home to IBM support via the built-in Service Advisor
  - Integration to IBM Systems Director for centralized alerting

For the configuration of SNMP traps and email alerts, see the *IBM BladeCenter Advanced Management Module Installation Guide,* which is available at the following website:

http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5073392

For the configuration of Service Advisor, see to 9.3.2, “Service Advisor” on page 458.

For information about managing and centralizing alerts for managed systems using the IBM Systems Director, see *Implementing IBM Systems Director 6.1*, SG24-7694.

The Problem Determination and Service Guide - IBM BladeCenter HX5

You can solve many problems without assistance from IBM by following the troubleshooting procedures in the *Problem Determination and Service Guide - IBM BladeCenter HX5*, which is available at the following website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5084529

The *Problem Determination and Service Guide - IBM BladeCenter HX5* describes the diagnostic tests that you can perform, troubleshooting procedures, and explanations of error messages and error codes.

If you have completed the diagnostic procedure and the problem remains and verified that all firmware code is at the latest level and all hardware and software configurations are valid, contact IBM or an approved warranty service provider for assistance.

Light path diagnostics

The light path diagnostics provide visual assistance to locate faulty hardware on the HX5:

- When the HX5 is removed from the chassis and the cover is open, pressing and holding the power button illuminates the lights of any faulty components.
- When the MAX5 is removed from the chassis and the cover is opened, you can use the light path diagnostics panel that is clearly located on the MAX5 system board to locate faulty hardware by pressing and holding the light path diagnostics button.

System event log

This log contains POST and system management interrupt (SMI) events and all events that are generated by the Baseboard Management Controller (BMC) that is embedded in the IMM. You can view the system event log through the UEFI utility by pressing F1 at system start-up and selecting **System Event Logs** → **System Event Log**. Also, you can view the system event log through the Dynamic System Analysis (DSA) program (as the IPMI event log).

POST event log

This log contains the three most recent error codes and messages that were generated during POST. You can view the POST event log through UEFI by pressing F1 at system start-up and navigating to **System Event Logs** → **POST Event Viewer**.

8.8.2 Reinserting the Speed Burst card for extended outages

In the unlikely scenario that a 4-socket HX5 configuration might be reduced to a 2-socket configuration due to a hardware failure, you can remove the IBM HX5 2-node scalability kit and connect the IBM HX5 1-node Speed Burst kit to increase the processor performance. This step is not mandatory for the remaining system to continue to operate. However, connect the IBM HX5 1-node Speed Burst kit can boost CPU performance temporarily for CPU-bound applications by doubling the QPI links between the CPUs on the remaining HX5.

This same solution also applies if you have HX5 with MAX5 attached and the MAX5 fails completely.

MAX5: In the unlikely event that a MAX5 fails completely, you must detach the MAX5 from the HX5 by removing the IBM HX5 MAX5 1-Node Scalability kit to allow the HX5 to resume normal operations.

8.8.3 Effects of power loss on HX5 2-node or MAX5 configurations

BladeCenter (BC) H by design is split into two power domains to provide redundancy to the chassis. From a blade bay perspective, power domain 1 supplies power to blades 1 - 7 and power domain 2 supplies power to blades 8 - 14.

If you install the primary blade server of a scalable blade complex in blade server bay 7 of a BC H Type 8852 chassis, the secondary blade or MAX5 is installed in bay 8. This configuration causes the blades to be split between two separate power domains. The following situations can occur if there is a power loss to either power domain, depending on how the scalable blade complex is implemented.

You might see the following effects of split power domains for HX5 4-socket configurations:

- A loss of power to power domain 1 results in both blade servers of a scalable partition going down. The remaining primary node in bay 7 that still has power to it might be powered on as a single complex system. You might receive errors on reboot stating that a non-maskable interrupt (NMI) error occurred or that there are scalability problems. This behavior is expected due to losing the power to the processors abruptly and the second node being still physically connected via the scalability card but unavailable.
- If the scalable blade complex is implemented in stand-alone mode, a loss of power to power domain 2 results in only the secondary blade server in bay 8 going down. The primary blade remains powered on but still generates an error.

You might see the following effects of split power domains for HX5 with MAX5 configurations:

- A loss of power to either power domain results in complete system shutdown.
- The HX5 will not be able to power on if the attached MAX5 in slot 8 has no power, because the HX5 treats the entire system as a single unit. You must remove the IBM HX5 MAX5 1-Node Scalability kit to allow the HX5 to resume normal operation if power cannot be restored to the second power domain that contains the MAX5.

Where possible, avoid placing HX5 4-socket or HX5 and MAX5 configurations across blade bays 7 and 8.

BladeCenter design: The loss of power to an entire power domain is extremely unlikely due to the redundant design of the BladeCenter H chassis. We provide this information merely for guidance in the case of this unlikely event.

9

Management

In the information technology sector, server systems management has received greater focus over recent years. The ability to maintain and manage systems efficiently is essential to business IT operations. We briefly describe the embedded hardware and external applications that are available to manage and maintain the eX5 range of systems. And, we demonstrate several of them.

This chapter contains the following topics:

9.1 Introduction

IBM provides a number of tools to successfully deploy, manage, and maintain the eX5 range of systems. The collective name for these tools is IBM ToolsCenter. You can access these tools at the following link:

http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-CENTER

We group these tools in the following high-level categories:

- System deployment
- System configuration
- System updates
- System diagnosis
- System management

We provide a summary of these tools in Table 9-1. In this chapter, we walk you through the process of using several of these tools.

Table 9-1 High-level overview of the available tools to manage the eX5 range of systems

IBM tools	Covered here	Firmware deployment	OS installation assistance	System management	Problem diagnosis
Integrated Management Module	p. 449	Yes[a]	No	Yes	Yes
Advanced Management Module	p. 454	Yes	No	Yes	Yes
Dynamic System Analysis	No	No	No	No	Yes
Bootable Media Creator	p. 514	Yes	No	No	No
Scripting Toolkit	p. 507	Yes	Yes	No	No
IBM ServerGuide	p. 501	No	Yes	No	No
Start Now Advisor	No	Yes	No	No	No
Update Express System Pack Installer	p. 511	Yes	No	No	No
IBM Systems Director	p. 467	Yes	Yes[b]	Yes	Yes
IBM Electronic Service Agent	p. 493	No	No	No	Yes
Storage Configuration Manager	No	Yes	No	Yes	Yes
Remote Control	p. 462	No[c]	No[c]	Yes	No
Advanced Settings Utility	p. 495	No[c]	No[c]	Yes	No
MegaRAID Storage Manager	p. 521	No	No	Yes[d]	Yes
Serial over LAN	p. 525	No	No	Yes	No

a. You can only update Integrated Management Module (IMM) firmware using the IMM web browser interface.
b. Only when the Tivoli Provision Manager for OS Deployment IBM Systems Director edition is installed.
c. This tool provides the ability to mount the media containing the firmware or operating system.
d. This tool provides the management of RAID controllers and hard disks only.

9.2 Integrated Management Module (IMM)

The Integrated Management Module (IMM) offers these overall features and functions:

- Provides diagnostics, virtual presence, and remote control to manage, monitor, troubleshoot, and repair from anywhere
- Securely manages servers remotely, independently of the operating system state
- Helps remotely configure and deploy a server from bare metal
- Auto-discovers the scalable components, ports, and topology
- Provides one IMM firmware for a new generation of servers
- Helps system administrators easily manage large groups of diverse systems
- Requires no special IBM drivers
- Works with IBM Systems Director to provide secure alerts and status, helping to reduce unplanned outages
- Uses standards-based alerting, which enables upward integration into a wide variety of enterprise management systems

In general, there are two methods to manage an IMM:

- Out-of-band management: All management tasks are passed directly to the system's IMM via a network connection. No drivers are required for the IMM, because it is configured with its own IP address and is connected directly to the network.
- In-band management: All management tasks are passed to the system using the operating system installed on it. The tasks can apply to the operating system, or they can apply to the IMM that is installed on the system. If the tasks are to be passed to the IMM, the relevant operating system driver must be installed for the IMM.

In the following section, we look at both the out-of-band and in-band initial configuration.

9.2.1 IMM out-of-band configuration

The x3690 X5 and the x3850 X5 vary slightly from the HX5 in the way that their IMMs are managed out-of-band. You can configure the IMM to send hardware alerts, for example, directly to a management system. Although the HX5 has its own IMM, there is no direct external network access to it and the Advanced Management Module (AMM) in the BladeCenter chassis must be used to manage the blade.

Configuring an x3850 X5 or x3690 X5 for out-of-band management

The IMM for both the x3690 X5 and the x3850 X5 can be managed over one of two ports:

- The dedicated system management port on the rear of each chassis.

  This port allows the IMM to be connected to an isolated management network for improved security.
- The onboard Ethernet port 1 on the rear of each chassis. The IMM shares this port with the operating system.

  This option is provided to allow the IMM to be managed out-of-band without using additional ports on an external network switch.

Enabling Ethernet port 1: The onboard Ethernet port 1 is not shared for use with the IMM, by default. You must enable this feature in the Unified Extensible Firmware Interface (UEFI) of the server in both the x3690 X5 and the x3850 X5. The following sections contain the configuration instructions.

Figure 9-1 shows the location of the available ports to manage the IMM on the x3690 X5.

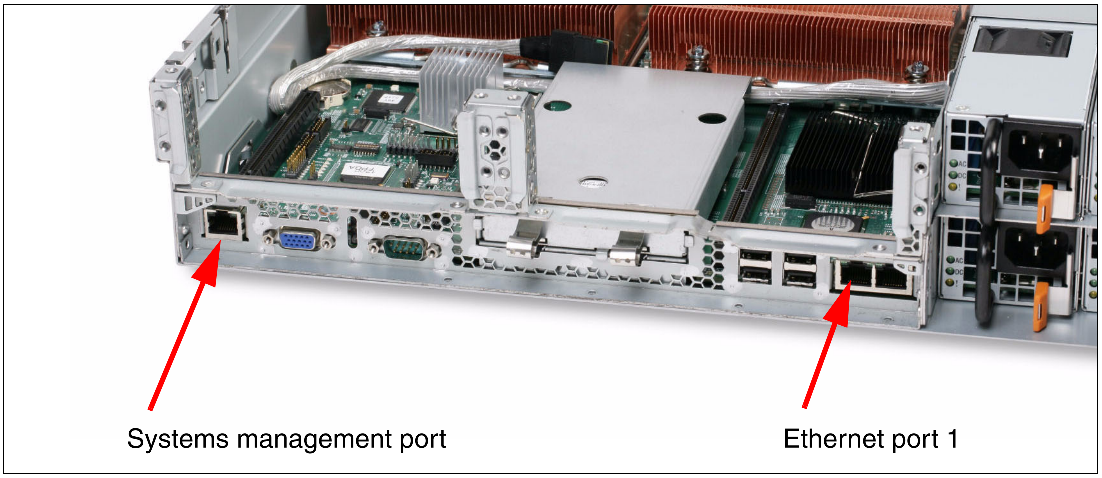


Figure 9-1 Ports over which the IMM can be managed out-of-band on an x3690 X5

Figure 9-2 shows the location of the ports available to manage the IMM on the x3850 X5.

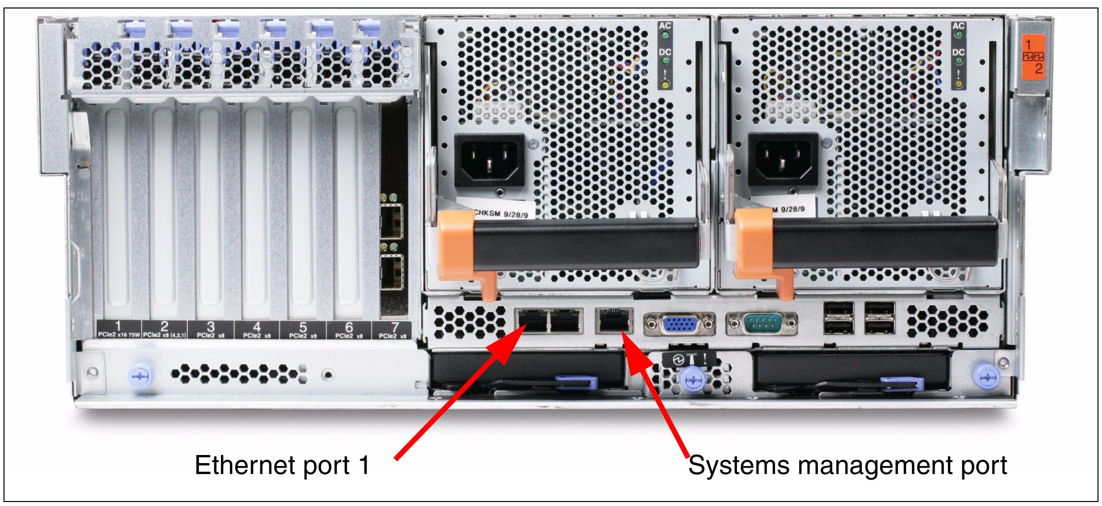


Figure 9-2 Ports over which the IMM can be managed out-of-band on an x3850 x5

To enable the IMM to use Ethernet port 1, you must complete the following instructions:

1. Boot up the server and press F1 when prompted.
2. Select **System Settings** from the System Configuration and Boot Management menu.
3. Select **Integrated Management Module** from the System Settings menu.
4. Select **Network Configuration** from the Integrated Management Module. This menu allows you to configure the network settings for the IMM. Also, you can configure the IMM to share the use of Ethernet port 1 for out-of-band management from this menu, as shown in Figure 9-3 on page 451.

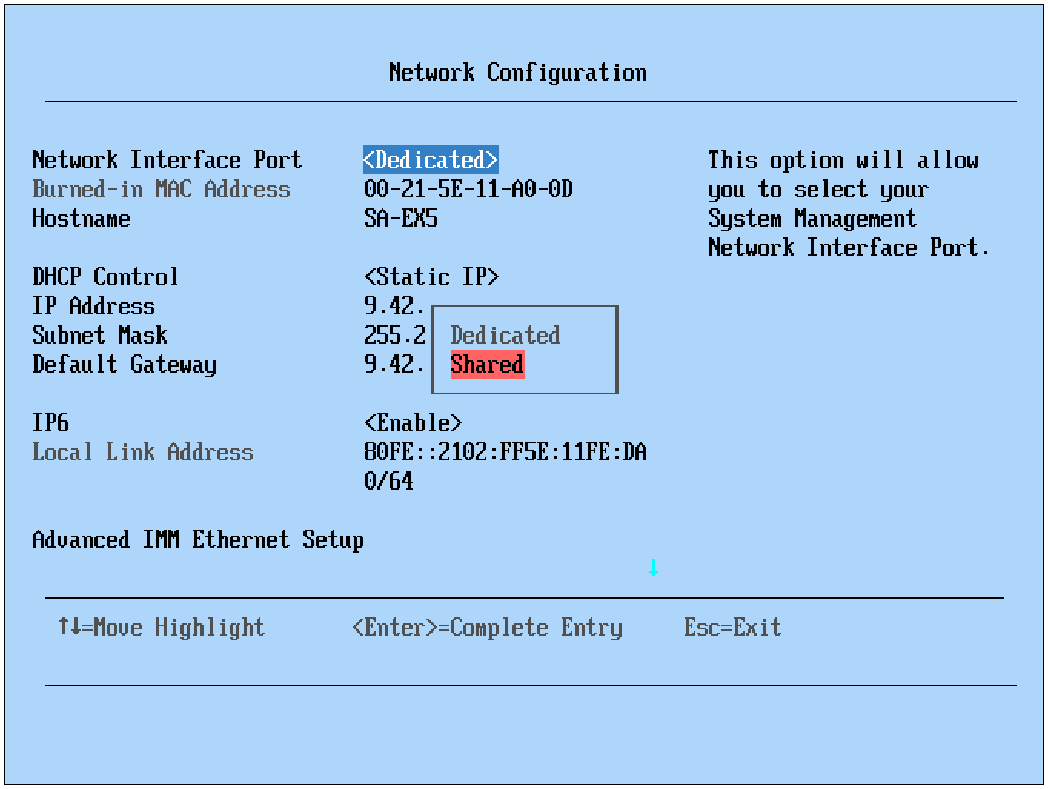


Figure 9-3 Network Configuration menu showing the Network Interface Port settings

5. Set the Network Interface Port setting to **Shared** to allow the IMM to use Ethernet port 1.
6. For DHCP Control, choose the **Static IP** option.
7. For IP Address, enter the relevant IP address.
8. For Subnet Mask, enter the required subnet mask.
9. For Default Gateway, enter the required default gateway address.
10. When you have completed the IP address configuration, press Esc three times to return to the System Configuration and Boot Management menu.
11. For Exit Setup, press the Y key when prompted to save and exit the Setup utility. At this point, the server reboots with the new settings.
12. Plug a network cable into either the dedicated system management port or Ethernet port 1, if you set the IMM to share its use as per instructions. Make sure that you can ping the IP address of the IMM on the connected network port.

After the IMM is available on the network, you can log in into the IMM web interface by typing its IP address in a supported web browser, as shown in Figure 9-4 on page 452.

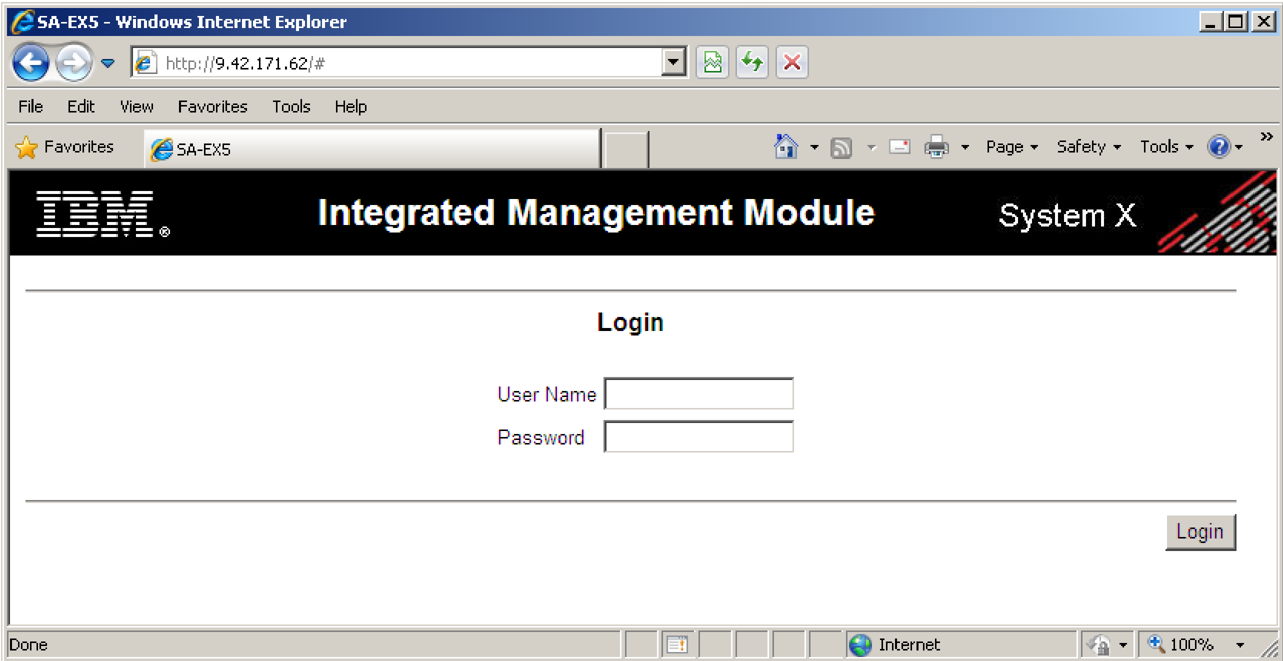


Figure 9-4 IMM Login web page

Enter the default user name, which is `USERID`. This user name is case-sensitive. Enter the default password, which is `PASSW0RD`. You must use a zero in place of the letter "o" in the password.

See the *User's Guide for Integrated Management Module - IBM BladeCenter and System x* at the following website for the additional configuration settings of the IMM:

`http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5079770`

Accessing an HX5 blade for out-of-band management

The HX5 has an IMM on board but the IMM does not have direct access to the external network. All management tasks directed to the IMM are passed using the AMM. As long as the AMM is connected to the network, the HX5 can be managed out-of-band via the AMM.

To manage an HX5, log in to the AMM by typing the IP address of the AMM into a supported web browser, as shown in Figure 9-5 on page 453.

Enter the default user name, which is `USERID`, for the AMM. This user name is case-sensitive. Enter the default password, which is `PASSW0RD`. You must use a zero in place of the letter "o" in the password.

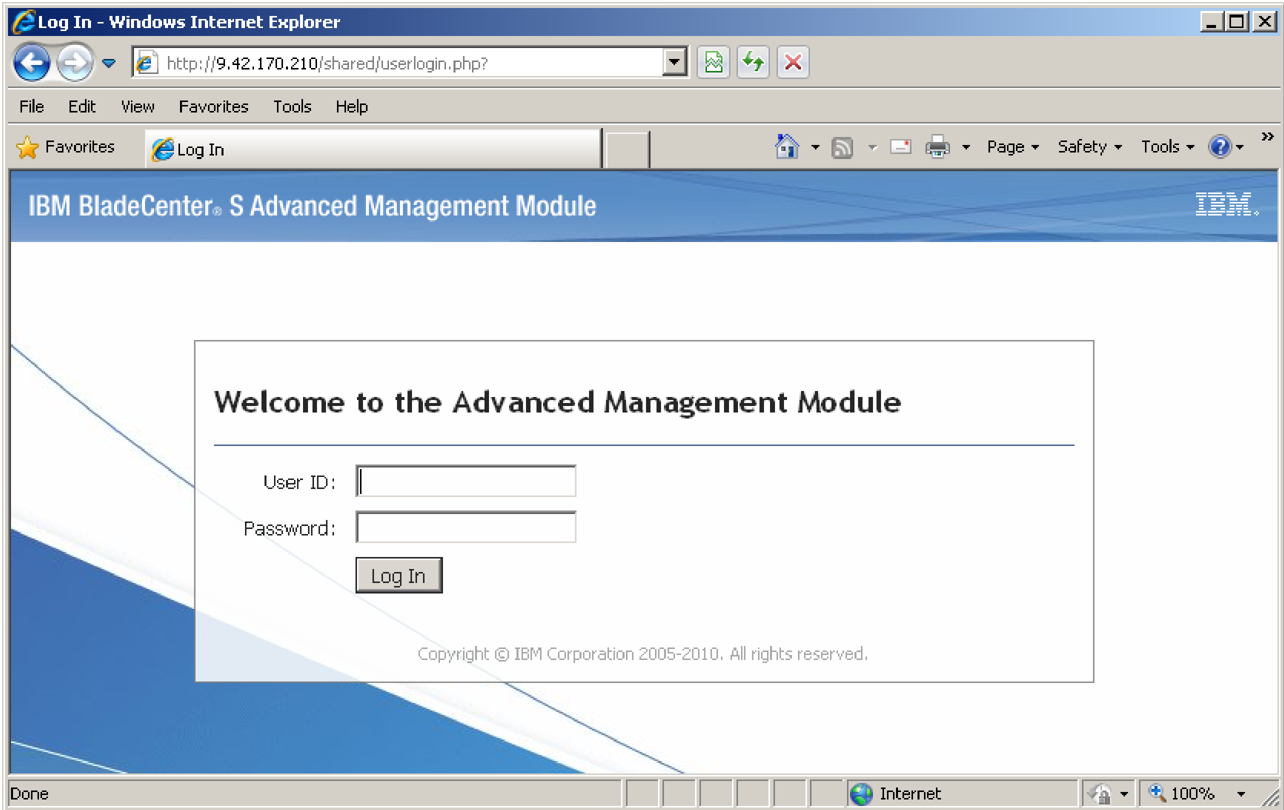


Figure 9-5 AMM login page

After you log in to the AMM, you can manage and retrieve information from the HX5, such as hardware vital product data (VPD), power usage, and so on. We describe more of this information in 9.3, "Advanced Management Module (AMM)" on page 454.

9.2.2 IMM in-band configuration

Managing an IMM in-band means to manage the IMM through the operating system. IBM Systems Director, for example, can update the firmware of an IMM via the operating system. The benefit of this approach is that you do not need to configure the IMM with its own dedicated IP address if there are insufficient IP addresses to allocate.

For in-band management, the methods of managing the x3690 X5 and the x3850 X5 are the same as the HX5.

There is no actual configuration required within the IMM web interface to allow the IMM to be managed in-band. However, you must ensure that the prerequisite drivers are installed to allow the operating system to recognize the IMM. All supported versions of Microsoft Windows Server 2008, VMware ESX, and Linux now include the prerequisite drivers for the eX5 systems.

See IBM ServerProven at the following website for a list of the supported operating systems:

http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/matrix.shtml

If IBM Systems Director is used to manage an eX5 system in-band, you must install an agent on the operating system of the managed system. We provide more detail about IBM Systems Director in 9.5, "IBM Systems Director 6.2" on page 467.

9.2.3 Updating firmware using the IMM

Through the Integrated Management Module (IMM), you can apply the following updates:

- IMM
- UEFI
- Field Programmable Gate Array (FPGA)
- Preboot Dynamic System Analysis (pDAS)

You can use the `.exe` or `.sh` file to make an update. For the UEFI update, you must power up and boot the server into the operating system or in the UEFI Menu.

Update the firmware in this order:

1. IMM
2. UEFI
3. FPGA
4. pDSA

> **IMM firmware update:** After the IMM firmware update, the IMM will be reset. The restart of the IMM after the update can take up to 9 minutes (13 minutes on a scalable complex system). After this restart, the IMM is ready for any further firmware updates.

Use the IMM to update firmware:

1. Log in into the IMM web interface.
2. Select **Tasks** → **Firmware Update**, as shown in Figure 9-6.

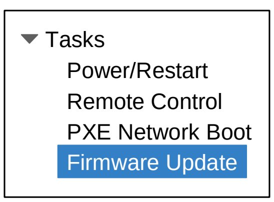


Figure 9-6 IMM menu

3. Click **Browse** and select the update file.
4. Click **Update** to start the update process. A progress indicator opens as the file is transferred to the temporary storage of the IMM.

When the transfer is completed, click **Continue** to complete the update process. A progress indicator opens as the firmware is flashed. A confirmation page opens to verify that the update was successful.

9.3 Advanced Management Module (AMM)

The Advanced Management Module (AMM) is a hot-swap module that you use to configure and manage all installed BladeCenter components. The AMM provides system management functions and keyboard/video/mouse (KVM) multiplexing for all blade servers in the BladeCenter unit that support KVM. It controls a serial port for remote connection, the external keyboard, mouse, and video connections for use by a local console, and a 10/100 Mbps Ethernet remote management connection.

Figure 9-7 shows an AMM.

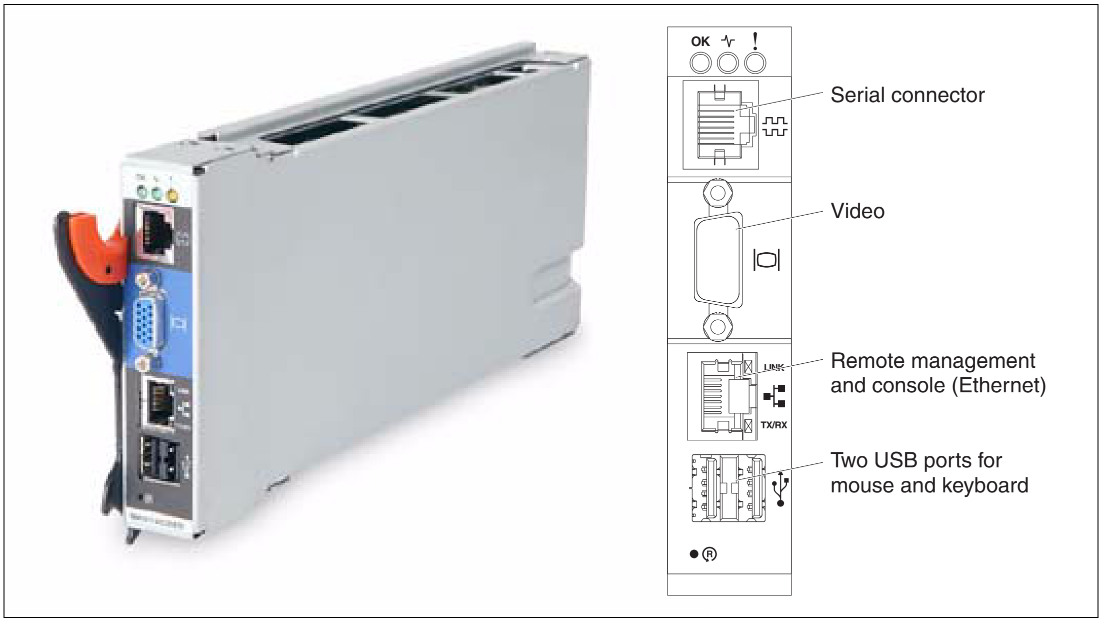


Figure 9-7 AMM

All BladeCenter chassis come standard with at least one AMM. Each chassis also supports a second management module for redundancy. One of the management modules is active, and the second management module, if installed, remains on standby until the management functions are manually switched over to it or if the primary management module fails.

The service processor in the management module communicates with the service processor in each blade server. The IMM is the service processor on the HX5. The service processor allows support features, such as blade server power-on requests, error and event reporting, KVM requests, and requests to use the BladeCenter shared media tray (removable-media drives and USB connector).

You configure BladeCenter components by using the management module to set information, such as IP addresses. The management module communicates with all components in the BladeCenter unit, detecting their presence or absence, reporting their status, and sending alerts for error conditions when required.

With the AMM, you can perform the following tasks:

- Define the login IDs and passwords
- Configure the security settings, such as data encryption and user account security
- Select recipients for the alert notification of specific events
- Monitor the status of the BladeCenter unit, blade servers, and other BladeCenter components:
  - Event log
  - LEDs
  - Hardware and firmware VPD
  - Fan speeds
  - Temperatures
  - Power usage

- Discover other BladeCenter units on the network and enable access to them through their management-module Web interfaces
- Control the BladeCenter unit, blade servers, and other BladeCenter components:
  - Power on/off
  - Firmware update
  - Configuration settings
  - Serial over LAN
- Configure power management for BladeCenter unit
- Access the I/O modules to configure them
- Change the start-up sequence in a blade server
- Set the date and time
- Use a remote console for the blade servers
- Mount remote virtual media for the blade servers
- Change ownership of the keyboard, video, and mouse
- Change ownership of the removable-media drives and USB ports (the removable-media drives in the BladeCenter unit are viewed as USB devices by the blade server operating system)
- Set the active color of the critical (CRT) and major (MJR) alarm LEDs (for BladeCenter T units only)
- Use BladeCenter Open Fabric Manager functions
- Scale the HX5 systems
- Use Service Advisor functions to autonomously inform IBM support about any critical events happening

The AMM supports the following management methods:

- Web-based interface with Secure Sockets Layer (SSL) support
- Command-line interface (CLI) through Telnet/Secure Shell (SSH)
- Systems Management Architecture for Server Hardware (SMASH) Command-Line Protocol
- Simple Network Management Protocol (SNMP)

9.3.1 Accessing the Advanced Management Module

Log in to the AMM by typing the IP address of the AMM into a supported web browser, as shown in Figure 9-8 on page 457.

The default user name is for an AMM is `USERID`. This user name is case-sensitive.

The default password is `PASSW0RD`. You must use a zero in place of the letter “o” in the password.

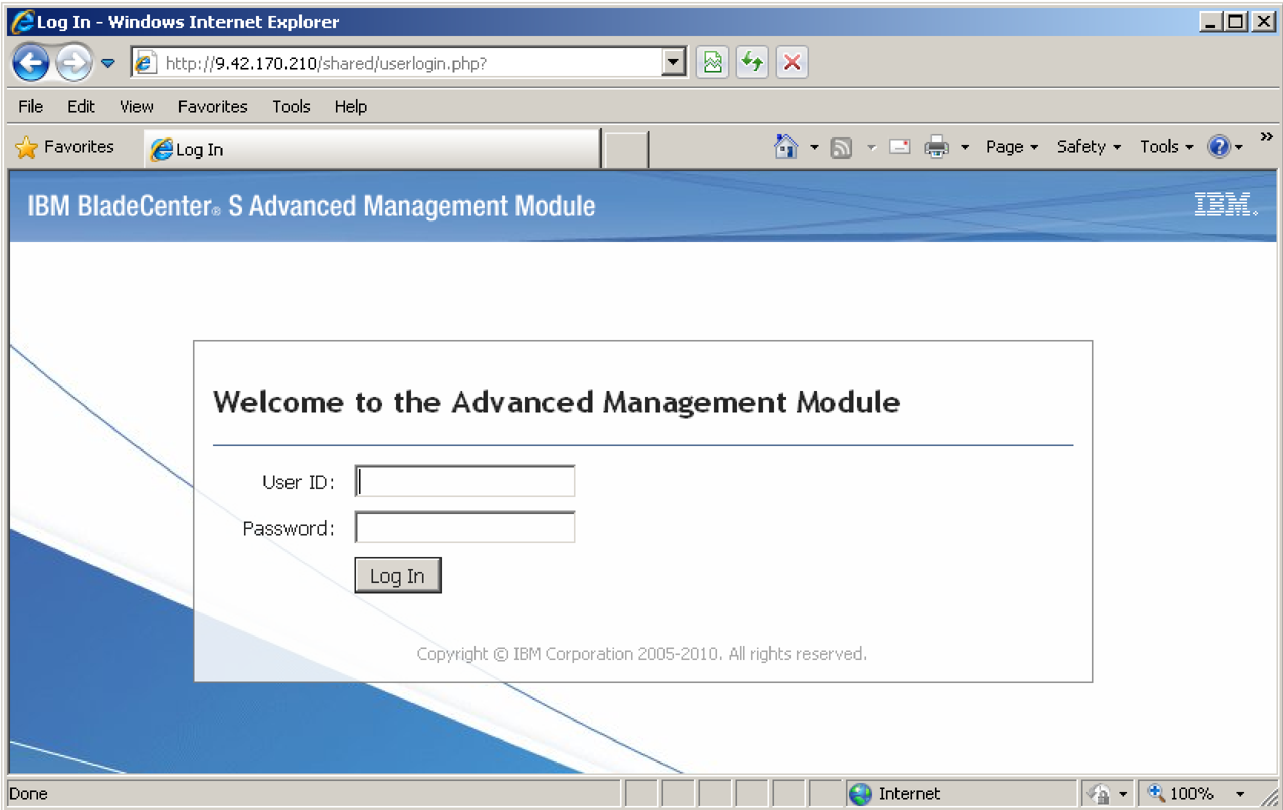


Figure 9-8 Advanced Management Module login web page

The System Status summary page displays. This page provides you with information about the overall health of the chassis and its installed components, as shown in Figure 9-9 on page 458.

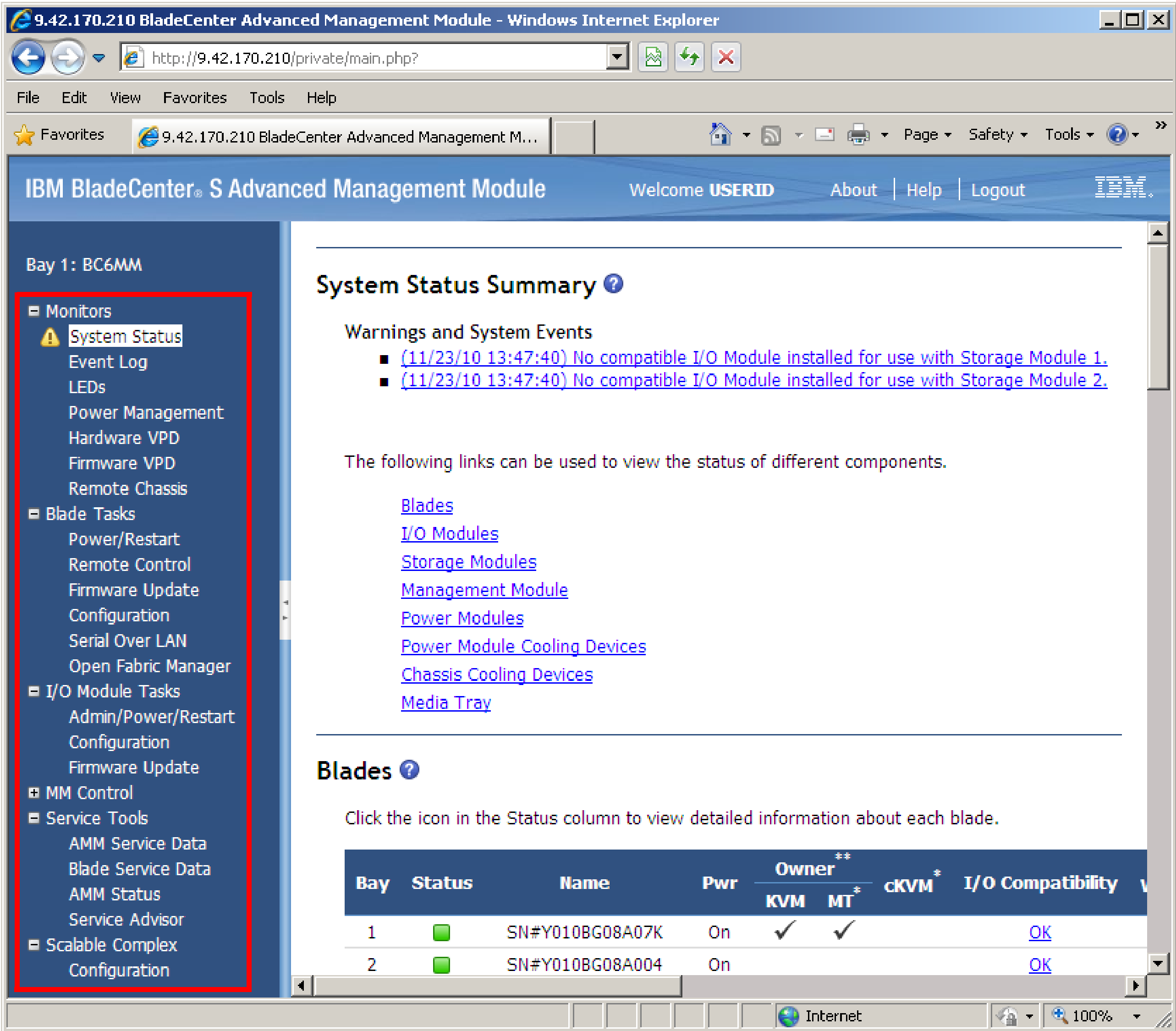


Figure 9-9 System Status Summary page

All high-level tasks that can be performed against the various components within the BladeCenter chassis are shown in the left pane, as highlighted in Figure 9-9. Of particular importance to this IBM Redbooks publication is the Scalable Complex section. In the Scalable Complex section, you can configure the HX5 blades to form 2-node scalable systems. We describe the scaling configuration of the HX5s in 8.6, “Creating an HX5 scalable complex” on page 402.

For more information about the AMM and its configuration, see the *IBM BladeCenter Advanced Management Module Installation Guide* at the following website:

`http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5073392`

9.3.2 Service Advisor

The AMM contains a component called the Service Advisor that allows the chassis to automatically notify IBM service and support of select hardware issues. When a serviceable event that has been designated as a call home event is detected, a message is written in the event log and any configured alerts are sent. The information that is gathered by Service Advisor is the same information that is available if you save service data from the AMM web interface.

After gathering the information, Service Advisor automatically initiates a call to IBM. Upon receipt of the information, IBM returns a service request ID, which is placed in the call home activity log.

> **Tip:** Although Service Advisor can send alerts 24 hours a day, 7 days a week, your service provider responds according to the arrangement that you have in place with the service provider.

On the Event Log page of the AMM web interface, you can choose to select the Display Call Home Flag check box. If you select the check box, events are marked with a `C` for call home events and an `N` for events that are not called home. In addition, you can filter the event log based on this setting.

> **Tip:** None of the information in the Call Home report contains client data from the servers or the I/O modules.

Before you configure the BladeCenter Service Advisor, you must ensure that ports 80 and 443 are open for the AMM to call home.

Complete the steps that follow to successfully configure the Service Advisor:

1. Log in to the AMM on which you want to activate the Service Advisor.
2. In the left navigation pane, click **Service Tools** → **Service Advisor.** If this is the first time that you select this option, or if the AMM firmware was reset to the default values, you need to view and accept the license agreement.
3. Click **View terms and conditions** to view the Service Advisor agreement. Click **I accept the agreement** on the terms and conditions page to close the page.
4. Click the **Service Advisor Settings** tab, as shown in Figure 9-10 on page 460, and complete all relevant details. Ensure that you select the correct IBM support center for your geographical location. Also, note that the FTP/TFTP Server of Service Data must only be configured if an approved service provider other than IBM provides your hardware warranty.
5. After all data has been entered, click **Save IBM Support**. At this point, the Service Advisor is not enabled.
6. To enable Service Advisor, click the **Service Advisor Settings** tab again and click **Enable IBM Support**. The Service Advisor Disabled status, as shown in Figure 9-10 on page 460, now shows Enabled.

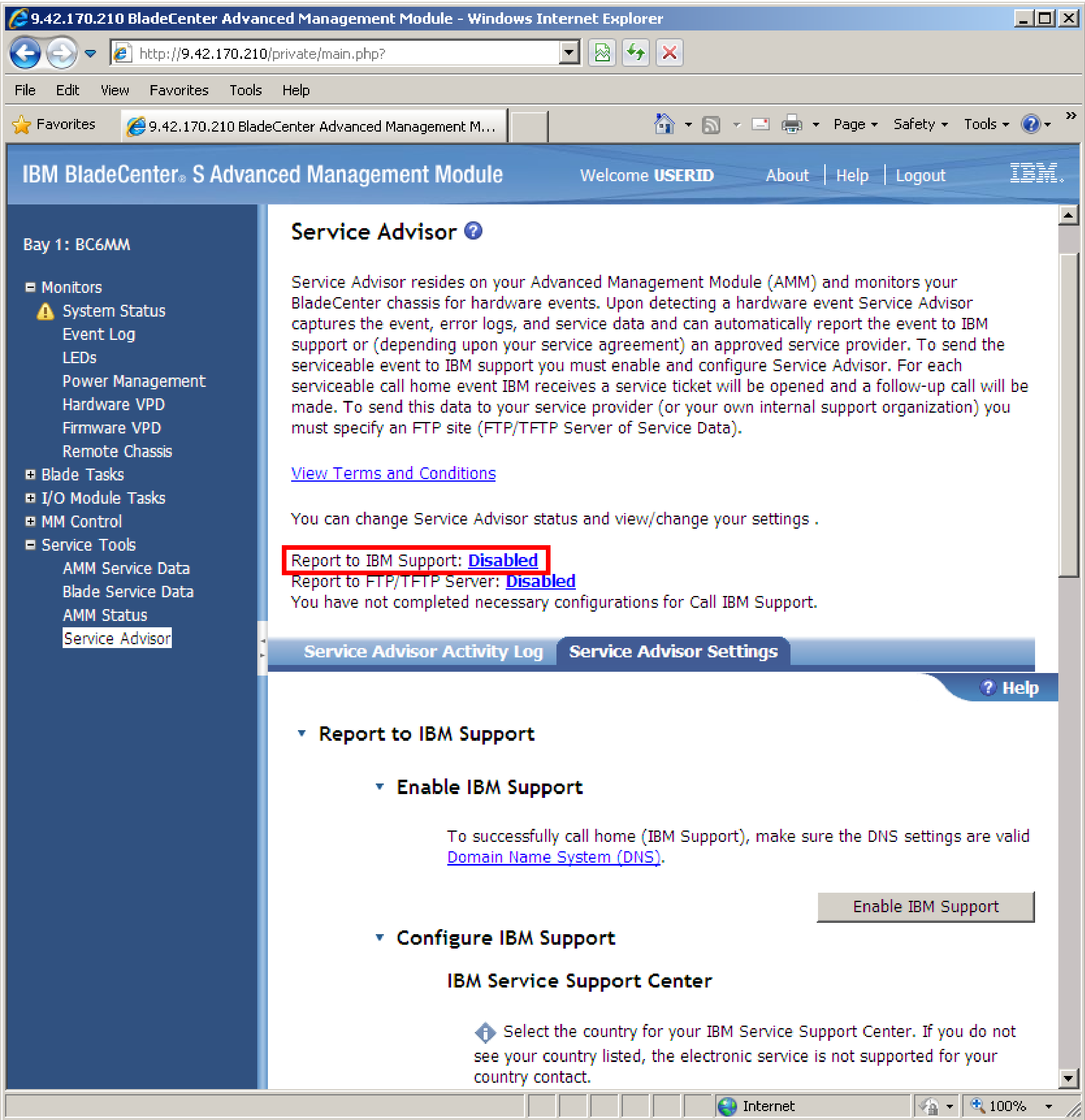


Figure 9-10 Service Advisor settings

7. It is advisable to generate a test call to IBM to ensure that the BladeCenter chassis can call home correctly. Select the **Test Call Home** tab and click **Test Call Home**. The Test Call Home tab only appears if the Service Advisor is set to **Enabled**.
8. You will be returned to the Service Advisor Activity Log tab after you click Test Call Home.
9. Click **Refresh** on the activity log until a success or failure is registered in the Send column of the activity log. If the call was successful, a ticket number appears in the Assign Num column. The ticket that is opened at IBM is identified as a test ticket. No action is required from IBM support for a test ticket, and the call will be closed. See the Connectivity security for Service Advisor section of the *IBM BladeCenter Advanced Management Module*

Installation Guide if you experience difficulties configuring the Service Advisor. You can obtain this guide at the following website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5073392

9.3.3 Updating firmware using the AMM

Through the AMM, you can apply following updates:

- IMM
- UEFI
- FPGA
- pDSA

You can use the `.exe` or `.sh` file to make an update.

Use this preferred order to update the firmware:

1. IMM
2. UEFI
3. FPGA
4. pDSA

Use these recommendations for updating through the AMM:

- After the IMM update is complete, wait at least 15 minutes before you initiate any planned UEFI or DSA Preboot firmware updates.
- Ensure that TFTP is enabled. Select **MM Control** → **Network Protocols** → **Trivial File Transfer Protocol (TFTP)**, as shown in Figure 9-11.

Figure 9-11 How to enable TFTP in the AMM

- If you have an AMM level 50G or earlier and you will update multiple blades in the chassis, restart the AMM once before beginning multiple updates.
- For the UEFI update, the server must be powered up and booted into the operating system or in the UEFI menu.

Use the AMM to update firmware:

1. Log in to the AMM web interface.
2. Select **Blade Tasks** → **Firmware update.**
3. Select the target blade, as shown in Figure 9-12 on page 462.

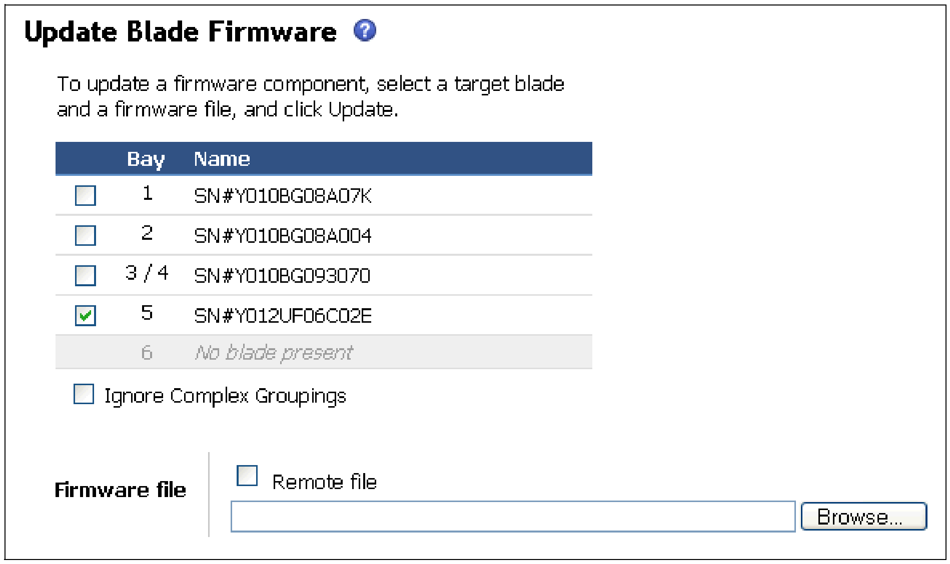


Figure 9-12 Overview of the blades that you select for update

4. Use **Browse** to locate the update file.
5. Click **Update** to start the update.

9.4 Remote control

The ability to control the eX5 systems remotely is provided by the IMM, in conjunction with the IBM virtual media key, for the x3690 X5 and the x3850 X5. The IBM virtual media key adds graphical remote control functionality to the IMM. The IBM virtual media key ships standard with the x3690 X5 and the x3850 X5.

The AMM, in conjunction with the IMM, provides the remote control functionality for the HX5. The available features for all three systems are similar with regard to using the graphical remote control. You can perform the following common tasks with the remote control function:

- Control the power of the systems
- Mount remote media, which includes CD/DVD-ROMs, supported ISO and firmware images, and USB devices
- Create your own customized keyboard key sequences using the soft key programmer
- Customize your viewing experience

9.4.1 Accessing the Remote Control feature on the x3690 X5 and the x3850 X5

Follow these steps to use the Remote Control feature on the IMM for the x3690 X5 and the x3850 X5:

1. Log in to the IMM of the specific system that you want to control.
2. Select **Tasks** → **Remote Control,** as shown in Figure 9-13 on page 463.
3. To protect sensitive disk and KVM data during your session, click the **Encrypt disk and KVM data during transmission** check box before starting Remote Control. For complete security, use Remote Control in conjunction with SSL. You can configure SSL at **IMM Control** → **Security**.

4. If you want exclusive remote access during your session, click **Start Remote Control in Single User Mode**. If you want to allow other users remote console (KVM) access during your session, click **Start Remote Control in Multi-user Mode**.

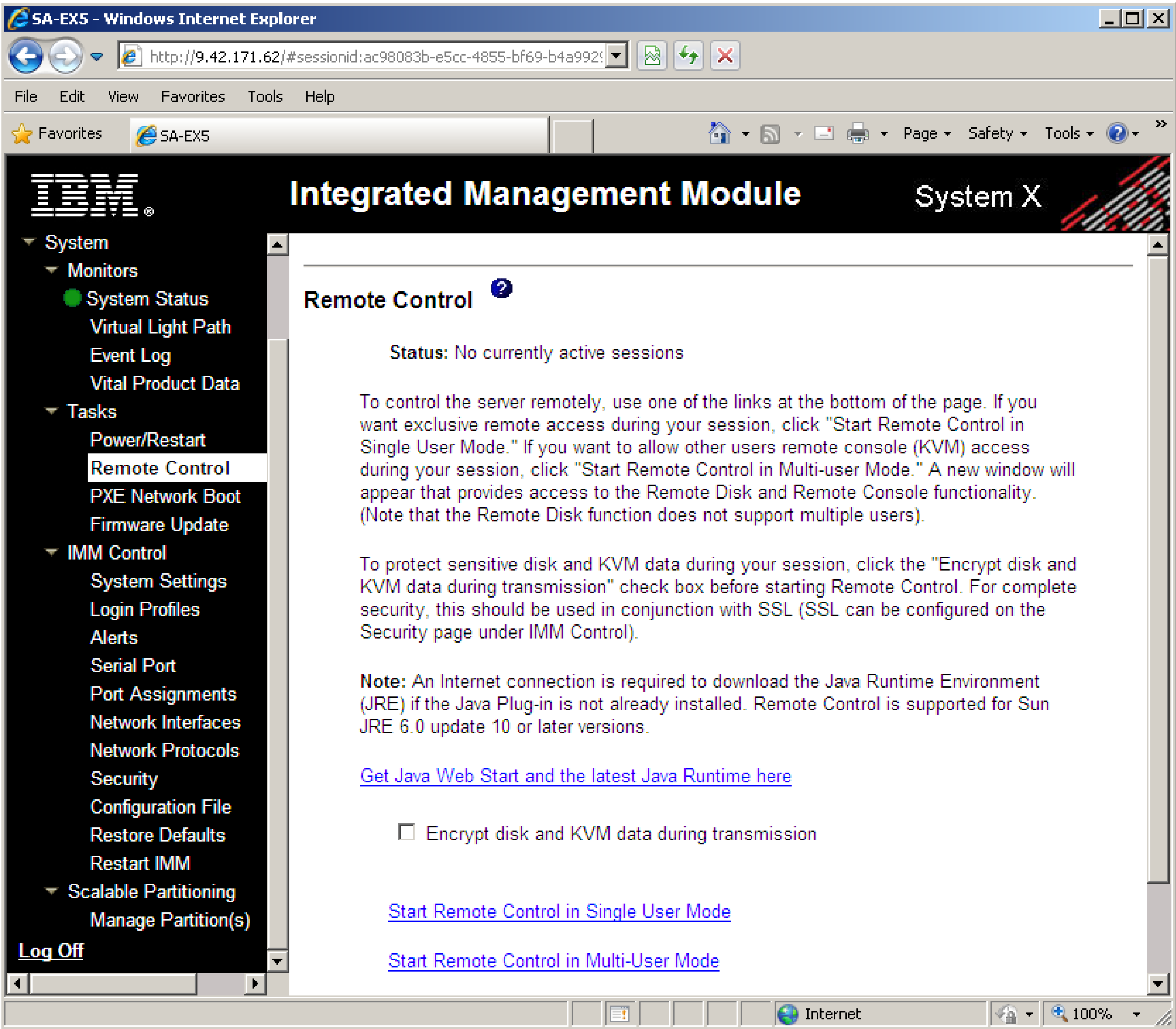


Figure 9-13 IMM Remote Control page

5. Two separate windows open: one window for the Video Viewer (as shown in Figure 9-14 on page 464) and the other window for the Virtual Media Session (as shown in Figure 9-15 on page 465). The various controls that you have to control the server are all contained in the *User's Guide for Integrated Management Module - IBM BladeCenter and System x*, which you can access at the following website:

   http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5079770

Figure 9-14 IMM Remote Control Video Viewer showing power control options

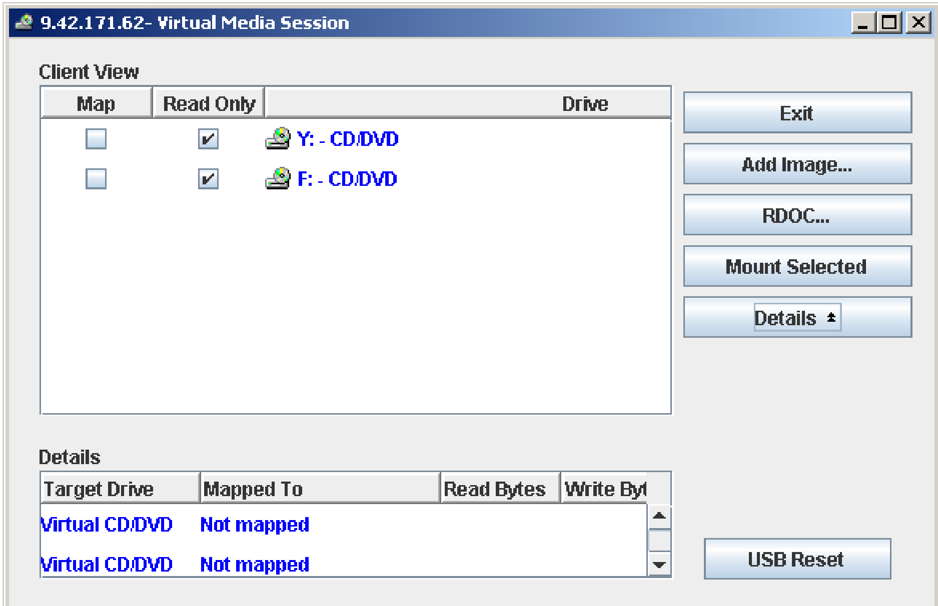


Figure 9-15 IMM Remote Control Virtual Media Session

9.4.2 Accessing the Remote Control feature for the HX5

You manage the Remote Control feature for the HX5 through the Advanced Management Module. Perform these steps:

1. To gain control of an HX5 blade, log in to the AMM.
2. In the left navigation pane, select **Blade Tasks** → **Remote Control**. The available remote control tasks are contained in three sections:
   – Remote Control Status: This section shows which blade is the current kvm and media tray owner of the chassis and allows you to change the blade ownership. Also, you can perform this task directly from the Remote Control pane.
   – Start Remote Control: This section allows you to start a remote control session to any of the blades contained within the chassis.
   – Remote Control Settings: This section controls how KVM access is managed for the blades within the chassis. As with the IMM, you can also specify whether to allow multiple concurrent sessions to the same blade. If you want exclusive remote access to a blade, clear **Allow multiple concurrent remote video sessions per blade**, as shown in Figure 9-16 on page 466.

> **Multiple concurrent remote video sessions:** The “Allow multiple concurrent remote video sessions per blade” setting is a system-wide setting, which applies to all of the blades in the respective BladeCenter chassis.

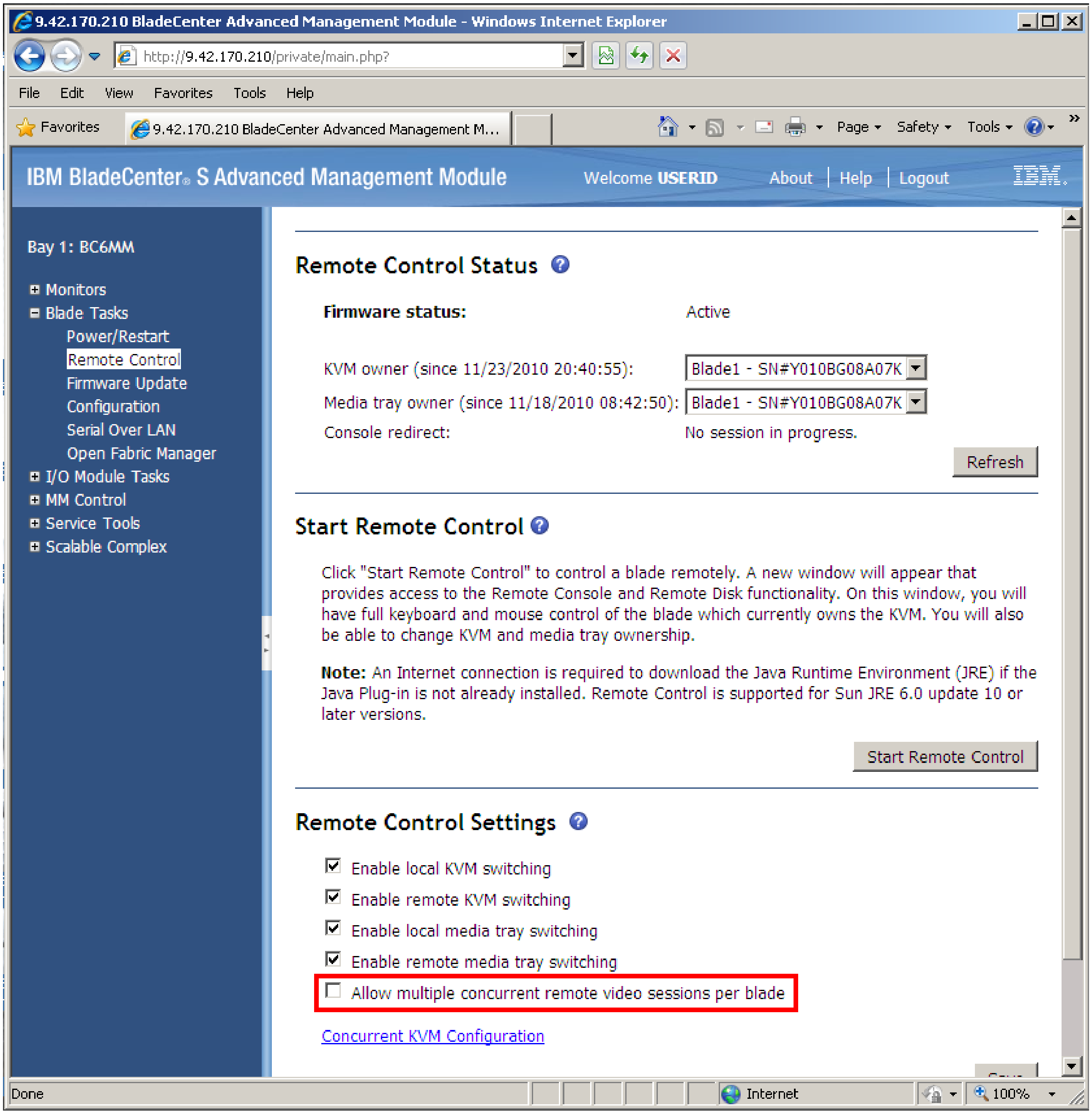


Figure 9-16 Remote Control options page

3. To select a blade that you want to control, click **Start Remote Control**.
4. To view the video of a blade, select the blade from the pull-down list box, as shown in Figure 9-17 on page 467. You can obtain additional instructions for power control, mounting remote media, and soft key programming in the *IBM BladeCenter Advanced Management Module Installation Guide* at the following website:

   http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5073392

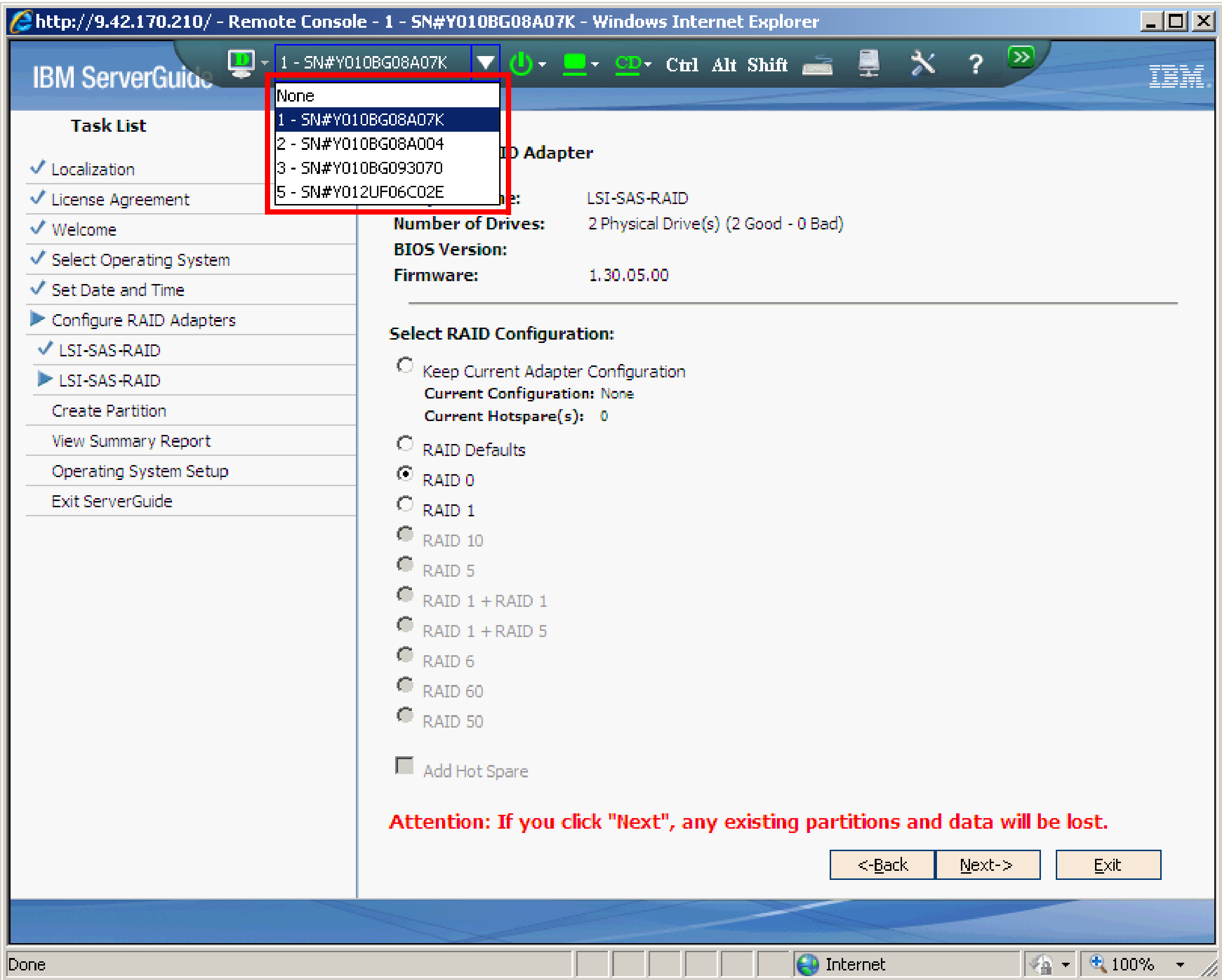


Figure 9-17 Selecting which blade's video output to display

9.5 IBM Systems Director 6.2

IBM Systems Director is a platform manager that offers the following benefits:

- Enables the management of eX5 physical servers and virtual servers that are running on the eX5 platform.
- Helps to reduce the complexity and costs of managing eX5 platforms. IBM Systems Director is the platform management tool for the eX5 platform that provides hardware monitoring and management.
- Provides a central control point for managing your eX5 servers and managing all other IBM servers.

You connect to IBM Systems Director server through a web browser. You can install IBM Systems Director server on the following systems: AIX®, Windows, Linux on Power, Linux on x86, or Linux on System z®.

IBM Systems Director provides the following functionality for eX5 systems, as well as many other system types:

- Discovery and inventory
- Monitoring and reporting
- Software updates
- Configuration management
- Virtual resource management
- Remote control, which includes power control
- Automation

> **Tip:** Your IBM Systems Director server must be at Version 6.2.1 to support all the eX5 systems, including their ability to scale into a complex. This version is also required to support the MAX5.

In 9.2, "Integrated Management Module (IMM)" on page 449, we described the concepts of in-band and out-of-band management of the IMM. In this section, we demonstrate how to discover the eX5 systems via IBM Systems Director.

9.5.1 Discovering the IMM of a single-node x3690 X5 or x3850 X5 out-of-band via IBM Systems Director

After your IMM has been configured for out-of-band management, as described in "Configuring an x3850 X5 or x3690 X5 for out-of-band management" on page 449, you can discover the system from within IBM Systems Director. After the IMM has been added to the IBM Systems Director console, you can perform management tasks against it.

Perform these steps to add the IMM to the IBM Systems Director console:

1. Log in to the IBM Systems Director web interface by navigating to the following website, where *servername* is the Domain Name System (DNS) registered name of your IBM Systems Director:

   `http://servername:8421/ibm/console`

   For example: `http://director6.ibm.com:8421/ibm/console`

   You can also connect to the IBM Systems Director web interface using its IP address:

   `http://ipaddress:8421/ibm/console`

   For example: `http://182.168.1.10:8421/ibm/console`

   > **Tip:** We advise that you configure your IBM Systems Director server correctly to use DNS for name resolution. We also recommend that you register the IMM in DNS to simplify the management of the IMM out-of-band.

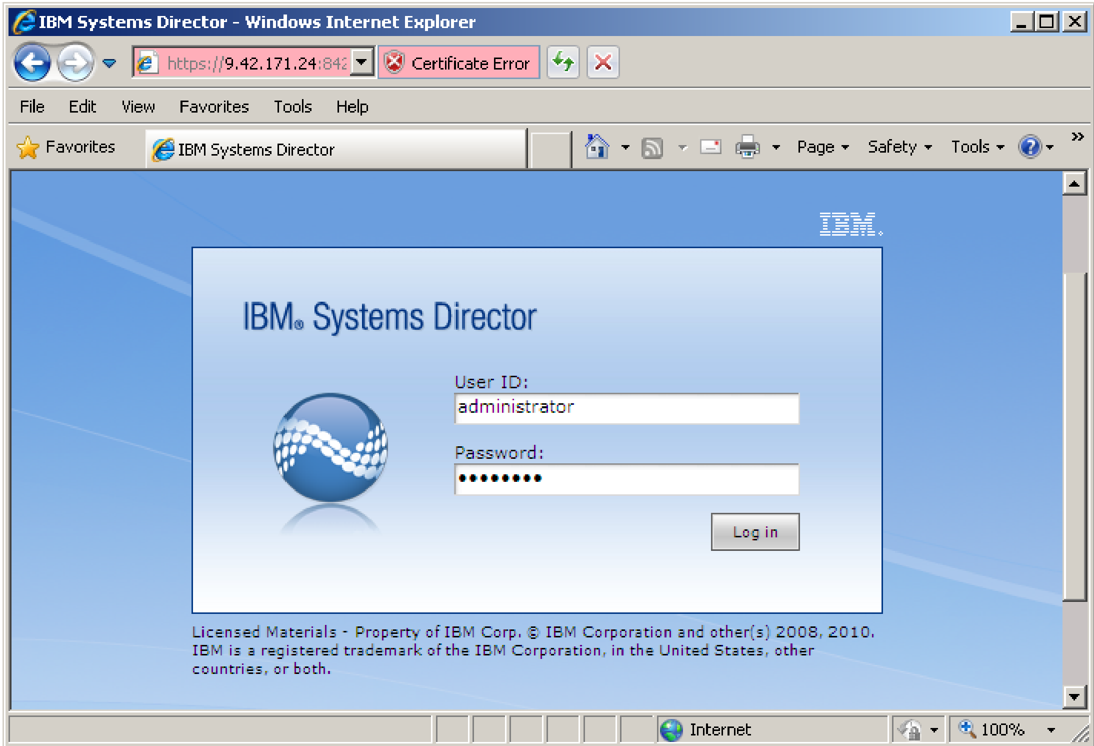


Figure 9-18 IBM Systems Director login web page

2. After you log in to the console, navigate to **Inventory** → **System Discovery** in the left navigation pane. In the right pane, select a discovery option in the Select a discovery option pull-down list. IBM Systems Director defaults to discovering a single system using IPv4 address resolution.
3. Enter the IP address of the IMM in the space that is provided under IP address.
4. An IMM is considered a server object in IBM Systems Director. To specify the IMM as a Server object, click the **Select the resource type to discover** list box and select **Server** as the resource type, as shown in Figure 9-19 on page 470.
5. Click **Discover Now** when you are ready to discover the IMM.
6. An informational message indicating that the discovery job has commenced appears at the top of the right pane. Click **Close Message** to acknowledge the message.
7. The discovered IMM appears at the bottom of the System Discovery pane under Discovered Manageable Systems. At this point, the IMM has been discovered but has not been authenticated to. Notice the `No Access` status under the Access column. Before authenticating to the IMM, we advise that you rename it first. The renaming process can be performed at a later stage as well, but we advise that you perform this process at discovery time, because it is easier to identify individually discovered IMMs for renaming than when they are in a group.

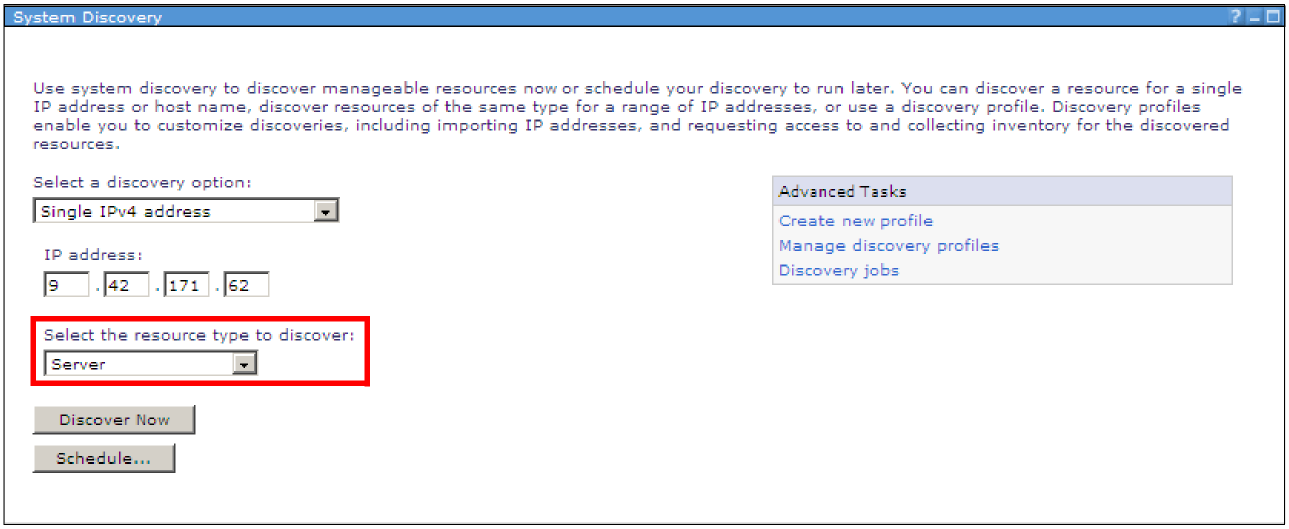


Figure 9-19 IBM Systems Director System Discovery pane

8. To rename the IMM, right-click the IMM server object in the Discovered Manageable Systems area and click **Rename,** as shown in Figure 9-20.

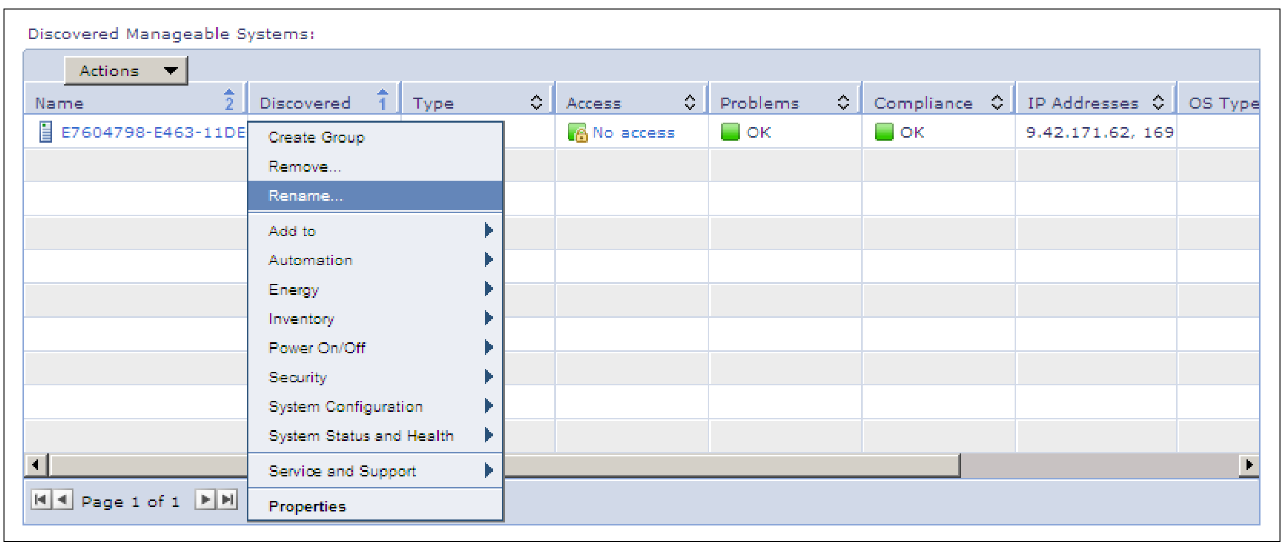


Figure 9-20 Renaming an IMM

9. A **Rename** display box opens, as shown in Figure 9-21. Provide a meaningful name for the IMM in the text box provided and click **OK** when finished.

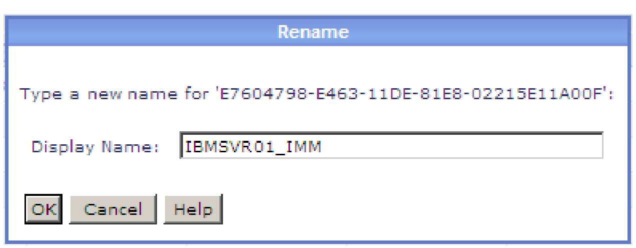


Figure 9-21 Rename box

10.To authenticate to the IMM, right-click **No access** under the Access column and click **Request Access,** as shown in Figure 9-22.

> **Tip:** If the access status appears as unknown, right-click the **Unknown** icon and select **Verify Connection**. The status changes to No access if IBM Systems Director can communicate correctly with the IMM.

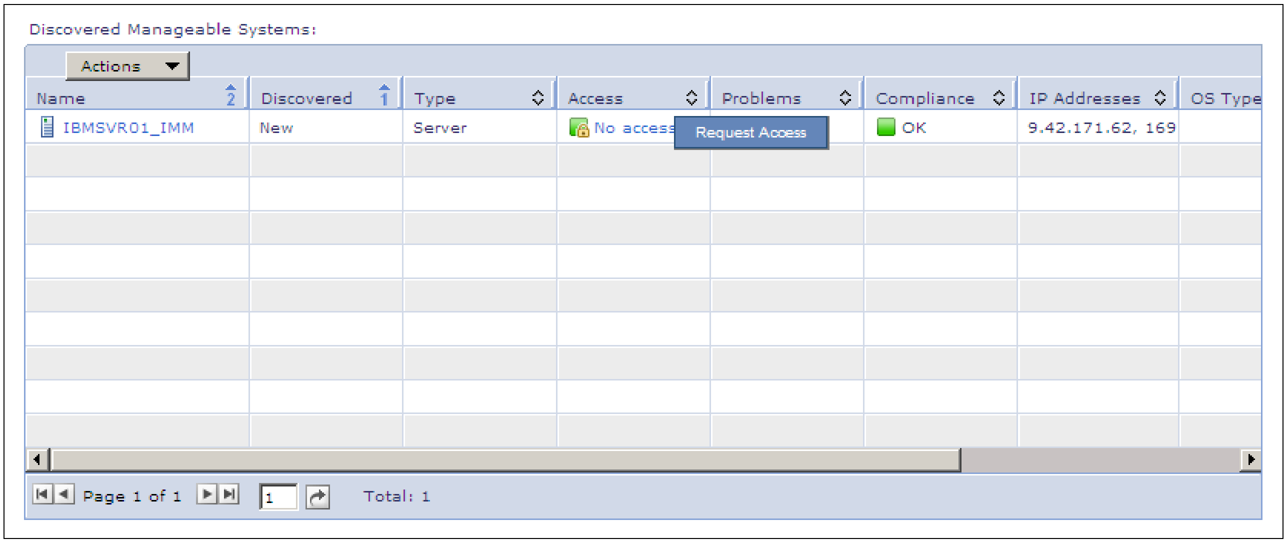


Figure 9-22 Request Access option

> **Tip:** You can also right-click the IMM and select **Security** → **Request Access** as an alternate method of authenticating to the IMM or any other supported object.

11.The Request Access pane opens, as shown in Figure 9-23. Enter the user name and password of an account that has supervisor access on the IMM in the text boxes provided and click **Request Access**.

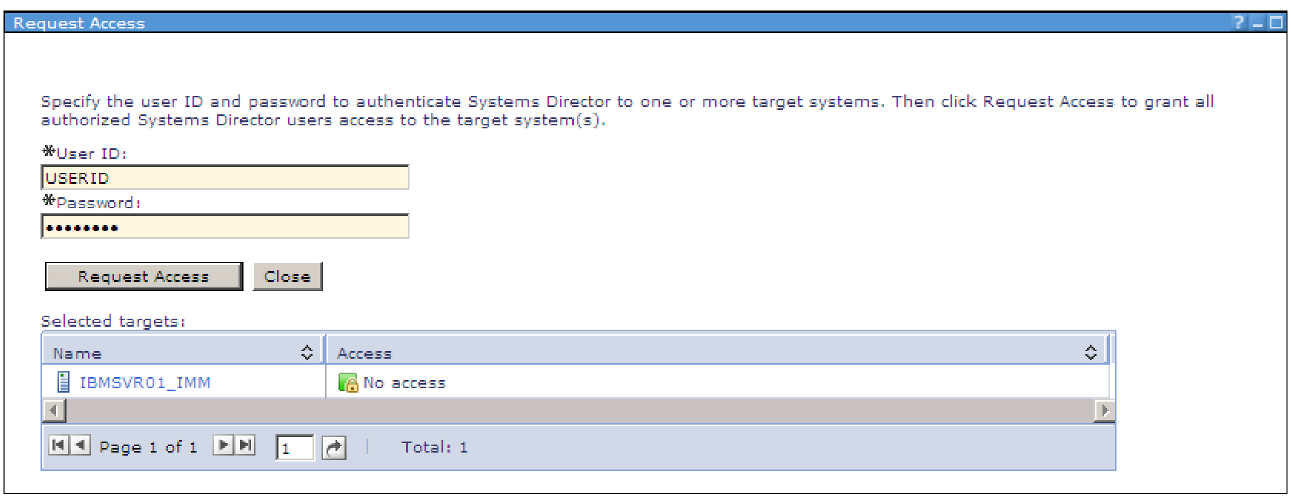


Figure 9-23 Request Access user credentials pane

12. A successful authentication to an IMM yields an OK access status, as shown in Figure 9-24. Click **Close** when finished.

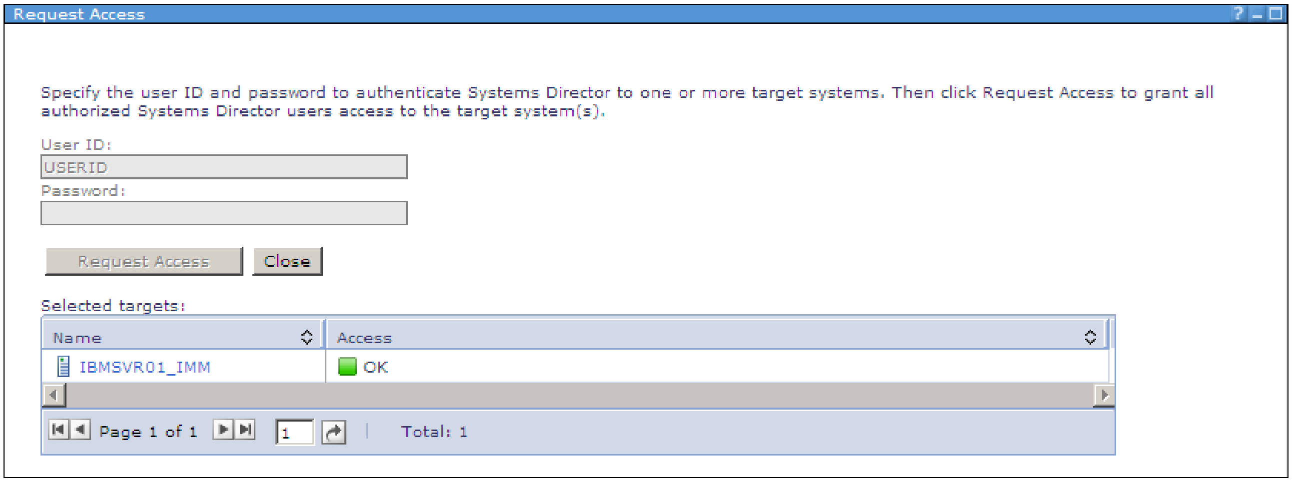


Figure 9-24 Successful authentication to an IMM

13. Click **Navigate Resources** in the left pane of IBM Systems Director. In the right pane, under **Groups (View Members)**, click the **All Systems** group. Right-click the IMM that you have discovered and select **Inventory** → **View and Collect Inventory**.

14. Your IMM will be displayed in the Target systems list box, as shown in Figure 9-25.

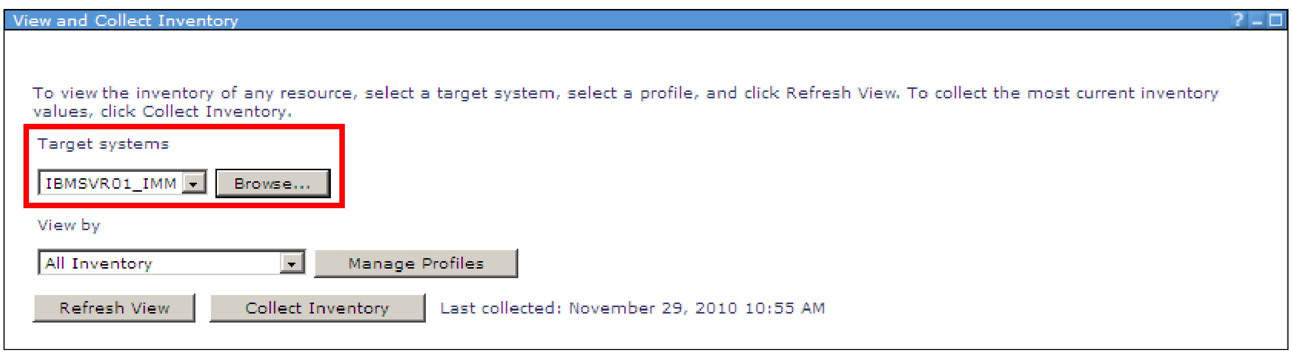


Figure 9-25 Collecting inventory for a managed system

15. Click **Collect Inventory**. Click **OK** to run the task immediately.

> **Tip:** Always collect the inventory of a system as soon as it is added to IBM Systems Director to ensure that all system data is captured immediately.

16. Click **Display Properties** to view the status of the running inventory collection task. After the task is 100% complete, you will then be able to perform tasks against the IMM.

9.5.2 Discovering a 2-node x3850 X5 via IBM Systems Director 6.2.x

The methods to discover a 2-node x3850 X5 (called a *complex*) are similar to discovering a single-node system.

Discovery tip: To correctly discover and manage a complex via IBM Systems Director, you must discover the system *both* out-of-band via the IMM *and* in-band via the operating system. A V6.2.1 or greater common or platform agent must also be installed.

Complete these steps to discover and manage an x3850 X5 complex running a supported operating system via IBM Systems Director:

1. Verify that your systems are configured correctly as a complex by logging in to the IMM web console and navigating to **Scalable Partitioning** → **Manage Partitions**. Verify that the values that are shown under System Partition Mode match the values that are shown in Table 9-2.

   Table 9-2 System Partition Mode status panel

System	Partition	Mode
Started	Valid	Multinode
Started	Valid	Multinode

   See 6.5, "Forming a 2-node x3850 X5 complex" on page 235 for instructions to configure a complex if you have not configured a complex already.
2. Ensure that you have installed the x3850 X5 complex with a supported operating system.
3. Discover and request access to the IMM to the primary node, as described in "Discovering the IMM of a single-node x3690 X5 or x3850 X5 out-of-band via IBM Systems Director" on page 468.
4. Notice that a Scalable Partition object and a Scalable System object appear in the **Navigate Resources** → **All Systems** group view. Also, you can see these objects by navigating to **Navigate Resources** → **Groups by System Type** → **System x** → **Scalable Systems** → **Scalable Systems and Members**. We explain the Scalable System and the Scalable Partition objects:

   Scalable System Refers to the system containing all the physical nodes that have been cabled together.

   Scalable Partition Refers to a logical partition running on the scalable system. Older generation multinode IBM systems scaled together can contain multiple partitions. Each partition can then run its own operating system and function independently, even though it is physically cabled to other nodes.
5. The access statuses for these objects appear as Partial access, because you have not yet authenticated to the second IMM in the complex.
6. Discover and request access to the IMM of the secondary node, as described in "Discovering the IMM of a single-node x3690 X5 or x3850 X5 out-of-band via IBM Systems Director" on page 468.
7. Wait one or two minutes before checking the access statuses of the Scalable System and Scalable Partition objects. Their statuses change to OK, as shown in Figure 9-26 on page 474.

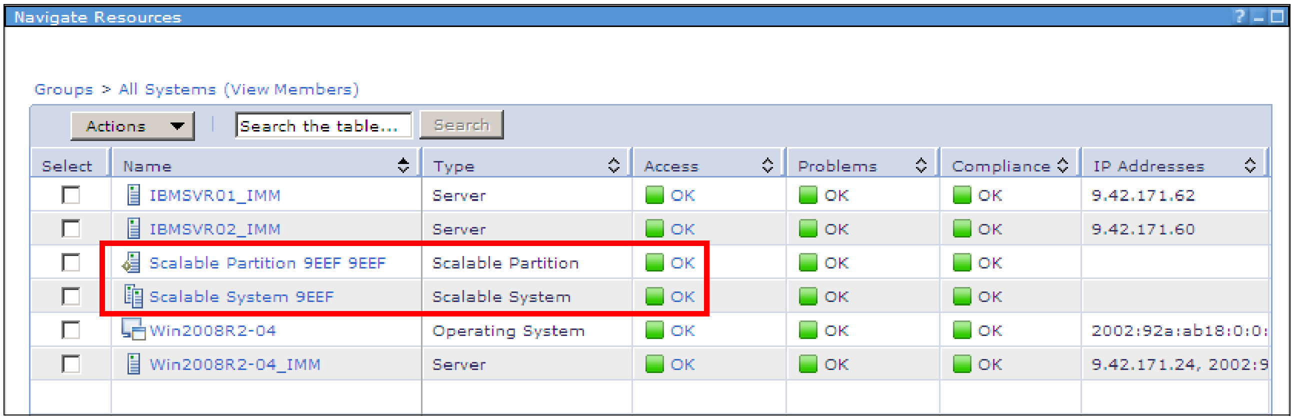


Figure 9-26 Scalable complex with correctly configured access in IBM Systems Director

8. Discover the IP address of the operating system running on the primary node by navigating to **Inventory** → **System Discovery**.

> **Importing a common or platform agent:** To deploy a common or platform agent via IBM Systems Director, you must import the agent first. See the following website for this procedure, because this procedure is out of the scope of this IBM Redbooks publication:
>
> `http://publib.boulder.ibm.com/infocenter/director/v6r2x/index.jsp?topic=/com.ibm.director.agent.helps.doc/fqm0_t_working_with_agent_packages.html`

9. Specify the IP address of the operating system under Select a discovery option.
10. Specify **Operating System** as the resource type to discover under the Select the resource type to discover list box. You must specify Operating System to enable IBM Systems Director to use the correct discovery protocols for the respective resource type. Click **Discover Now** when finished.
11. Rename the system if required and request access to it, as described in step 7 on page 469 to step 12 on page 472.
12. After the operating system is discovered, you need to deploy an agent to it to ensure that IBM Systems Director can manage the complex correctly. Select **Navigate Resources** → **All Operating Systems** under the Groups (View Members) pane.
13. Right-click the operating system object that you have discovered and select **Release Management** → **Install Agent,** as shown in Figure 9-27 on page 475.

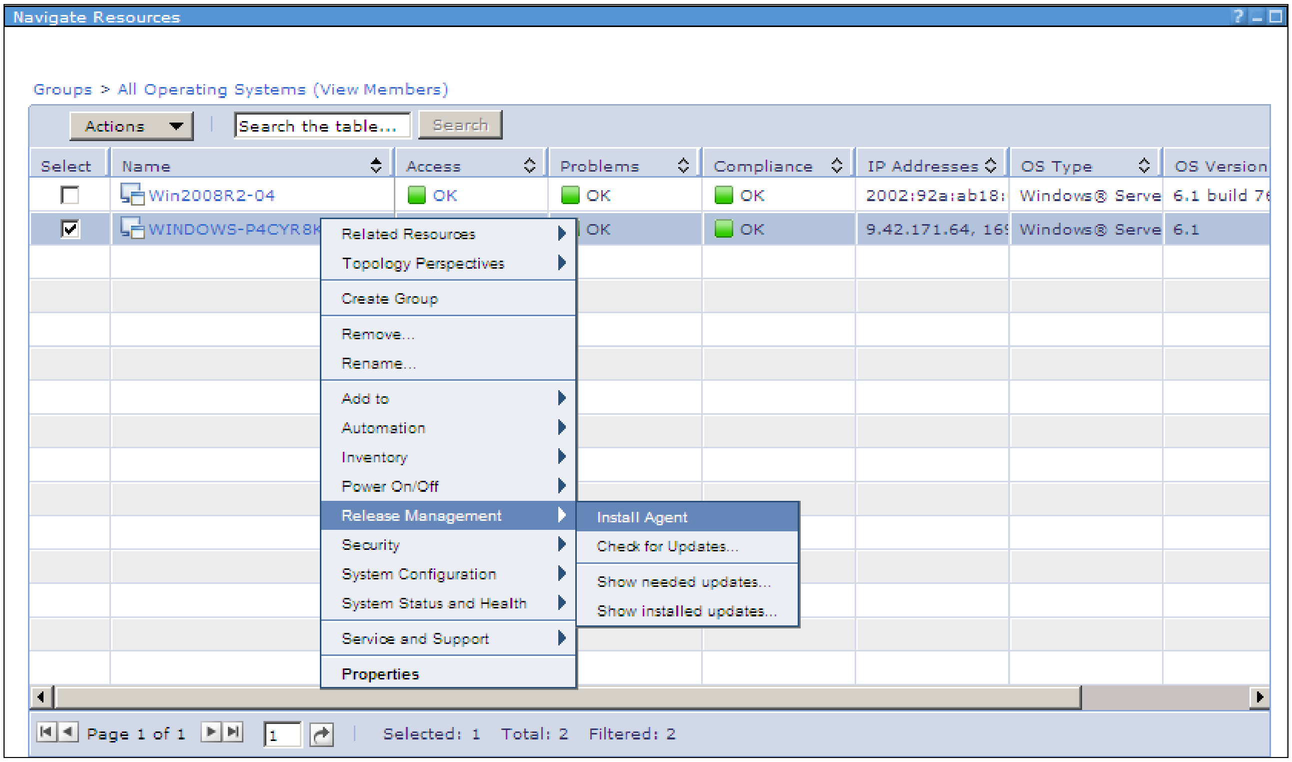


Figure 9-27 Deploying an agent to a managed system

14. The agent installation wizard window opens. Click **Next**.
15. Click the **Platform Agent Packages** group in the left pane and select the platform agent package that is relevant to your operating system. For Linux and WMware operating systems, select the PlatformAgent 6.2.1 Linux package. In this example, a Windows-based server was used; therefore, we selected the **PlatformAgent 6.2.1 Windows** package by clicking the radio button next to the agent and clicking **Add**. The selection options look like Figure 9-28 on page 476.

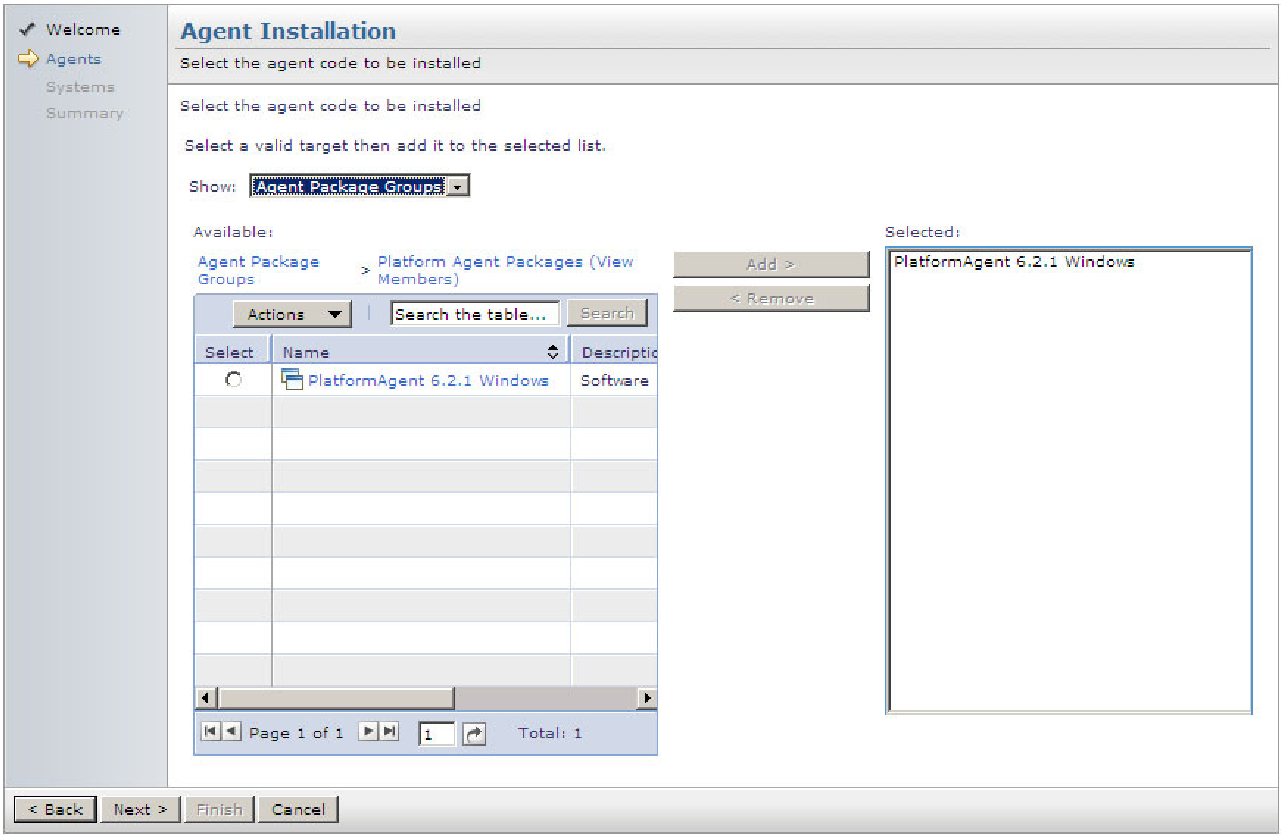


Figure 9-28 Platform agent selection pane

16. Click **Next**.
17. Click **Next** again.
18. Click **Finish**.
19. You are prompted to either schedule the deployment of the agent or deploy it immediately. For our example, we deployed the agent immediately. IBM Systems Director defaults to **Run Now** for all tasks unless specified otherwise. Click **OK** when ready to deploy the agent.

    Click **Display Properties** to view the status of the agent deployment process.

    > **Progress indicator:** The progress indicator might remain on 5% for a period of time before its value increases, which is normal.

20. After the agent deployment has completed, inventory the operating system object by navigating to **Navigate Resources** → **All Operating Systems**. Right-click the respective operating system object and click **Inventory** → **View and Collect Inventory**.
21. Click **Collect Inventory** and click **OK** to run the task immediately.
22. Click **Display Properties** again to view the progress of the inventory task.
23. After the task has completed, you are ready to view and manage the scalable systems.

9.5.3 Discovering a single-node HX5 via IBM Systems Director

Managing an HX5 with IBM Systems Director provides you added benefits, including centralized management of hardware alerts and inventory capabilities. The methods that are used to discover the HX5 IMM are similar to the methods that are used to discover the IMM of a stand-alone server. The only difference is that the IMM on the HX5 is not configured with an external IP address; therefore, discovery of the HX5 is carried out via the AMM.

Complete these steps to discover a single-node HX5:

1. Log in to the IBM Systems Director web interface by navigating to the following URL, where *servername* is the DNS registered name of your IBM Systems Director:

   `http://`*servername*`:8421/ibm/console`

   For example: `http://director6.ibm.com:8421/ibm/console`

   You can also connect to the IBM Systems Director web interface using its IP address:

   `http://`*`ipaddress`*`:8421/ibm/console`

   For example: `http://182.168.1.10:8421/ibm/console`

2. Select **Inventory** → **System Discovery** in the left pane of the IBM Systems Director console.
3. Enter the IP address in the text box.
4. In the Select the resource type to discover list box, make sure that you select **BladeCenter Chassis** as the resource type. This selection allows IBM Systems Director to use the correct discovery protocols to locate the chassis, as shown in Figure 9-29.

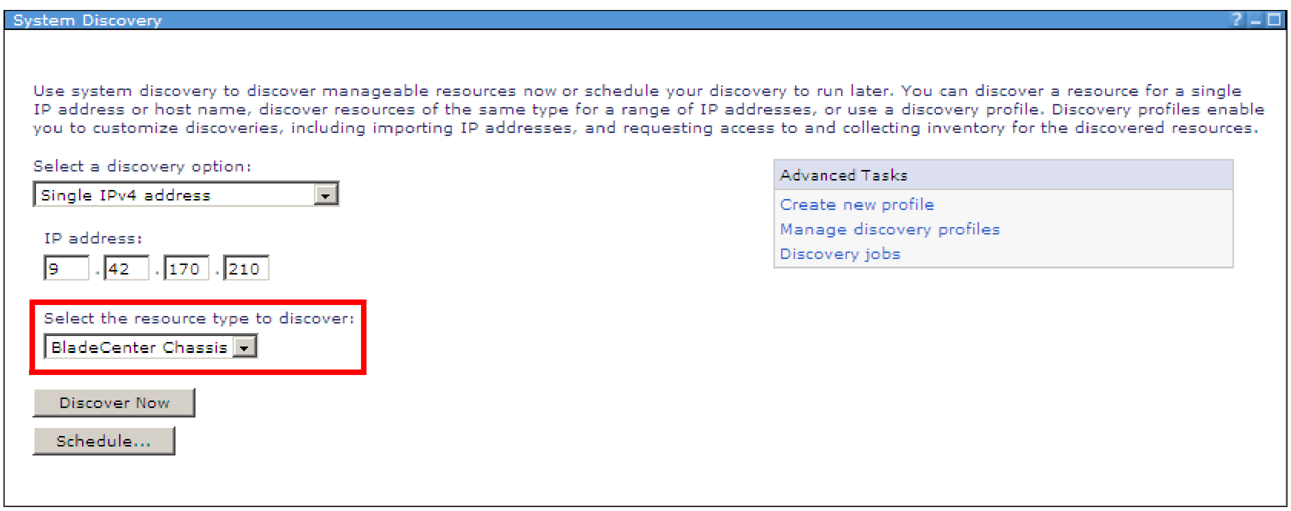


Figure 9-29 Selecting the correct resource type to discover a BladeCenter chassis

5. Click **Discover Now**. The discovered chassis appears in the Discovered Manageable Systems area at the bottom of the System Discovery pane.
6. Right-click the **No access** icon under the Access column and click **Request Access**.

> **Tip:** If the access status appears as Unknown, right-click the **Unknown** icon and select **Verify Connection**. Click **Verify Connection**. The status changes to No access if IBM Systems Director can communicate correctly with the IMM.

7. Enter the user name and password credentials of an account that has supervisor access to the AMM and click **Request Access** when finished.

> **Wait time:** Requesting access to a BladeCenter chassis might take time to complete. IBM Systems Director has to discover all the components within the chassis, including blades, I/O modules, power supplies, and so on.

8. Click **Close** after the process completes. You return to the System Discovery pane.
9. Close this pane and select **Navigate Resources**. In the Groups (View Members) pane, click **Groups by System Type** → **BladeCenter Systems** → **BladeCenter Chassis and Members**. All blade service processors (the IMM in the case of the HX5) and I/O module switches are displayed here. Notice the scaled HX5 that has been discovered in this group view, as shown in Figure 9-30.

   Also, note that the MAX5 does not show up as a separate item. The blade `IBM:7872-AC1-06EC578` has a MAX5 attached. Installing a platform agent onto the operating system of the HX5 allows you to view the memory that is installed in the HX5, as well as the memory that is installed in the MAX5.

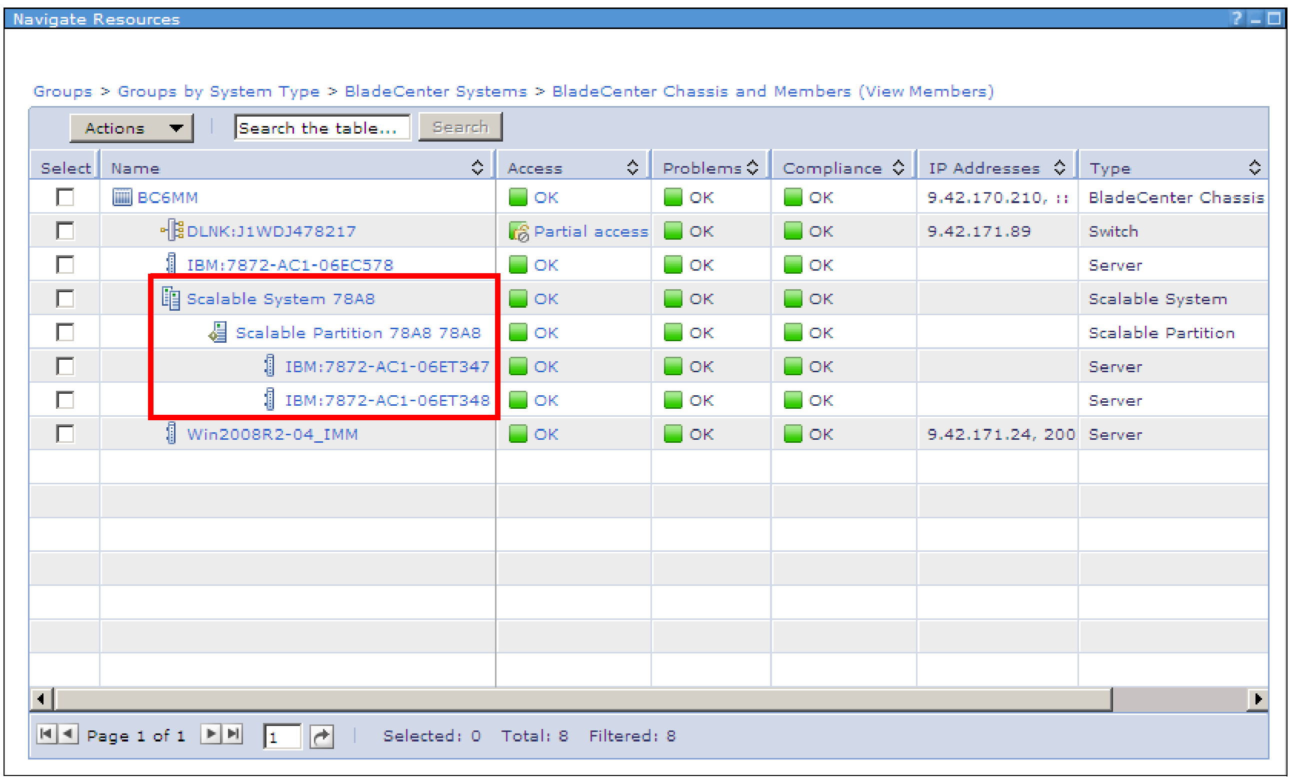


Figure 9-30 BladeCenter Chassis and Members (View Members) group view

You have completed the exercise of discovering a single-node HX5 out-of-band via IBM Systems Director.

9.5.4 Discovering a 2-node HX5 via IBM Systems Director 6.2.x

The methods to discover a 2-node HX5 (called a *complex*) are similar to discovering a single-node system.

Discovery tip: To correctly discover and manage a complex via IBM Systems Director, you must discover the system *both* out-of-band via the IMM *and* in-band via the operating system. A V6.2.1 or greater common or platform agent must also be installed.

For the purposes of this exercise, we demonstrate how to install a platform agent on an HX5 blade running Windows Server 2008. The method of deploying a platform agent to a Linux or WMware ESX server is the same.

Complete these steps to discover and deploy a platform agent to an HX5 complex:

1. Install Windows Server 2008 on the HX5 complex. We recommend using the IBM ServerGuide, because it installs all necessary drivers as part of the process. You can download the ServerGuide CD at the following website:

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

2. Ensure that you have configured the complex correctly by accessing the AMM and navigating to **Scalable Complex** → **Configuration**.
3. On the Scalable Complex Information window, click the **Complex** tab that contains the scalable systems. The numbers in parentheses on the tabs refer to the slots that contain the blades, as shown in Figure 9-31.

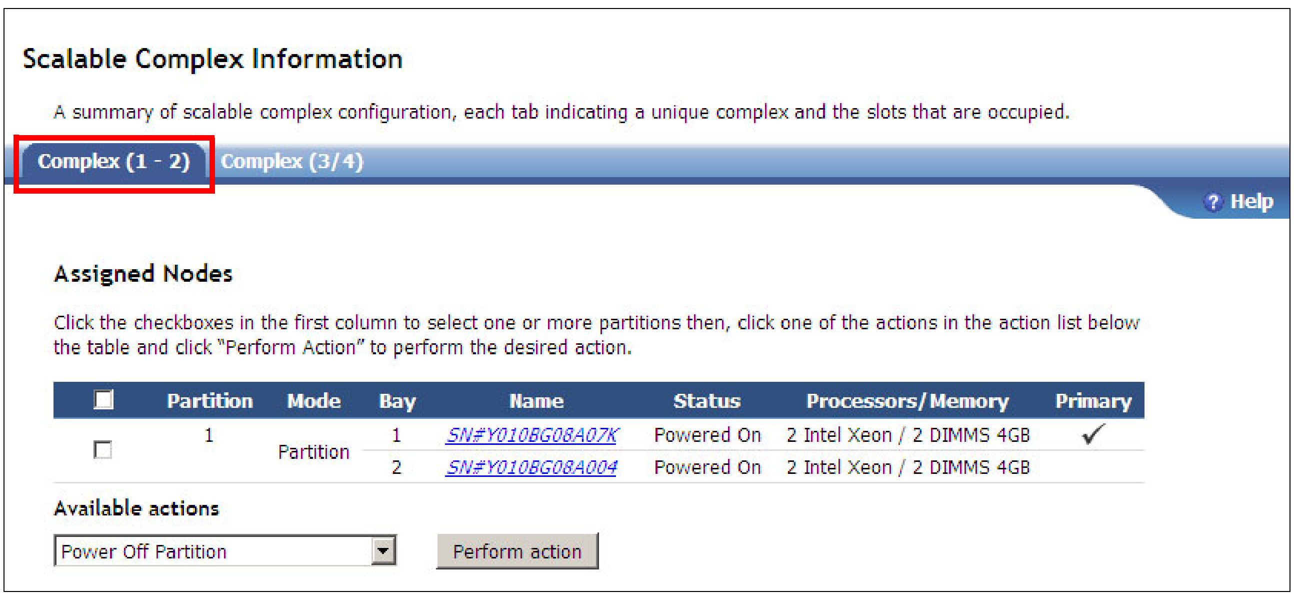


Figure 9-31 Selecting an HX5 complex to partition

4. Ensure that the Mode column status for the respective complex indicates Partition, as shown in Figure 9-31. If it does not indicate the Partition mode, see 8.6, "Creating an HX5 scalable complex" on page 402.
5. Complete the instructions in "Discovering a 2-node HX5 via IBM Systems Director 6.2.x" if you have not done so already.
6. You are still logged in to the IBM Systems Director web console at this point. Discover the IP address of the operating system running on the primary HX5 node by navigating to **Inventory** → **System Discovery**.

Importing a platform agent: To deploy a platform agent via IBM Systems Director, you must import the agent first. See the following website for this import procedure, because it is beyond the scope of this IBM Redbooks publication:

http://publib.boulder.ibm.com/infocenter/director/v6r2x/index.jsp?topic=/com.ibm.director.agent.helps.doc/fqm0_t_working_with_agent_packages.html

7. Specify the IP address of the operating system under Select a discovery option.
8. Specify **Operating System** as the resource type to discover under the Select the resource type to discover list box. You must specify **Operating System** to enable IBM Systems Director to use the correct discovery protocols for the respective resource type. Click **Discover Now** when finished.
9. Rename the system if required and request access to it, as detailed in step 7 on page 469 to step 12 on page 472.
10. After the operating system is discovered, you need to deploy an agent to it to ensure that IBM Systems Director can manage the complex correctly. Select **Navigate Resources** → **All Operating Systems** under the Groups (View Members) pane.
11. Right-click the operating system object that you have discovered and select **Release Management** → **Install Agent**.
12. The Agent Installation wizard window opens. Click **Next**.
13. Click the **Platform Agent Packages** group in the left pane and select the **PlatformAgent 6.2.1 Windows** package by clicking the radio box next to the agent. When deploying to a Linux or WMware ESX server, you need to select the PlatformAgent 6.2.1 Linux agent package. Click **Add**. Your selection options look like Figure 9-32.

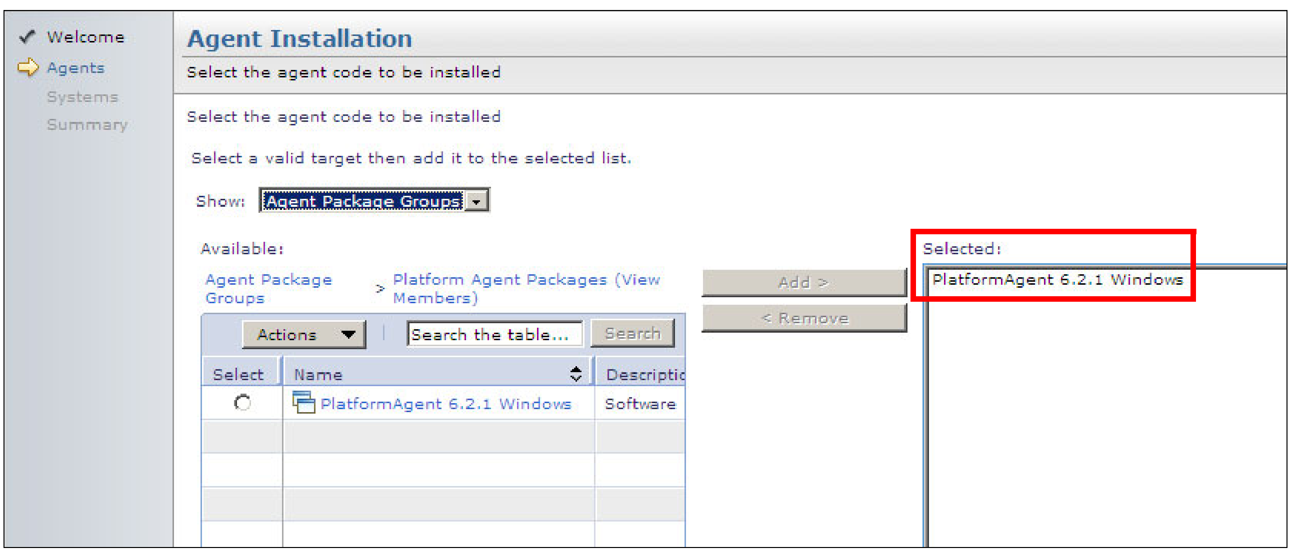


Figure 9-32 Selecting the platform agent to install onto Windows

14. Click **Next**.
15. Click **Next** again.
16. Click **Finish**.

17.You are prompted to either schedule the deployment of the agent or deploy it immediately. For our example, we deploy the agent immediately. IBM Systems Director defaults to **Run Now** for all tasks unless specified otherwise. Click **OK** when ready to deploy the agent.

Click **Display Properties** to view the status of the agent deployment process.

> **Progress indicator:** The progress indicator might remain on 5% for a period of time before its value increases, which is normal.

18.After the agent deployment completes, inventory the operating system object by going to **Navigate Resources** → **All Operating Systems**. Right-click the respective operating system object and click **Inventory** → **View and Collect Inventory**.

19.Click **Collect Inventory** and click **OK** to run the task immediately.

20.Click **Display Properties** again to view the progress of the inventory task.

21.After the task completes, you are ready to view and manage the scalable systems.

9.5.5 Service and Support Manager

Service and Support Manager is a no-charge plug-in to IBM Systems Director. It automatically captures hardware errors and reports them to IBM support for you on a 24/7 basis. It forms part of IBM Electronic Services, as described in 9.6, “IBM Electronic Services” on page 493. The tool is quick, easy to configure, and provides the following benefits:

- Automatic problem reporting
- Direct routing 24/7 of reported problems to IBM technical support
- Reduced personnel time required for gathering and reporting service information
- Higher availability and shorter downtime
- Custom IT management tools enabled
- Secure Internet access
- Accurate solutions with reduced human error in gathering and reporting service information
- Secure Web access to your service information
- Consistent IBM worldwide service and support process

You can access the Service and Support Manager plug-in on the IBM Systems Director 6.2.1 DVD or you can download it from the following website:

https://www14.software.ibm.com/webapp/iwm/web/reg/pick.do?lang=en_US&source=dmp

> **Version tip:** Service and Support Manager versions are specific to the version of the IBM Systems Director server that is installed. In our example, we have IBM Systems Director Version 6.2.1; therefore, we must use Service and Support Manager Version 6.2.1.

You need to register at no charge if you do not have an IBM ID to access this website.

Service and Support Manager prerequisites

Before installing Service and Support Manager, you must meet the following prerequisites:

- The IBM Systems Director server must have access to the Internet. The access can be granted via a proxy server, but the following configuration rules must be followed for Service and Support Manager and Update Manager to work correctly.

  Configure the proxy server to use basic authentication if it is configured for digest or NT LAN Manager (NTLM) authentication. The update manager task supports only basic authentication with the proxy server. If digest or NTLM authentication is required, the update manager is unable to access update packages from IBM.

- Service and Support Manager requires access to the host names, IP addresses, and ports that are shown in Table 9-3.

  Table 9-3 Remote servers to which Service and Support Manager needs access

Remote server	IP addresses	Port
eccgw01.boulder.ibm.com	207.25.252.19	443
eccgw02.rochester.ibm.com	129.42.160.51	443
www-945.ibm.com	129.42.26.224 129.42.34.224 129.42.42.224	443
www6.software.ibm.com	170.225.15.41	443
www.ecurep.ibm.com	192.109.81.20	443

> **Tip:** IP addresses are subject to change, so ensure that you use host names whenever possible.

- Collect inventory for the systems before you start the installation. This task is not mandatory, but it is recommended.

Installing and configuring Service and Support Manager

In this section, we describe how to install and configure Service and Support Manager from the IBM Systems Director 6.2.1 DVD:

1. Place the IBM Systems Director DVD into the IBM Systems Director server. The DVD runs automatically.

   > **Tip:** If you have disabled the auto-run feature on your IBM Systems Director server, you can find the Service and Support Manager software in the `SSM` directory directly in the root of the DVD. Open the `SSM` directory and double-click **SysDir6_2_1_Service_Support_Windows.exe** to start the installation.

2. Select the language that you want to use from the list box and click **OK**.
3. The IBM Systems Director Welcome page appears. Click **IBM Service and Support Manager**.
4. Click **Install IBM Service and Support Manager 6.2.1**, as shown in Figure 9-33 on page 483.

Figure 9-33 Installing Service and Support Manager

5. Select the language again that you want to use from the list box and click **OK**.
6. The Introduction pane appears. Accept the license terms if prompted and click **Next**.
7. Click the **Restart IBM Systems Director** check box if you want Service and Support Manager to be enabled immediately after installation, as shown in Figure 9-34 on page 484. Leave the Restart IBM Systems Director check box unchecked if you will restart the IBM Systems Director server at a later stage. Click **Next** when ready.

Figure 9-34 Restart IBM Systems Director option

8. Click **Install**. The Service and Support Manager installer stops the IBM Systems Director server service. The installation completes. Click **Done** when finished. The IBM Systems Director server service starts after the installation of the Service and Support Manager completes.
9. Log in to the IBM Systems Director server console after the IBM Systems Director service has started.
10. Click the **Manage** tab in the right pane and scroll down to the Service and Support Manager plug-in, as shown in Figure 9-35 on page 485. The Service and Support Manager icon is blue, which indicates that additional configuration is required for this plug-in to operate.

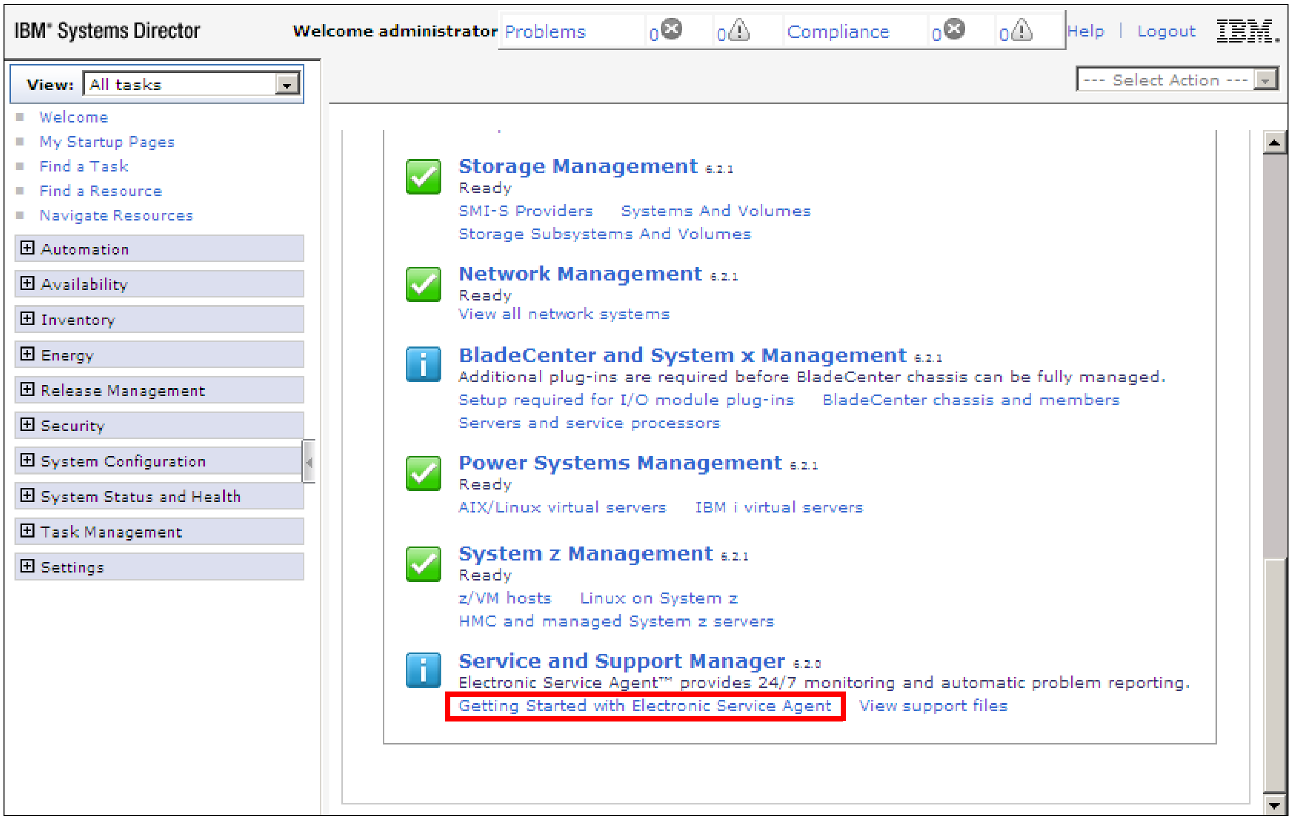


Figure 9-35 Service and Support Manager plug-in

11. Click **Getting Started with Electronic Service Agent**, as shown in Figure 9-35.
12. The Service and Support Manager wizard Welcome pane opens. Click **Next**.
13. Enter your company contact information in the provided fields. The more information that you provide, the easier it will be for IBM support to assist you. Pay particular attention to ensuring that the Country or region field is completed correctly, as shown in Figure 9-36 on page 486.
14. Click **Next**.

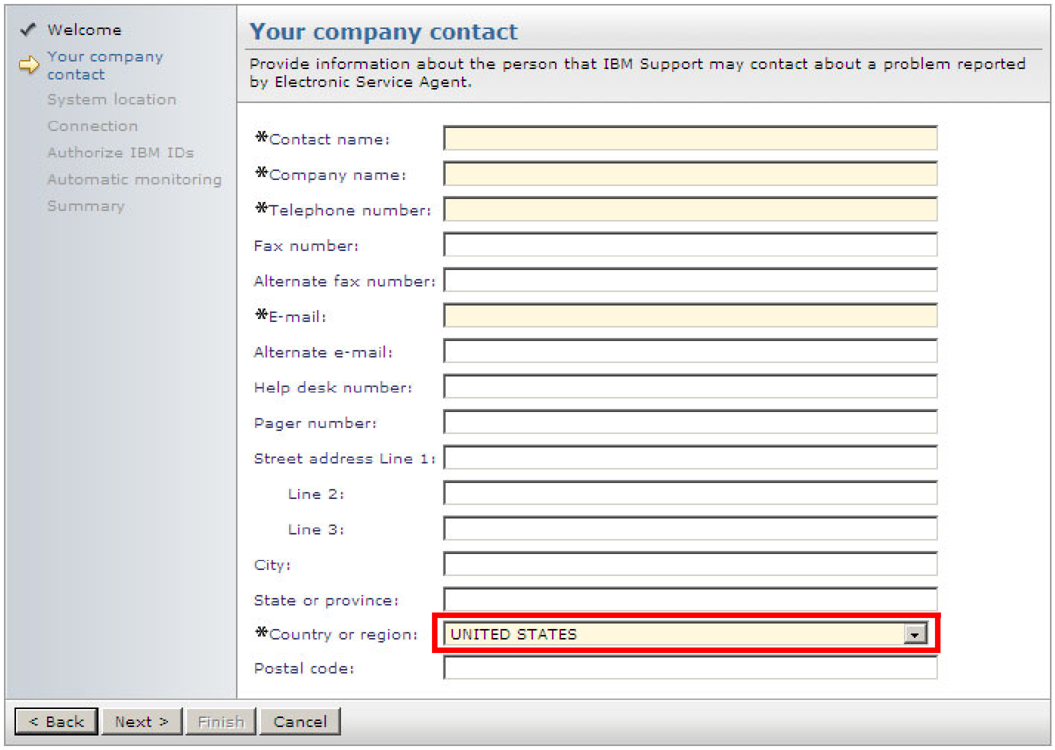


Figure 9-36 Service and Support Manager company contact details

15. Provide default details for the physical location of your systems, as shown in Figure 9-37. You can change individual system location details at a later stage using the details that are provided in "Changing the location settings of a system" on page 488.

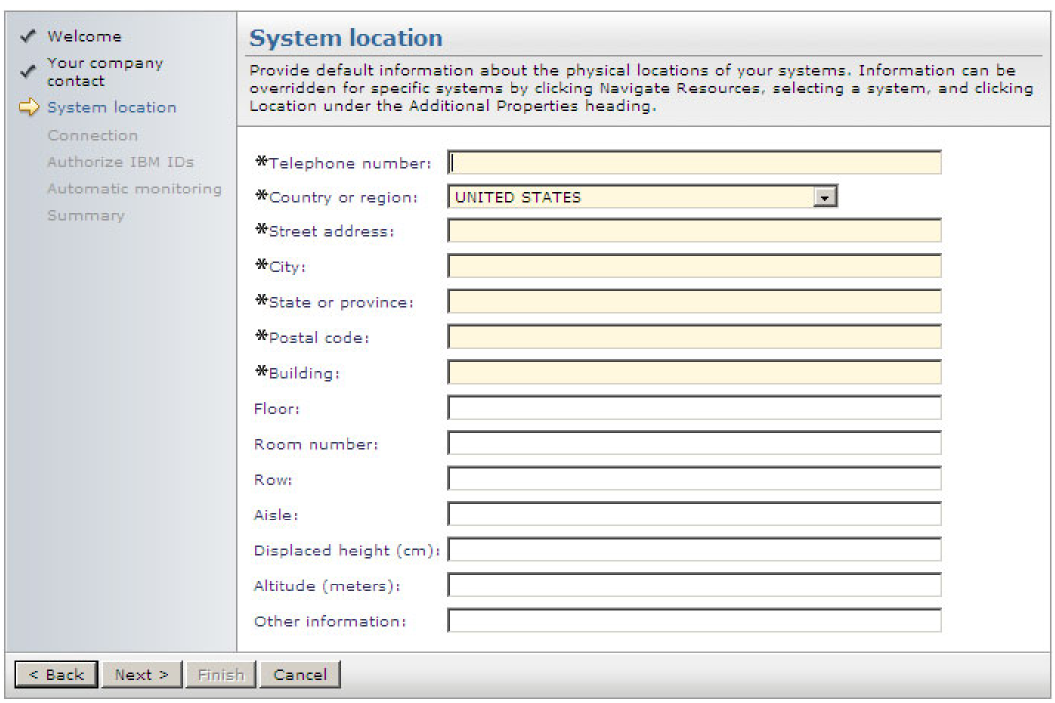


Figure 9-37 System location details

16.On the Connection pane, leave the default settings if your IBM Systems Director server has a direct connection to the Internet, as shown in Figure 17. Enter the proxy server details if you are required to connect to the Internet via a proxy server. Always ensure that you test the Internet connection by clicking **Test Connection** when finished. Click **Next** after a successful Internet connection has been confirmed.

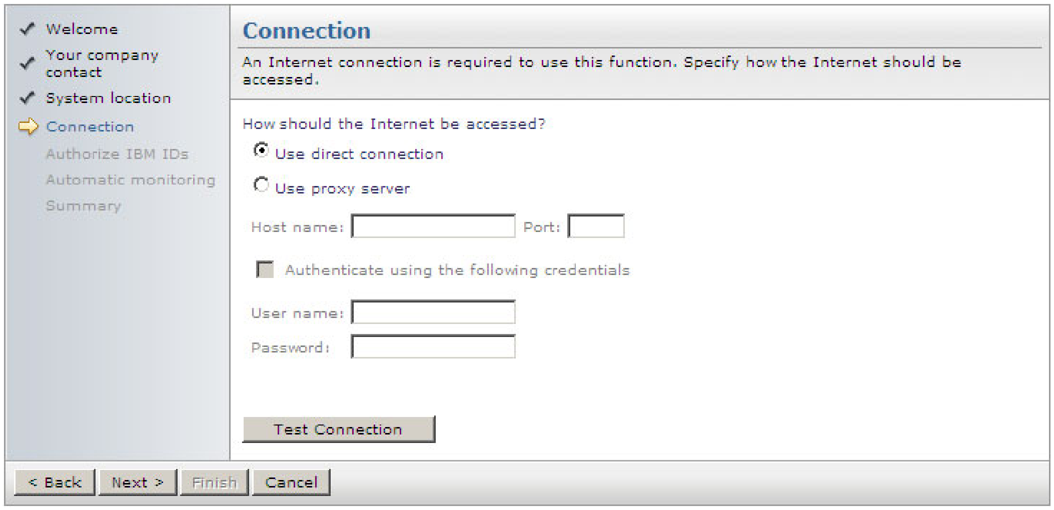


Figure 9-38 Service and Support Manager Connection configuration

17.Provide the authorized IDs of the personnel who need access to the service information that is transmitted to IBM. This information is not a requirement to activate Service and Support Manager.

If you have not already created the IDs, you can do so by clicking the link, as shown in Figure 9-39.

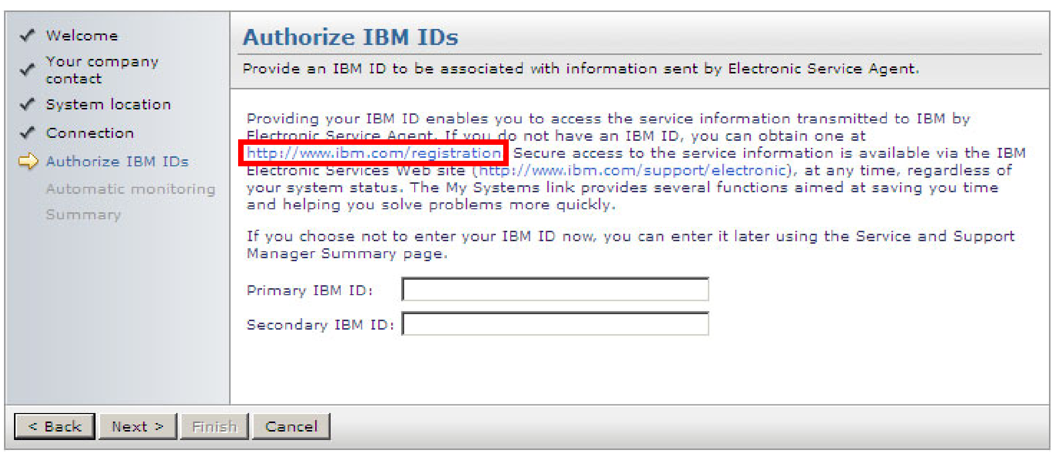


Figure 9-39 Authorize IBM IDs pane

18.On the Automatic monitoring pane, leave the check box checked if you want all newly discovered systems to be monitored by Service and Support Manager. Click **Next**.

19.Click **Finish** when done. You return to the Manage tab and the status of the Service and Support Manager plug-in is now green.

The service now actively monitors all eligible systems that are monitored by IBM Systems Director. You can click the **Service and Support Manager** plug-in on the Manage tab to view systems that might have a serviceable problem. See the IBM Systems Director 6.2.x

Information Center, which is available at this website, for further details regarding Service and Support Manager:

http://publib.boulder.ibm.com/infocenter/director/v6r2x/index.jsp

Changing the location settings of a system

Perform these steps to change the location settings of a system:

1. Click **Navigate Resources** → **All Systems**.
2. Right-click a system and select **Properties**.
3. Click **Location** on the Properties pane.
4. Click **Edit** to provide the location details. Click **OK** when finished.

9.5.6 Performing tasks against a 2-node system

After the 2-node system (complex) has been discovered and a platform agent has been installed onto the scaled system, you can perform the following additional tasks:

- Power control of the complex
- Capability to view the inventory of the complex
- System identification
- Firmware management of the complex using Update Manager

Partitioning: You cannot control the partitioning of a complex via IBM Systems Director. You must use the IMM or the AMM where applicable. You can, however, launch the IMM or AMM web interface via IBM Systems Director.

We demonstrate several of these tasks in the following sections.

Complex power control via IBM Systems Director 6.2.x

Many methods are available to control power to a scaled system within IBM Systems Director. You can perform the power on and off functions against the following objects from within the console:

- IMM (when it is not configured as part of a complex)
- Scalable partition
- Operating system

In most instances, use the operating system object to control the power of the complex unless an operating system has not been installed. Use the following procedure to power on a complex via IBM Systems Director using the operating system object:

1. Select **Navigate Resources** → **All Systems** in the Groups (View Members) view. Right-click the respective operating system object and select **Power On/Off** → **Power On**, as shown in Figure 9-40 on page 489.

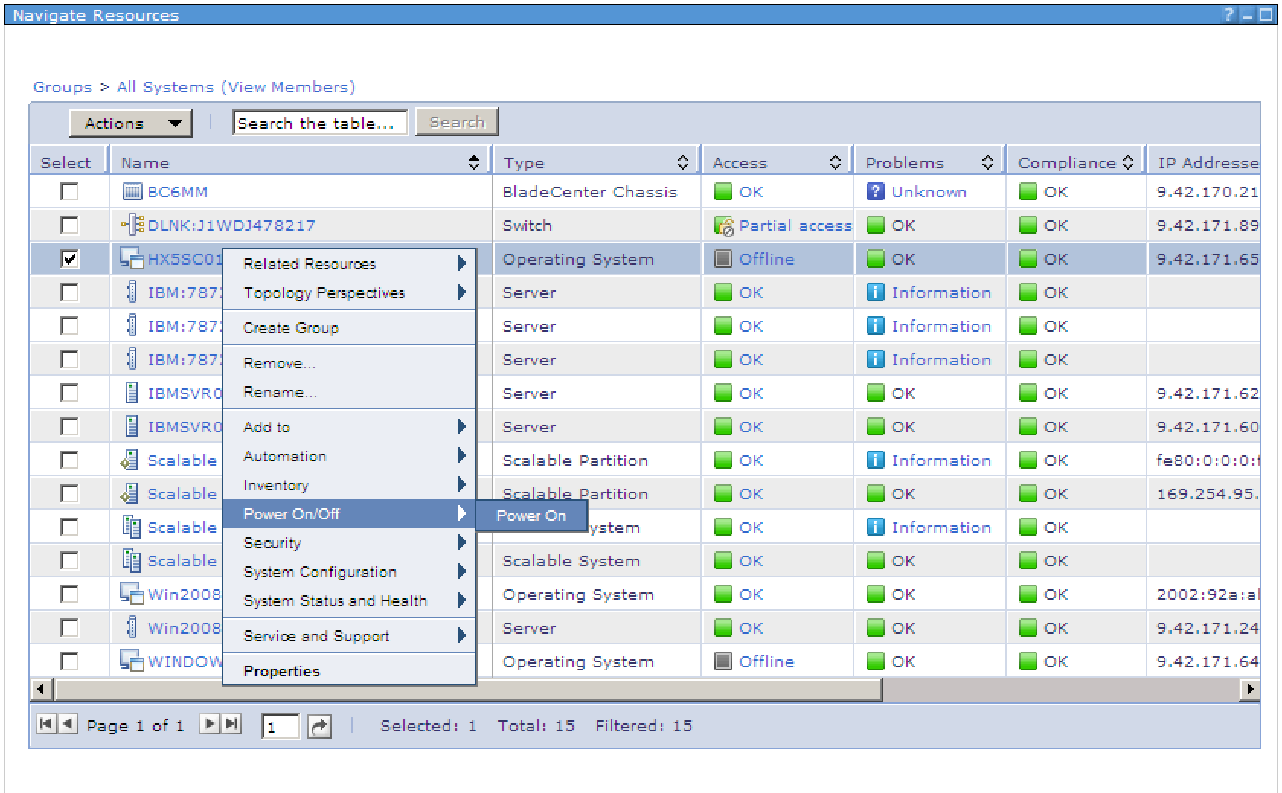


Figure 9-40 Powering on a complex via the operating system object

2. Click **OK** when the Task Launch Dialog pane appears to run the task immediately.

The system powers on normally. The power control menus within IBM Systems Director are adaptive, which means that they change based on the power state of the system. Figure 9-41 shows the available options when the same system is powered on.

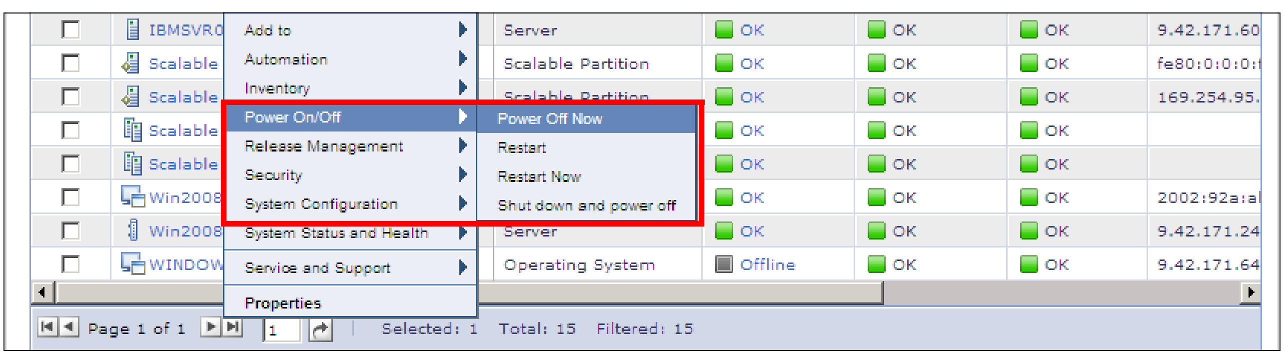


Figure 9-41 Power off controls for the operating system object

Viewing the installed hardware and software on a complex

Complete these steps to view information pertaining to the hardware and software components that are installed in a complex:

1. Select **Navigate Resources** → **All Systems** in the Groups (View Members) view.
2. Right-click the respective scalable partition object, select **Related Resources**, and select the hardware component for which you want to view the information.

3. To view the installed software, right-click the respective scalable partition object, select **Related Resources**, and then select **Installed Software**. All software, including system drivers, is displayed for that system.

You can also view additional information about the complex by either viewing the properties of the Scalable System or Scalable Partition objects. Use the following procedure:

1. Right-click either the **Scalable System** or **Scalable Partition** objects and select **Properties** from the menu. Figure 9-42 shows the properties of the Scalable Partition object indicating the number of nodes participating in the partition, as well as various other information.

Figure 9-42 Properties of a scalable partition

Identifying a system via IBM Systems Director

The ability to remotely identify systems becomes important when you have a large number of systems in one rack or many racks. You can, for example, illuminate the system identification light on a system in a remote location for someone else to locate that system to replace a part. Under normal circumstances, you need to switch on the system identification light via the IMM web console. IBM Systems Director allows you to perform this task for all your servers from a single location. Perform these steps to illuminate the system identification light for a system that is managed by IBM Systems Director:

1. Log in to the IBM Systems Director console.
2. Select **Navigate Resources** and click the **All Systems** group.

3. Right-click the IMM of the server or blade that you want to identify and select **System Status and Health** → **System Identification** → **LED On**, as shown in Figure 9-43.

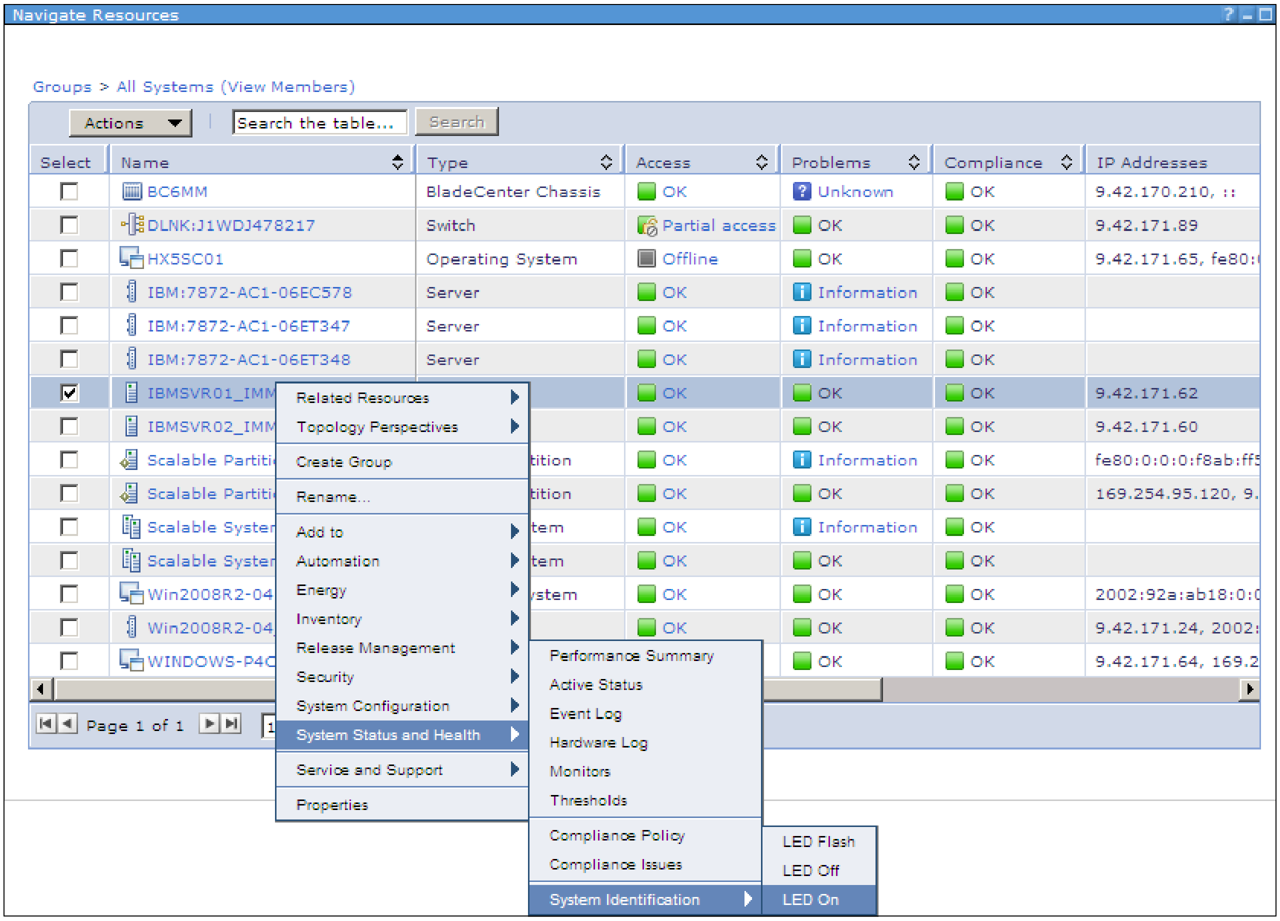


Figure 9-43 Identifying a system within the IBM Systems Director console

> **Tip:** You can only perform system identification against the IMM object for IMM-based systems. You cannot perform this operation against the operating system, Scalable System object, or Scalable Partition object.

Updating the firmware of a complex using IBM Systems Director

Managing the server firmware, as needed and when needed, is a necessary but time-consuming task. The IBM Systems Director Update Manager makes this process easier by providing centralized management of firmware deployment for the eX5 servers.

The ability to maintain identical firmware across servers in a complex is critical for their operation. In a scalable system, Update Manager is responsible for keeping the following four types of system firmware at the same level on all physical servers across the system:

- DSA
- FPGA
- IMM
- UEFI

If you update the firmware on the operating system on the primary node in a complex, the second node attached to it will also be updated.

Consider the following items when updating the system firmware on a complex or scalable partition:

- Before starting any system firmware update processes, ensure that the multinode systems are discovered with both in-band and out-of-band methods. For more information, see the following sections for in-band and out-of-band configuration:
  - 9.5.2, "Discovering a 2-node x3850 X5 via IBM Systems Director 6.2.x" on page 472
  - 9.5.4, "Discovering a 2-node HX5 via IBM Systems Director 6.2.x" on page 478
- The system firmware updates are installed to all the physical server systems within a system partition. All of the systems within the system partition are then rebooted after the installation.
- To update the firmware of either an x3850 X5 or an HX5 complex via IBM Systems Director, you need to configure Update Manager to download updates from the Internet. See Chapter 10, Update Manager, in *Implementing IBM Systems Director 6.1*, SG24-7694, which is available at this website:

  http://www.redbooks.ibm.com/abstracts/sg247694.html

Use the following procedure to update the firmware of either a x3850 X5 or an HX5 complex via IBM Systems Director:

1. Log in to the IBM Systems Director console.
2. Select **Navigate Resources** and click the **All Systems** group.
3. Right-click the operating system object of the complex that you want to update and select **Release Management** → **Show needed updates**. Any updates that are applicable to be deployed to the system will be displayed, as shown in Figure 9-44. For our example, we imported the latest UpdateXpress System Pack bundle for the HX5 into IBM Systems Director.

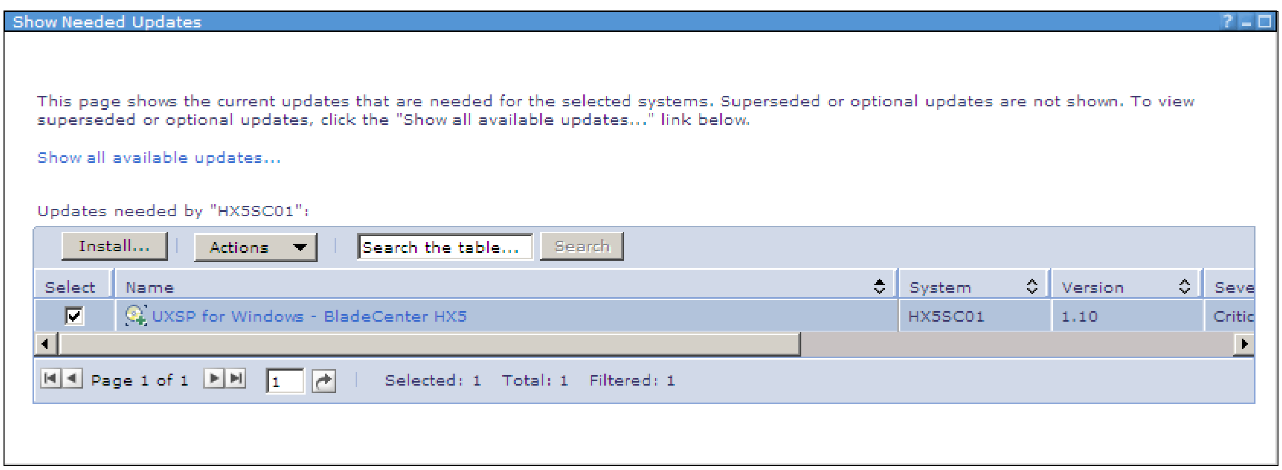


Figure 9-44 Available updates to be deployed to the HX5 scalable system

4. Click the check box next to the applicable update and click **Install**.
5. The update wizard window appears. Click **Next**.
6. Leave the default setting of **Automatically install missing required updates** on the Options window and click **Next**.
7. Update Manager will need to restart the systems that form the complex during the firmware update process. Leave **Automatically restart as needed during installation** as the default. Click **Next**.

8. A summary of the updates to be installed will be displayed, as shown in Figure 9-45. Click **Finish** when ready to proceed.

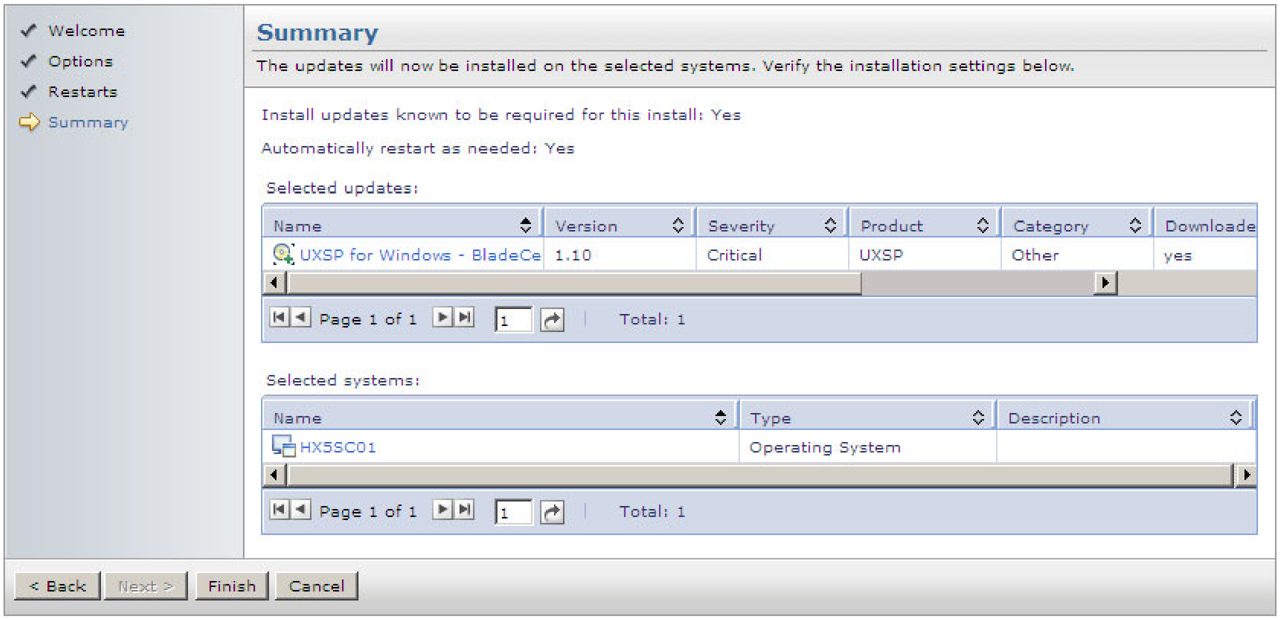


Figure 9-45 Summary of firmware to be installed

9. Click **OK** to run the firmware update immediately. The server will need to be rebooted so the update process can be scheduled if required. Wait for the firmware update process to complete, which might take time depending on the number of required updates to be installed.

9.6 IBM Electronic Services

Electronic Services is an IBM support approach that is made up of the Electronic Service Agent (ESA) and the Electronic Services website.

The Electronic Service Agent is a no-charge software tool that gets installed on your system to monitor events and securely transmit system inventory information to IBM. ESA's two key functions, automatic hardware problem reporting and service inventory information collection, enable proactive and predictive services, as well as faster problem resolution and call avoidance.

ESA tracks and captures machine inventory, hardware error logs, and automatically reports hardware problems to IBM if the server is under a service agreement or warranty. The ESA inventory includes the following information:

- Your support contact information, including names, phone numbers, and email addresses
- System utilization
- Performance data
- System failure logs
- Part feature codes, part number, part serial number, and part locations
- Software inventory, including operating system applications
- Program temporary fixes (PTFs), including the maintenance levels and configuration values

The ESA inventory does not include the following information:

- Collection or transmission of any of your company’s financial, statistical, or personnel data
- Client information
- Your business plans

The web component of IBM Electronic Services offers a single location for you to access many IBM Internet service and support capabilities. You can also view and use the ESA inventory information from any location around the world.

The Electronic Services website offers the following information:

- A single portal for hardware and software information and reference materials
- My Systems to view and use ESA service information in customized reports, such as hardware and software inventory, fixes, and system parameters
- My Search facility that uses ESA information to provide customized results for your specific machines from the IBM reference databases
- A single place to submit a service request for either hardware or software, in any country
- My Messages to receive information for specific platforms or individual profile definition
- Access to web-delivered premium services, such as Performance Management or Enhanced Technical Support (ETS) contracted services
- My Links to customize the web view by your selections of IBM system platforms
- Tutorials or demonstrations provided for all major areas of the website

For information about downloading and installing the relevant ESA for your systems and registering on the IBM electronic services website, see the following website:

https://www-304.ibm.com/support/electronic/portal/navpage?category=5

9.7 Advanced Settings Utility (ASU)

The Advanced Settings Utility (ASU) allows you to modify your server firmware settings from a command line. It supports multiple operating systems, such as Linux, Solaris, and Windows (including Windows Preinstallation Environment (PE)). The firmware settings that can be modified on the eX5 platform include UEFI and IMM settings.

You can perform the following tasks by using the ASU:

- Modify the UEFI CMOS settings without the need to restart the system and access the F1 menu.
- Modify the IMM setup settings.
- Modify a limited set of VPD settings.
- Modify the iSCSI boot settings. To modify the iSCSI settings through ASU, you must first manually configure the iSCSI settings through the server setup utility.
- Remotely modify all of the settings through an Ethernet connection.

ASU supports scripting environments through batch-processing mode. Download the latest version and the *Advanced Settings Utility User's Guide* from the Advanced Settings Utility website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-ASU

9.7.1 Using ASU to configure settings in IMM-based servers

ASU 3.x supports configuring settings on servers with integrated management modules (IMMs), such as the eX5 family of servers. The ASU uses the same set of commands and syntax that is used by previous versions of the ASU tool. Certain commands are enhanced to manage and display groups of settings, including new classes that are used as filters if you display the supported settings by using the **show** set of commands.

In IMM-based servers, you configure all firmware settings through the IMM. The ASU can connect to the IMM locally (in-band) through the keyboard console style (KCS) interface or through the LAN over USB interface. The ASU can also connect remotely over the LAN (out-of-band).

When the ASU runs any command on an IMM-based server, it attempts to connect and automatically configure the LAN over a USB interface, if it detects that this interface is not configured. The ASU also provides a level of automatic and default settings. You have the option of specifying that the automatic configuration process is skipped, if you have manually configured the IMM LAN over a USB interface. We recommend that the ASU configure the LAN over a USB interface.

See the *User's Guide for the Advanced Settings Utility* at the following website for more details:

http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-ASU

Tip: After you use the ASU to change settings, you must reset the IMM before you flash new firmware; otherwise, the changes to the settings might be lost. To reset the IMM, use the following ASU command:

```
asu rebootimm
```

Use the following procedure to download, install, and connect to the IMM using a Windows operating system:

1. Create a directory named ASU.
2. Download the ASU Tool for your operating system (32-bit or 64-bit) at the following website and save it in the ASU directory:

   `http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-ASU`

3. Unpack the utility:

   – For Windows, double-click **ibm_utl_asu_asut*XXX*_windows_*yyyy*.exe**, which decompresses the following files into the ASU folder, as shown in Figure 9-46.

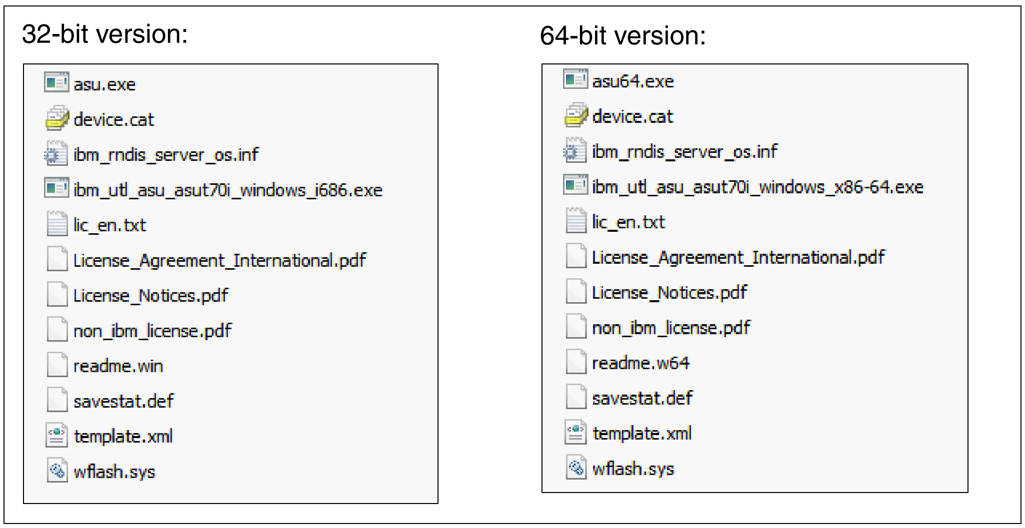


Figure 9-46 Files installed in the ASU folder

   – For Linux, open a terminal session and run the following command from the ASU directory:

```
tar zxf ibm_utl_asu_asutXXX_linux_yyyy.tgz
```

   Figure 9-47 on page 497 shows how the files are installed.

64 bit

```
-rwxrwxr-x 1 root root 3.4M Oct 29 10:28 asu64
-rwxrw-r-- 1 root root  53K Oct 29 10:28 cdc_interface.sh
-r-xr-xr-x 1 ibm  ibm   16M Nov 17 11:35 ibm_utl_asu_asut70i_linux_x86-64.tgz
-rwxrw-r-- 1 ibm  ibm  2.3M Oct 29 10:21 License_Agreement_International.pdf
-rwxrw-r-- 1 ibm  ibm  4.4M Oct 29 10:21 License_Notices.pdf
-rwxrw-r-- 1 ibm  ibm   60K Oct 29 10:21 lic_en.txt
-rwxrw-r-- 1 ibm  ibm  5.8M Oct 29 10:21 non_ibm_license.pdf
drwxr-xr-x 2 root root 4.0K Nov 22 15:47 rdcli64
-rwxrw-r-- 1 ibm  ibm   16K Oct 29 10:22 readme.l64
-rwxrw-r-- 1 root root 1.4K Oct 29 10:28 savestat.def
-rwxrw-r-- 1 ibm  ibm  4.2K Oct 29 10:20 template.xml
[root@localhost ASU]#
```

32 bit

```
-rwxrwxr-x 1 root root 3.2M Oct 29 10:51 asu
-rwxrw-r-- 1 root root  53K Oct 29 10:45 cdc_interface.sh
-r-xr-xr-x 1 ibm  ibm   16M Nov 17 11:34 ibm_utl_asu_asut70i_linux_i686.tgz
-rwxrw-r-- 1 ibm  ibm  2.3M Oct 29 10:21 License_Agreement_International.pdf
-rwxrw-r-- 1 ibm  ibm  4.4M Oct 29 10:21 License_Notices.pdf
-rwxrw-r-- 1 ibm  ibm   60K Oct 29 10:21 lic_en.txt
-rwxrw-r-- 1 ibm  ibm  5.8M Oct 29 10:21 non_ibm_license.pdf
drwxr-xr-x 2 root root 4.0K Nov 22 15:55 rdcli32
-rwxrw-r-- 1 ibm  ibm   16K Oct 29 10:22 readme.lin
-rwxrw-r-- 1 root root 1.4K Oct 29 10:45 savestat.def
-rwxrw-r-- 1 ibm  ibm  4.2K Oct 29 10:20 template.xml
[root@localhost ASU]#
```

Figure 9-47 Files installed in the ASU folder

4. Run a command, such as the **asu show** command, either in-band or out-of-band, using the commands that are listed in Table 9-4. This command confirms that the connection and utility work.

Table 9-4 Steps to run ASU both in-band and out-of-band

In-band	Out-of-band
1. Run the following command: - For 32 bit: **asu show** - For 64 bit: **asu64 show**	1. Ping the IMM to ensure that you have a network connection to the IMM. The default IP address is 192.168.70.125. Note: You might need to change your IP address. 2. Run the following command: - For 32 bit: **asu show --host IPADRESS** - For 64 bit: **asu64 show --host IPADRESS**

Example 9-1 shows the output from running the **asu show** command in-band.

Example 9-1 Output from the asu show command in-band

```
IBM Advanced Settings Utility version 3.61.70I
Licensed Materials - Property of IBM
(C) Copyright IBM Corp. 2007-2010 All Rights Reserved
Successfully discovered the IMM via SLP.
Discovered IMM at IP address 169.254.95.118
Connected to IMM at IP address 169.254.95.118
IMM.SSH_SERVER_KEY=Installed
```

Default IMM user and password parameters: When you change the default user and password of IMM, you must specify the `--user` and `--password` parameters, for example:

`asu show --user USER --password PASSWORD`

9.7.2 Common problems

This section describes the problems that you might encounter running ASU. The following reasons for errors are the most common:

- Incorrect password or user
- Firewall issues
- Insufficient rights for the operating system

Link error

Example 9-2 shows a connection link error that might occur.

Example 9-2 Connection link error

```
IBM Advanced Settings Utility version 3.61.70I
Licensed Materials - Property of IBM
(C) Copyright IBM Corp. 2007-2010 All Rights Reserved
Connection link error.
```

To resolve this error, try the following steps:

- Ensure that your firewall allows ASU.
- Check if you have entered the correct IMM IP address in the command.
- Check if you can ping the IMM IP address.
- Restart the IMM.

Wrong user and password error

Example 9-3 shows a password error that might occur.

Example 9-3 Wrong user and password error

```
IBM Advanced Settings Utility version 3.61.70I
Licensed Materials - Property of IBM
(C) Copyright IBM Corp. 2007-2010 All Rights Reserved
Unable to validate userid/password on IMM.
```

To resolve this error, ensure that you have entered the correct user ID and password for the IMM. Remember that both the user ID and password are case-sensitive.

Permission error

Example 9-4 shows a permission error that might occur in Windows.

Example 9-4 User rights in Windows error

```
IBM Advanced Settings Utility version 3.61.70I
Licensed Materials - Property of IBM
(C) Copyright IBM Corp. 2007-2010 All Rights Reserved
IPMI command error.  Please check your IPMI driver and IBM
mapping layer installation.
```

Example 9-5 on page 499 shows a permission error that might occur in Linux.

Example 9-5 User rights in Linux error

```
IBM Advanced Settings Utility version 3.61.70I
Licensed Materials - Property of IBM
(C) Copyright IBM Corp. 2007-2010 All Rights Reserved
User authority level is not sufficient. You must invoke ASU over the network (--host --user
--password)
```

To resolve this error, try the following steps:

- Ensure that the user has administrator or root equivalent operating system privileges.
- When you use ASU to connect remotely to the system (out-of-band), the user must have adequate IMM rights (Supervisor).

Firewall error

Example 9-6 shows a firewall error that might occur.

Example 9-6 Firewall does not allow ASU - Linux

```
IBM Advanced Settings Utility version 3.61.70I
Licensed Materials - Property of IBM
(C) Copyright IBM Corp. 2007-2010 All Rights Reserved
SLP request fail.
Try to configure usb-lan using DHCP.
Unable to connect to IMM via LAN : Could not configure in band network.
..............................
Failed to connect to local IMM using LAN IP  and KCS
```

To resolve this error, disable the firewall or allow ASU.

9.7.3 Command examples

In this section, we provide a brief overview for the most commonly used ASU commands. See the *Advanced Settings Utility User's Guide* at the following website for all the ASU commands:

http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-ASU

Show all settings

On the command line, enter **asu show**. See Example 9-7.

Example 9-7 Output

```
IMM.SSH_SERVER_KEY=Installed
IMM.SSL_HTTPS_SERVER_CERT=Private Key and Cert/CSR not available.
IMM.SSL_HTTPS_SERVER_CSR=Private Key and Cert/CSR not available.
IMM.SSL_LDAP_CLIENT_CERT=Private Key and Cert/CSR not available.
IMM.SSL_LDAP_CLIENT_CSR=Private Key and Cert/CSR not available.
IMM.SSL_SERVER_DIRECTOR_CERT=Private Key and Cert/CSR not available.
IMM.SSL_SERVER_DIRECTOR_CSR=Private Key and Cert/CSR not available.
IMM.SSL_CLIENT_TRUSTED_CERT1=Not-Installed
IMM.SSL_CLIENT_TRUSTED_CERT2=Not-Installed
IMM.SSL_CLIENT_TRUSTED_CERT3=Not-Installed
IMM.PowerRestorePolicy=Restore
IMM.ThermalModePolicy=Normal
```

Show all UEFI settings

On the command line, enter **`asu show uefi`**. See Example 9-8.

Example 9-8 Output

```
uEFI.OperatingMode=Custom Mode
uEFI.QuietBoot=Enable
uEFI.TurboModeEnable=Enable
uEFI.TurboBoost=Power Optimized
uEFI.ProcessorEistEnable=Enable
uEFI.ProcessorCcxEnable=Disable
uEFI.ProcessorC1eEnable=Enable
uEFI.HyperThreading=Enable
uEFI.EnableCoresInSbsp=All
uEFI.ExecuteDisableBit=Enable
uEFI.ProcessorVmxEnable=Enable
uEFI.ProcessorDataPrefetch=Enable
```

Show all IMM settings

On the command line, enter **`asu show imm`**. See Example 9-9.

Example 9-9 Output

```
IMM.SSH_SERVER_KEY=Installed
IMM.SSL_HTTPS_SERVER_CERT=Private Key and Cert/CSR not available.
IMM.SSL_HTTPS_SERVER_CSR=Private Key and Cert/CSR not available.
IMM.SSL_LDAP_CLIENT_CERT=Private Key and Cert/CSR not available.
IMM.SSL_LDAP_CLIENT_CSR=Private Key and Cert/CSR not available.
```

Set turbomode to enable

On the command line, enter **`asu set uefi.turbomodeenable enable`**. See Example 9-10.

Example 9-10 Output

```
uEFI.TurboModeEnable=Enable
Waiting for command completion status.
Command completed successfully.
```

Set QuickPath Interconnect (QPI) speed to maximum performance

On the command line, enter **`asu set uefi.qpispeed "Max Performance"`**. See Example 9-11.

Example 9-11 Output

```
uEFI.QPISpeed=Max Performance
Waiting for command completion status.
Command completed successfully.
```

Set memory speed to maximum performance

On the command line, enter **`asu set uefi.ddrspeed "Max Performance"`**. See Example 9-12.

Example 9-12 Output

```
uEFI.DDRspeed=Max Performance
Waiting for command completion status.
Command completed successfully.
```

Enable Hyper-Threading

On the command line, enter **`asu set uefi.hyperthreading enable`**. See Example 9-13.

Example 9-13 Output

```
uEFI.HyperThreading=Enable
Waiting for command completion status.
Command completed successfully.
```

Disable Energy Manager

On the command line, enter **`asu set uefi.energymanager disable`**. See Example 9-14.

Example 9-14 Output

```
uEFI.EnergyManager=Disable
Waiting for command completion status.
Command completed successfully.
```

9.8 IBM ServerGuide

IBM ServerGuide is an installation assistant for Windows installations that simplifies the process of installing and configuring IBM System x and BladeCenter servers. The wizard guides you through the setup, configuration, and operating system installation.

Table 9-5 shows the minimum ServerGuide versions that are required for the eX5 servers.

Table 9-5 Minimum required ServerGuide versions

Machine	Version
System x3850 X5/x3950 X5 (7145, 7146)	ServerGuide 8.22
System x3690 X5 (7148, 7149)	ServerGuide 8.3
System x3850 X5/x3950 X5 2-node (7145, 7146)	ServerGuide 8.3
BladeCenter HX5 (7872, 1909)	ServerGuide 8.3

Tip: If possible, use the latest version of IBM ServerGuide.

ServerGuide can accelerate and simplify the installation of eX5 servers in the following ways:

- Assists with installing Windows-based operating systems and provides updated device drivers that are based on the detected hardware
- Reduces rebooting requirements during hardware configuration and Windows operating system installation, allowing you to get your eX5 server up and running sooner
- Provides a consistent server installation using IBM best practices for installing and configuring an eX5 server
- Provides access to additional firmware and device drivers that might not be applied at installation time, such as adapter cards that are added to the system later

ServerGuide deploys the OS image to the first device in the boot order sequence. Best practices dictate that you have one device that is available for the ServerGuide installation process. If you boot from SAN, make sure that you have only one path to the device because ServerGuide has no multipath support. See "Booting from SAN" on page 295 for more details.

After the ServerGuide installation procedure, you can attach external storage or activate additional paths to the disk. For installation instructions about how to attach external storage or multipath drivers, see the respective User Guide.

The following procedure describes how to install Windows Server 2008 R2 with ServerGuide. The method to install LInux is similar. Follow these steps:

1. Download the latest version of ServerGuide from the following website.

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

2. Burn a CD with the ISO image or mount the image through the IMM.
3. Boot the server from ServerGuide.
4. The progress bar appears, as shown in Figure 9-48.

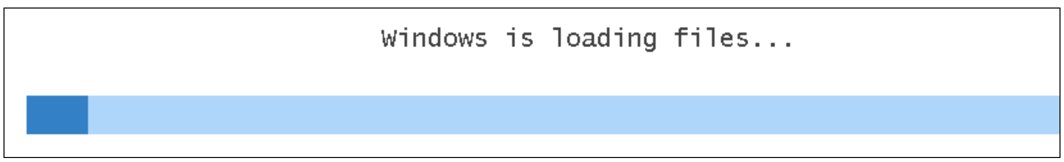


Figure 9-48 ServerGuide progress bar

5. After the files load, the Start window opens. Choose a language to continue.
6. Select your preferred keyboard layout and click **Next**.
7. Accept the license agreement and click **Next**.
8. The Welcome window opens and provides information about ServerGuide, which systems are supported, a readme file, and copyright and trademark information. Click **Next**.
9. Select the operating system that you want to install and click **Next**, as shown in Figure 9-49.

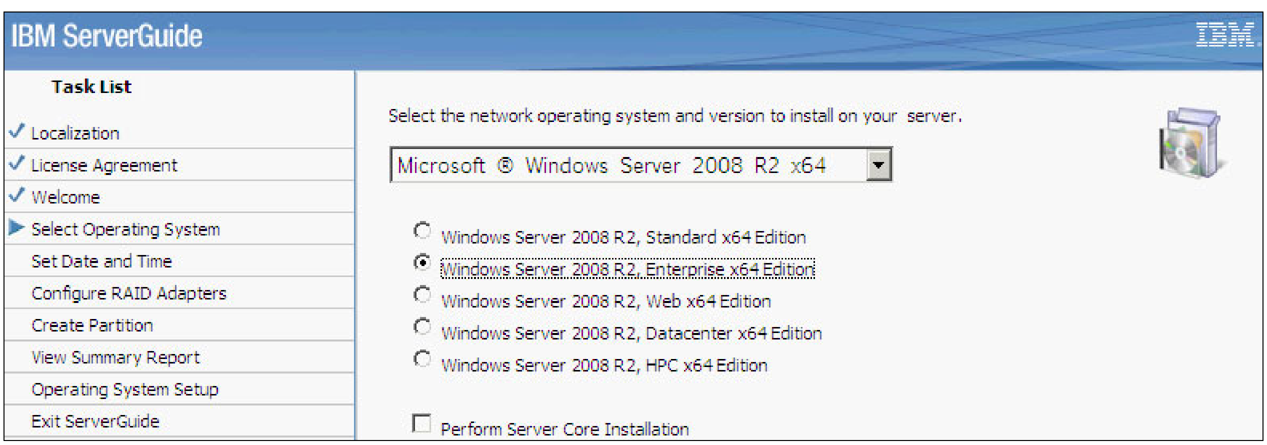


Figure 9-49 Selecting the operating system

10. Enter the current date and time and click **Next**, as shown in Figure 9-50 on page 503.

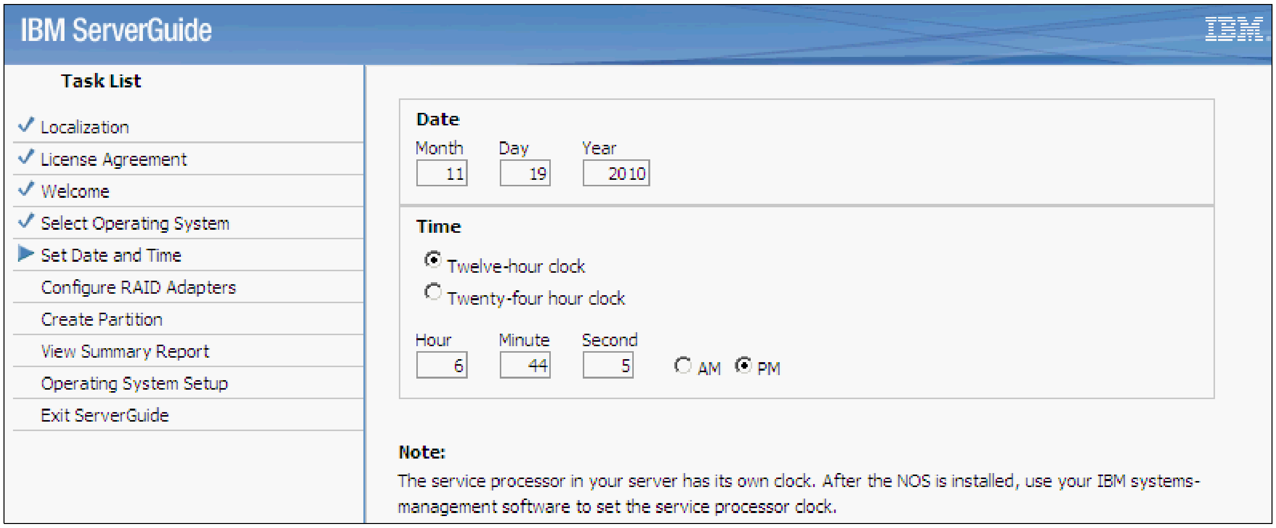


Figure 9-50 Date and time settings

11.Create a RAID configuration. If you previously created a RAID adapter, ServerGuide detects this configuration and displays it under Keep Current Adapter Configuration. Select a RAID configuration and click **Next,** as shown in Figure 9-51.

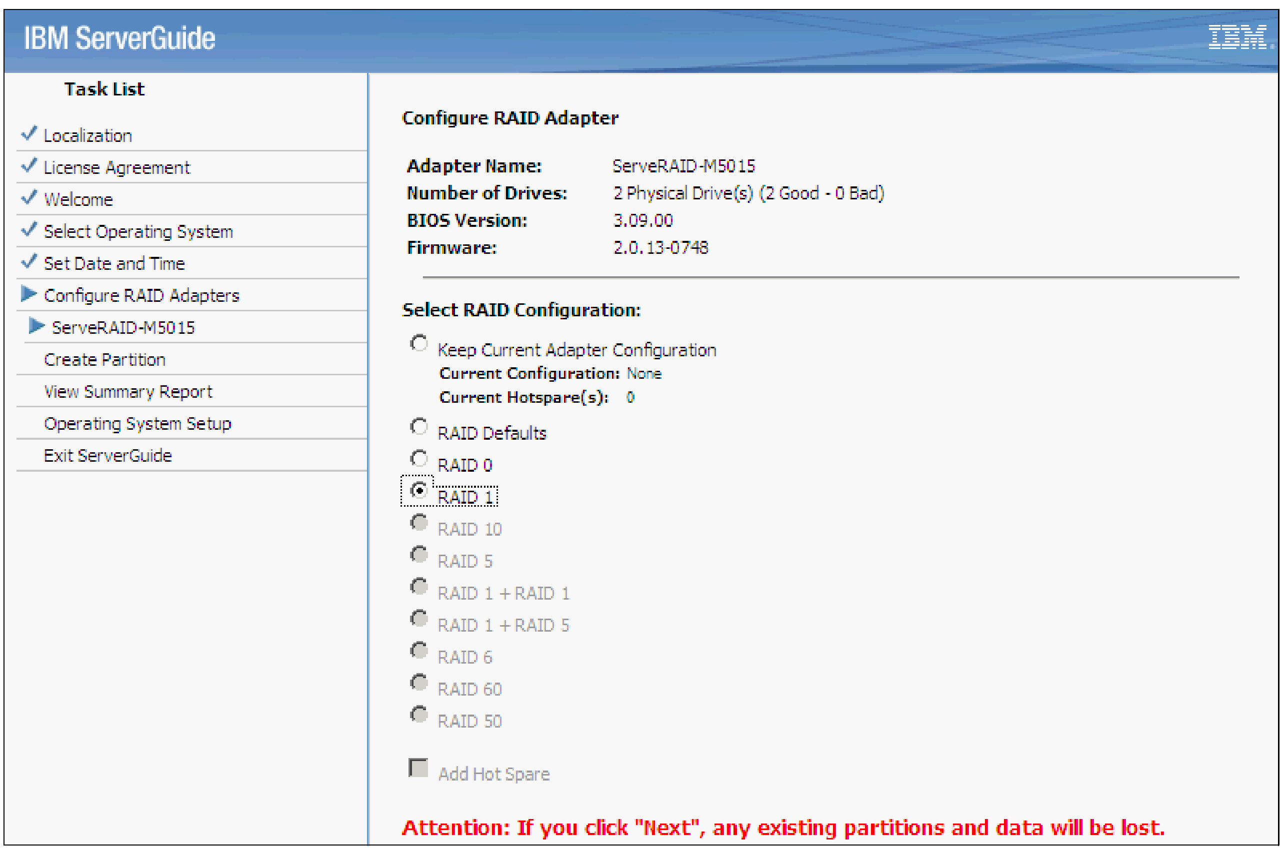


Figure 9-51 RAID configuration panel

12.A Confirmation window opens indicating that RAID configuration is complete (Figure 9-52 on page 504). Click **Next**.

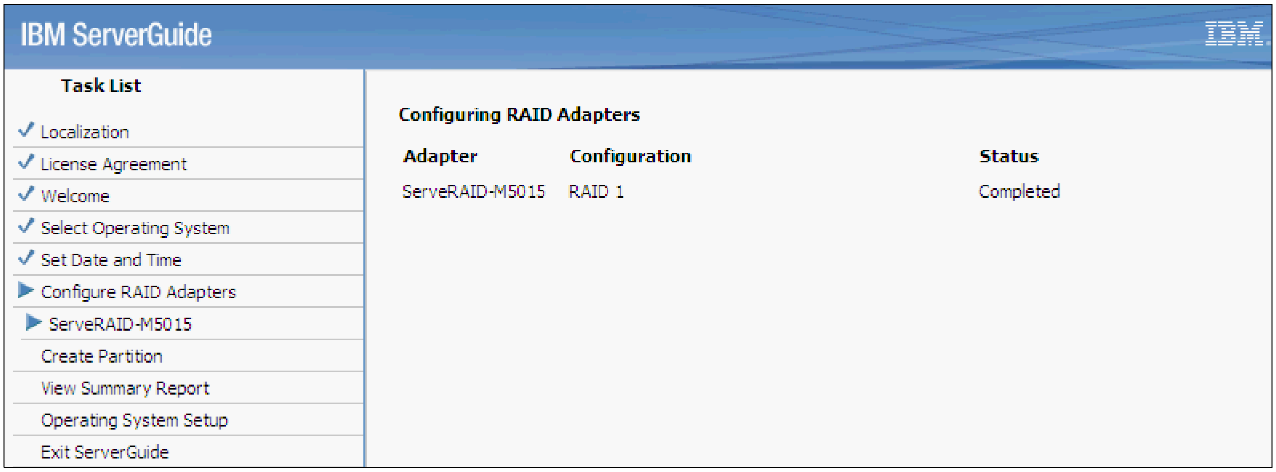


Figure 9-52 RAID configuration panel

13. Restart the server. After the server restarts, the server will restart again to complete the ServerGuide setup, as shown in Figure 9-53. If you previously created a RAID configuration and kept the configuration, restarting the server is not necessary. Click **Next**.

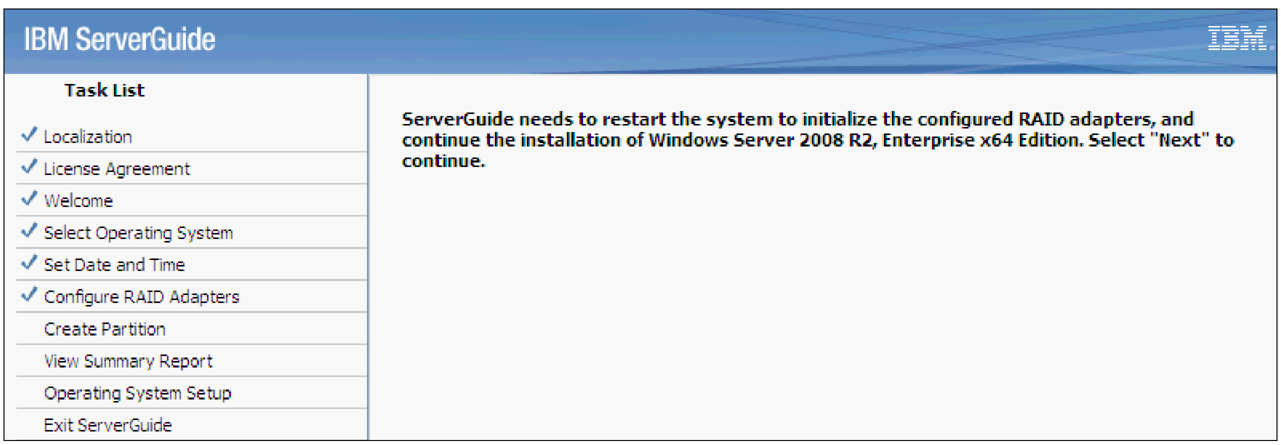


Figure 9-53 Information about the restart

14. After the server has restarted, you must create and format a partition. Choose your selection and click **Next** to start the process, as shown in Figure 9-54 on page 505.

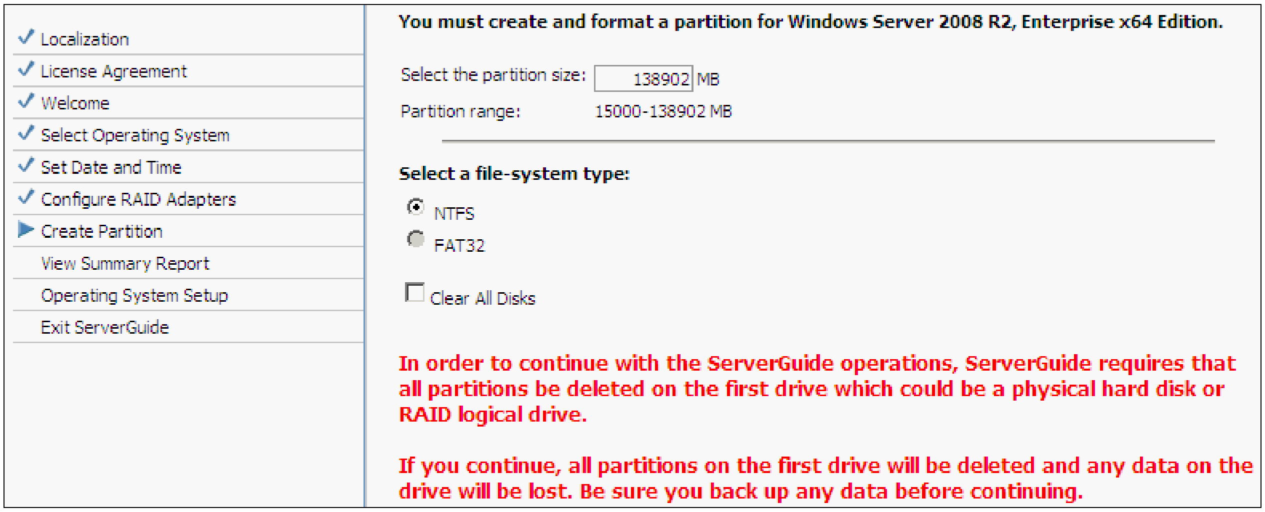


Figure 9-54 Selection for format and partition

15. When the process has completed, click **Next**, as shown in Figure 9-55.

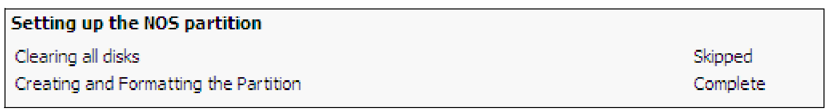


Figure 9-55 Confirmation about completing the creation and formatting of the partition

16. Review the configuration, as shown in Figure 9-56. Click **Next**.

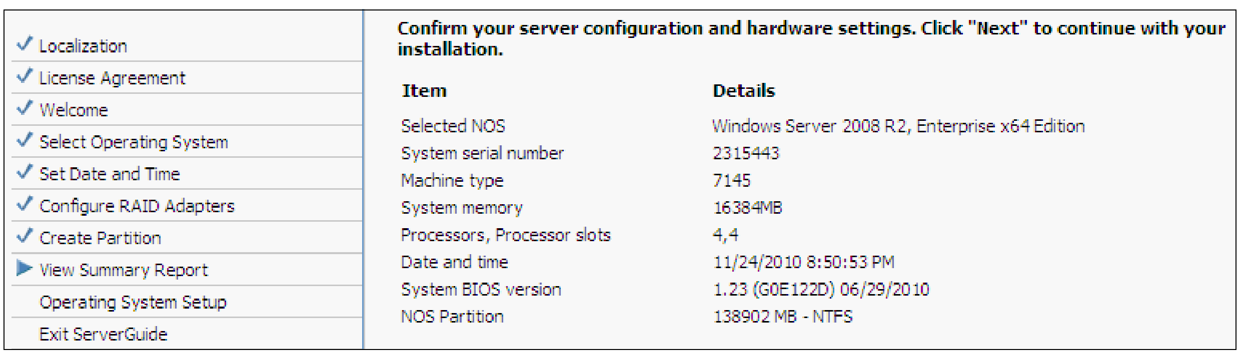


Figure 9-56 Summary of selections

17. ServerGuide copies the necessary files to the disk in preparation for the operating system installation, as shown in Figure 9-57 on page 506.

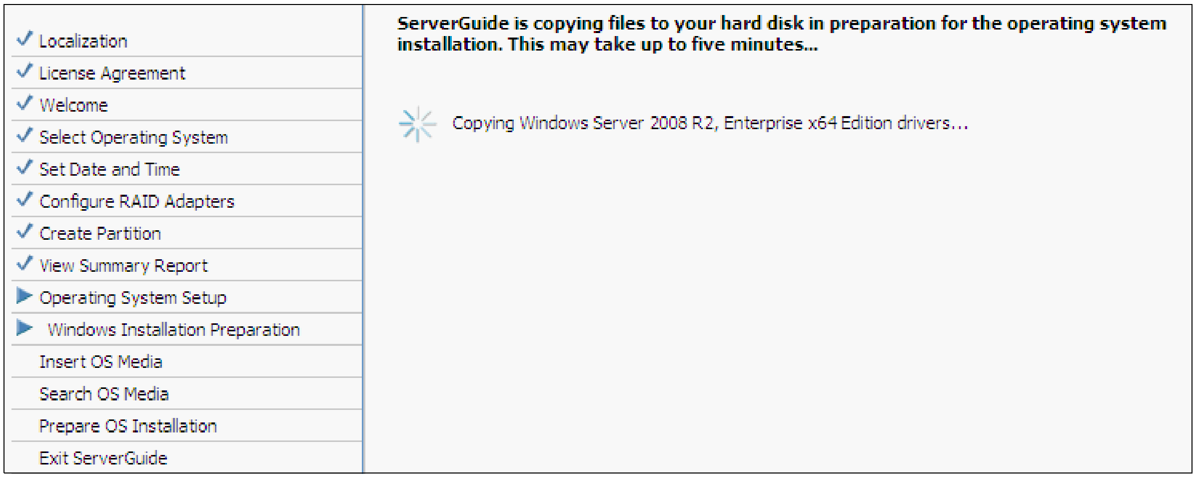


Figure 9-57 ServerGuide copying files

18. When the process is finished, click **Next**.

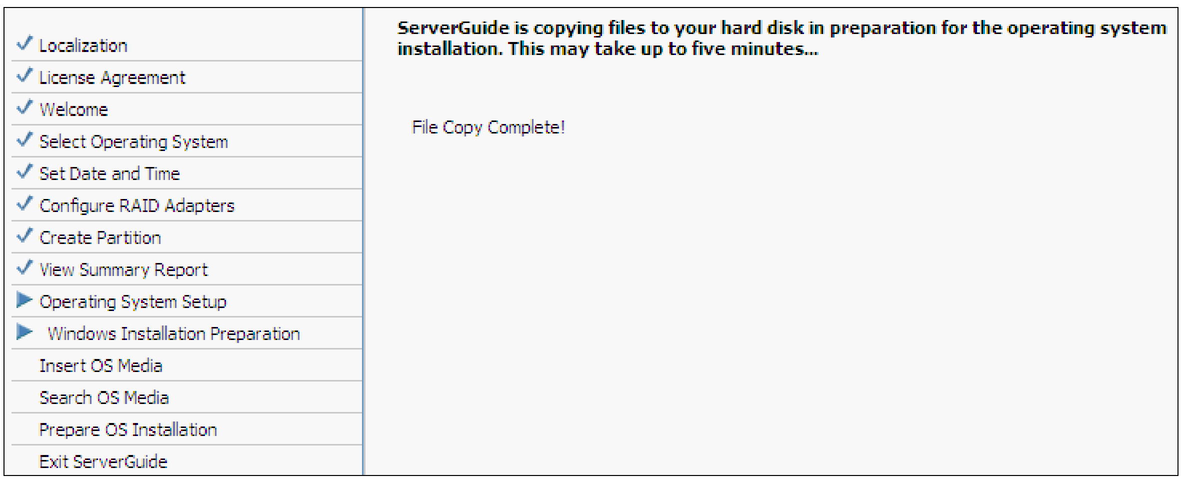


Figure 9-58 Confirmation about file copy success

19. Insert the operating system installation DVD and click **Next**, as shown in Figure 9-59 on page 507. ServerGuide searches for the disc.

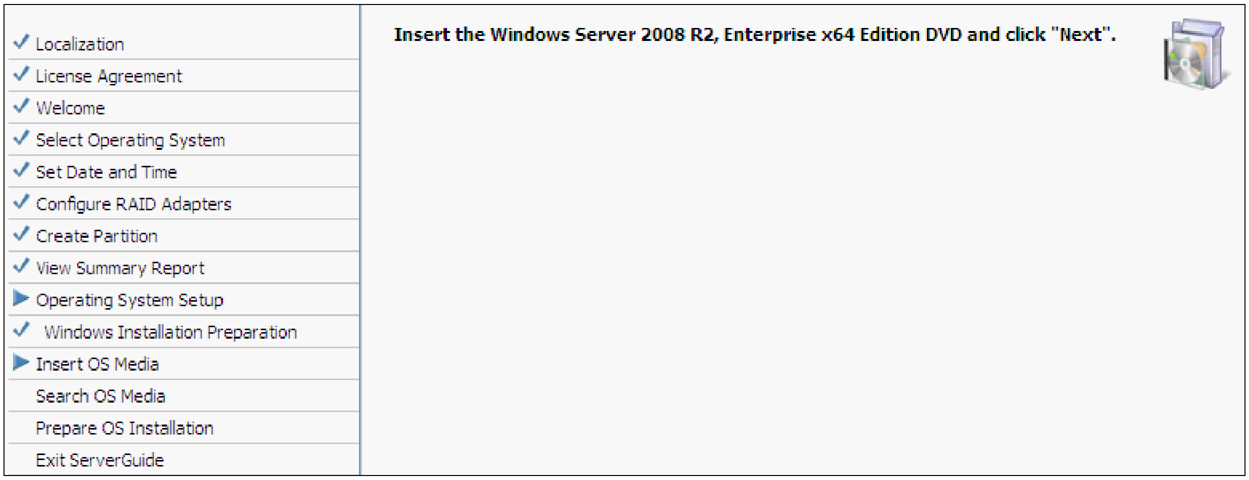


Figure 9-59 Prompt to insert the install medium

20. When ServerGuide successfully finds the installation medium, click **Next**, as shown in Figure 9-60.

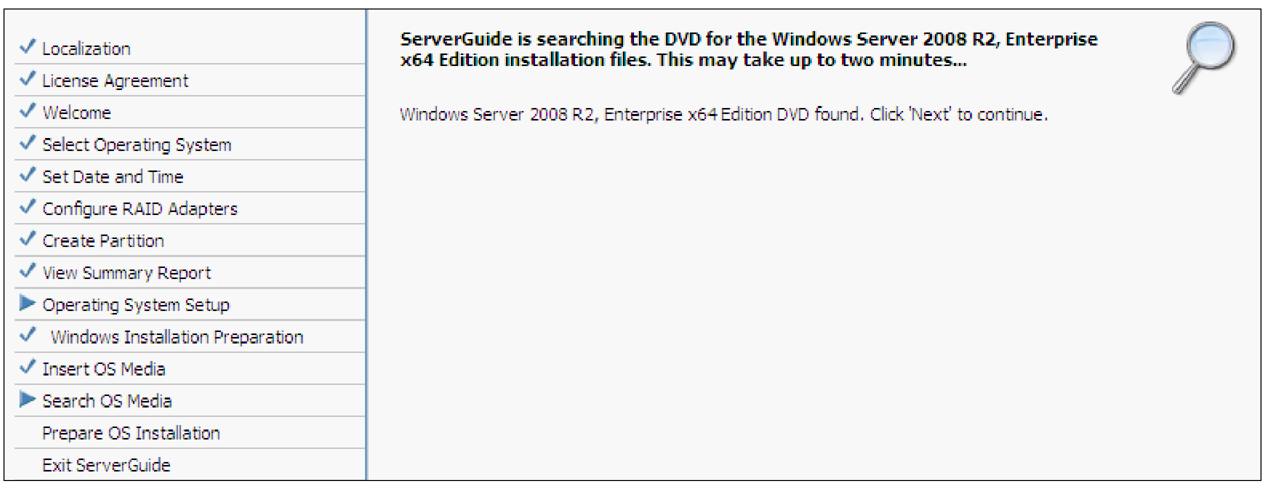


Figure 9-60 Confirmation about finding installation medium

21. The Windows setup installation procedure starts. Follow the installation procedure to complete the installation.

9.9 IBM ServerGuide Scripting Toolkit

You can use the IBM ServerGuide Scripting Toolkit to create deployable images using a collection of system-configuration tools and installation scripts. There are versions of the ServerGuide Scripting Toolkit for the Windows Preinstallation Environment (PE) and Linux platforms.

The ServerGuide Scripting Toolkit enables you to tailor and build custom hardware deployment solutions. It provides hardware configuration utilities and operating system (OS) installation examples for IBM System x and BladeCenter x86-based hardware.

If used with IBM ServerGuide and IBM UpdateXpress, the ServerGuide Scripting Toolkit provides a total solution for deploying IBM System x and BladeCenter x86-based hardware in an unattended mode.

The ServerGuide Scripting Toolkit enables you to create a bootable CD, DVD, or USB key that supports the following tasks and components:

- Network and mass storage devices
- Policy-based RAID configuration
- Configuration of system settings using Advanced Settings Utility (ASU)
- Configuration of fibre host bus adapters (HBAs)
- Local self-contained DVD deployment scenarios
- Local CD/DVD and network share-based deployment scenarios
- Remote Supervisor Adapter (RSA) II, Integrated Management Module (IMM), and BladeCenter Management Module (MM)/Advanced Management Module (AMM) remote disk scenarios
- UpdateXpress System Packs installation integrated with scripted network operating system (NOS) deployment
- IBM Director Agent installation, integrated with scripted NOS deployment

The ServerGuide Scripting Toolkit, Windows Edition supports the following versions of IBM Systems Director Agent:

- Common Agent 6.1 or later
- Core Services 5.20.31 or later
- Director Agent 5.1 or later

The Windows version of the ServerGuide Scripting Toolkit enables automated operating system support for the following Windows operating systems:

- Windows Server 2003, Standard, Enterprise, and Web Editions
- Windows Server 2003 R2, Standard and Enterprise Editions
- Windows Server 2003, Standard and Enterprise x64 Editions
- Windows Server 2003 R2, Standard and Enterprise x64 Editions
- Windows Server 2008, Standard, Enterprise, Datacenter, and Web Editions
- Windows Server 2008 x64, Standard, Enterprise, Datacenter, and Web Editions
- Windows Server 2008, Standard, Enterprise, and Datacenter Editions without Hyper-V
- Windows Server 2008 x64, Standard, Enterprise, and Datacenter without Hyper-V
- Windows Server 2008 R2 x64, Standard, Enterprise, Datacenter, and Web Editions

The Linux version of the ServerGuide Scripting Toolkit enables automated operating system support for the following operating systems:

- SUSE Linux Enterprise Server 9 32 bit SP4
- SUSE Linux Enterprise Server 9 x64 SP4
- SUSE Linux Enterprise Server 10 32 bit SP1/SP2/SP3
- SUSE Linux Enterprise Server 10 x64 SP1/SP2/SP3
- SUSE Linux Enterprise Server 11 32 bit Base/SP1
- SUSE Linux Enterprise Server 11 x64 Base/SP1
- Red Hat Enterprise Linux 4 AS/ES 32 bit U6/U7/U8
- Red Hat Enterprise Linux 4 AS/ES x64 U6/U7/U8
- Red Hat Enterprise Linux 5 32 bit U1/U2/U3/U4/U5
- Red Hat Enterprise Linux 5 x64 U1/U2/U3/U4/U5
- VMware ESX Server 3.5 U4/U5
- VMware ESX Server 4.0/4.0u1/4.0u2/4.1

To download the Scripting Toolkit or the *IBM ServerGuide Scripting Toolkit User's Reference*, select the IBM ServerGuide Scripting Toolkit website:

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

9.10 Firmware update tools and methods

Multiple methods exist for performing firmware updates. The preferred method to perform firmware updates is to use one of these tools:

- UpdateXpress System Pack Installer (UXSPI)
- Bootable Media Creator (BoMC)

These tools are able to perform these functions:

- Display an inventory of installed firmware and drivers
- Download firmware and drivers from http://www.ibm.com
- Download a UXSP from http://www.ibm.com
- Update all of the firmware and drivers in your system, including RAID, hard disk drives (HDDs), network interchange cards (NICs), and Fibre Channel devices
- Apply updates in the correct order to completely update a system with the fewest reboots
- Create a bootable CD/DVD/USB key/Preboot eXecution Environment (PXE) image to perform firmware updates (BoMC)

We cover these two tools in the following sections.

This section describes important points for the firmware update. You can also read the readme file in each update package and all recommendations and prerequisites for updating in the *Best Practice Firmware Update Guide* at the following website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=MIGR-5082923

> **Tip:** On systems that contain an FPGA, a power off/on cycle is required to activate the new firmware.

9.10.1 Configuring UEFI

To enable flashing, you must enable the LAN over USB interface first in UEFI:

For the x3850 X5 and x3690 X5 servers, in the UEFI menu, select **System Settings** → **Integrated Management Module** → **Commands on USB Interface Preference** and enable **Commands on USB interface**. See Figure 9-61 on page 510.

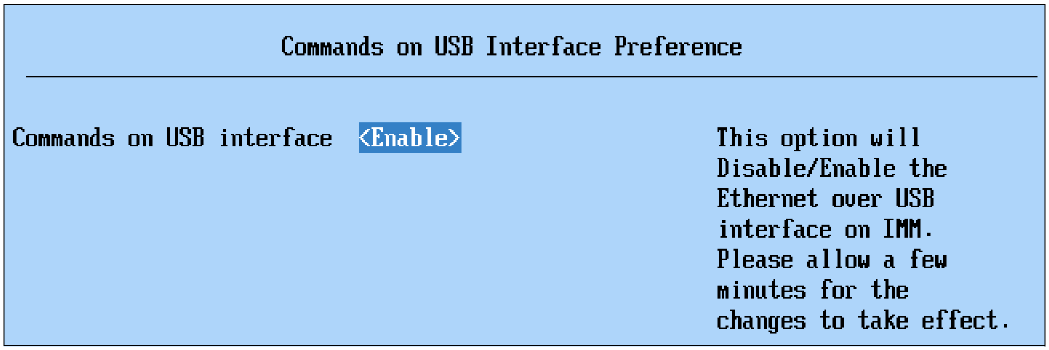


Figure 9-61 UEFI Setting for LAN over USB

For the HX5, from the AMM web interface, click in the left panel **Blade Tasks** → **Configuration** → **Advanced Blade Policy Settings** and ensure that the server's Ethernet over USB interface is enabled. See Figure 9-62.

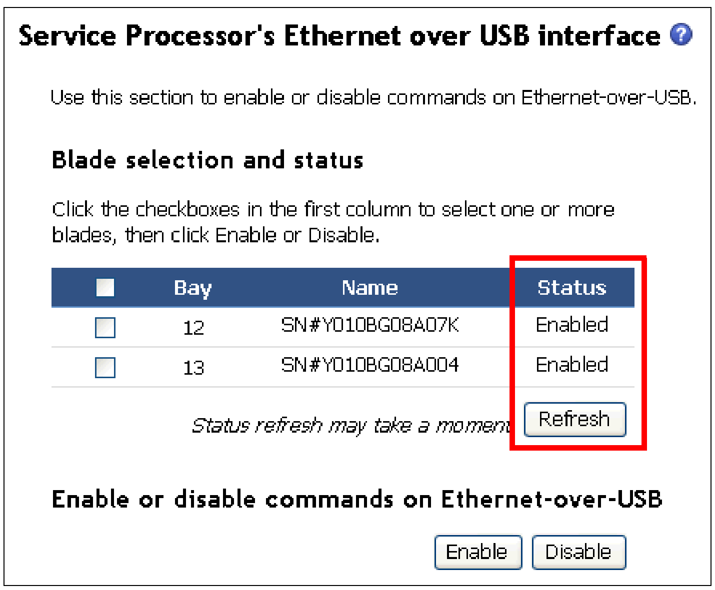


Figure 9-62 AMM setting for LAN over USB

9.10.2 Requirements for updating scalable systems

Before connecting two systems to each other with a QPI cable kit or a MAX5 to your system, you must ensure that the UEFI, IMM, FPGA, and DSA preboot firmware levels are the same.

Running various levels between both nodes can lead to unpredictable results. The preferred methods to perform firmware updates are to use the ToolsCenter UpdateXpress System Pack Installer (UXSPI) or Bootable Media Creator (BoMC). Both update methods ensure that the systems in a scalable complex are updated at the same time.

Updating using out-of-band through the IMM (x3850 X5 and x3690 X5) or AMM (for the HX5) is supported. You must ensure that the firmware for all systems has been updated successfully before rebooting the system.

The HX5 blade supports creating a scalable blade complex with the blade servers configured as two independent partitions. The online update utilities can perform the update only on the partition from which they are executed. You must update the firmware for each blade independently before rebooting either system.

With IMM firmware Version 1.15, a Dynamic Host Configuration Protocol (DHCP) server is included in the IMM. This DHCP assigns an IP address for the internal LAN over USB interface. Both LAN over USB interfaces in a scalable complex will get an IP address. You must enable DHCP on these interfaces in the operating system to enable firmware updating.

9.10.3 IBM Systems Director

You can also perform an update using IBM Systems Director, which we explain in “Updating the firmware of a complex using IBM Systems Director” on page 491. The update procedure for a 1-node or 2-node configuration is the same.

9.11 UpdateXpress System Pack Installer

The UpdateXpress System Pack Installer (UXSPI) gives you the ability to update the firmware and device drivers of the system under an operating system. You can deploy UpdateXpress System Packs (UXSPs) and the latest individual updates.

UXSPI uses the standard HTTP (port 80) and HTTPS (port 443) to get the updates from IBM. Your firewall must allow these ports. UXSPI is supported on Windows, Linux, and VMware operating systems. UXSPI is supported on both 32-bit and 64-bit operating systems.

The *IBM UpdateXpress System Pack Installer User’s Guide* provides a detailed list of supported operating systems at the following website:

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

Perform these steps to use UXSPI:

1. Start the UXSPI setup utility that you downloaded from the precious website.
2. Accept the license agreement.
3. The Welcome window opens, as shown in Figure 9-63 on page 512. Click **Next**.

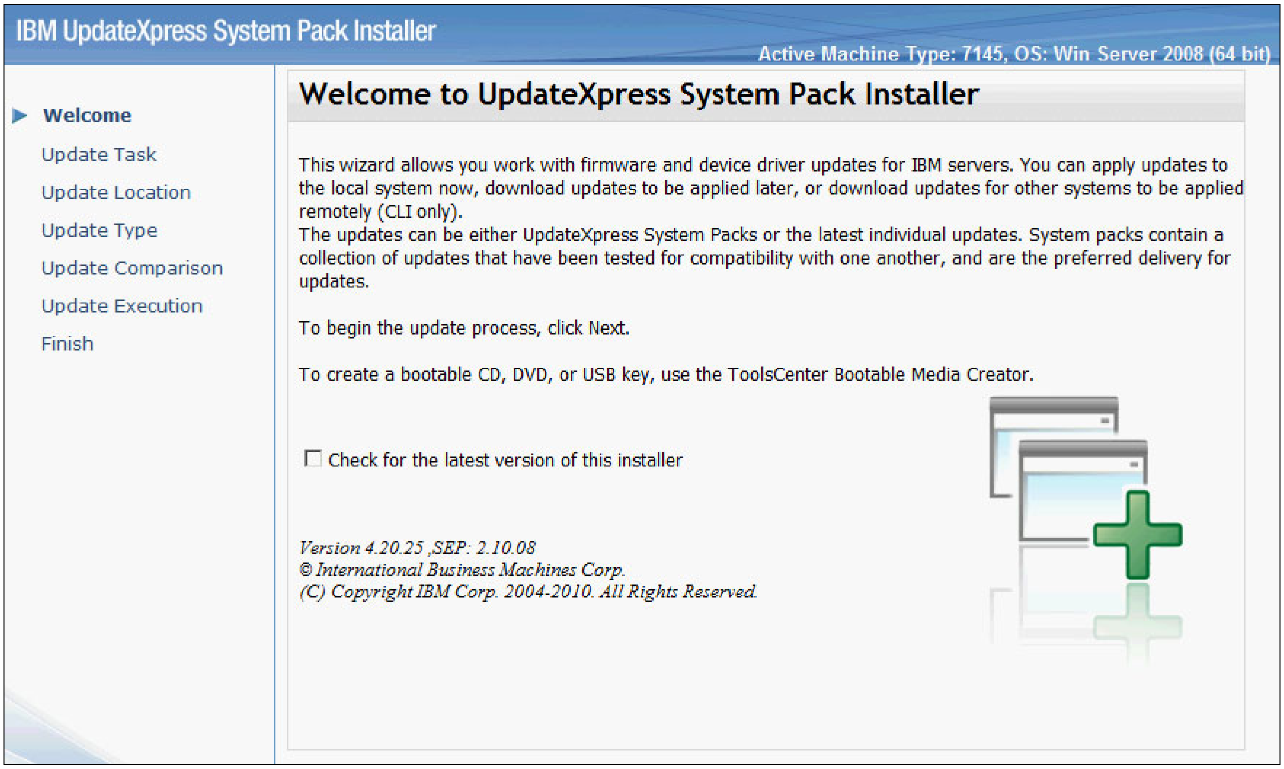


Figure 9-63 Welcome window

4. Accept the default **Update the local machine**, as shown in Figure 9-64. Click **Next**.

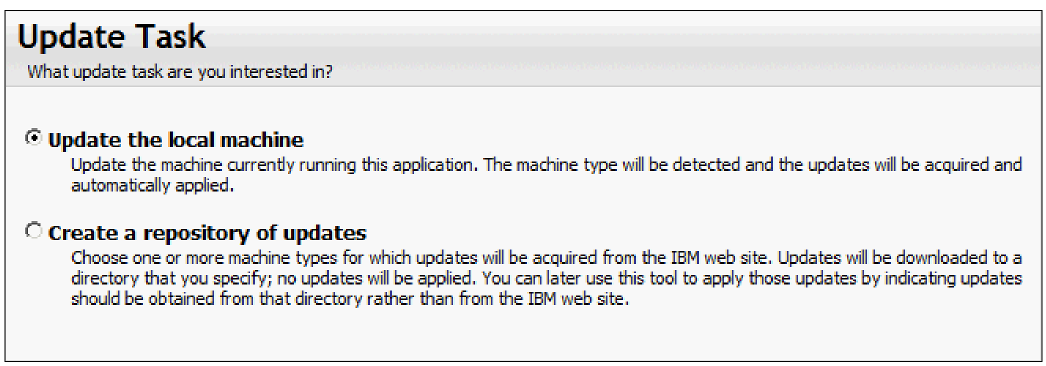


Figure 9-64 Selecting the update task

5. Accept the default **Check the IBM web site**, as shown in Figure 9-65. Click **Next**.

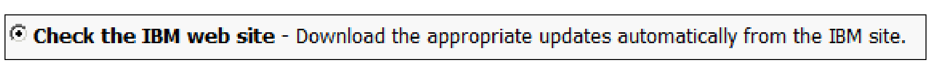


Figure 9-65 Checking the IBM website for updates

6. Select **Latest available individual updates**, as shown in Figure 9-66 on page 513. Click **Next**.

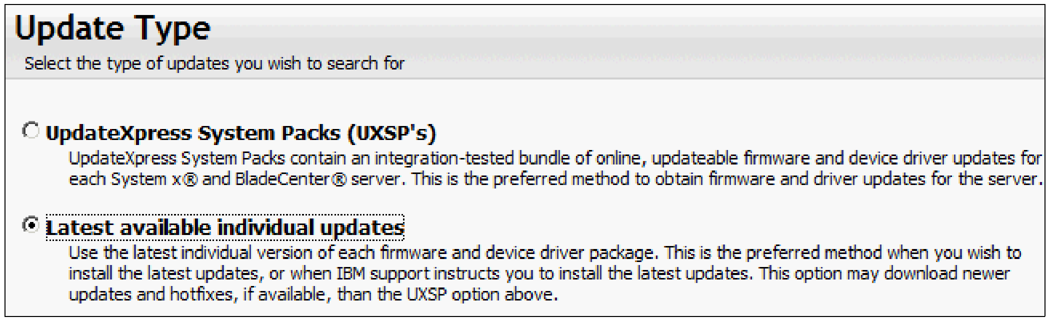


Figure 9-66 Selecting the type of updates

7. Select the directory in which you want to store the downloaded files, as shown in Figure 9-67. Click **Next**.

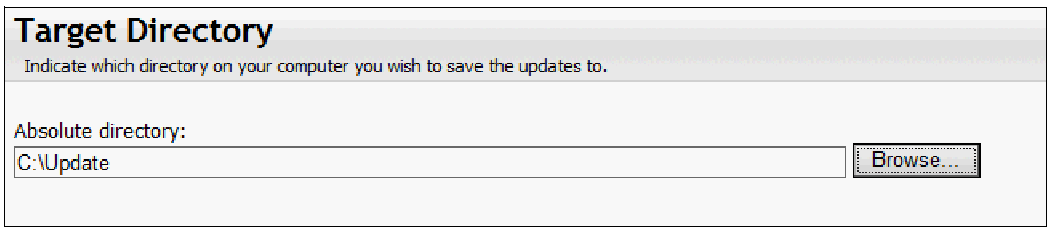


Figure 9-67 Selecting your target directory

8. Enter the settings for an HTTP proxy server, if necessary, or leave the check box unchecked, as shown in Figure 9-68. Click **Next**.

Figure 9-68 HTTP proxy settings

9. A message displays showing that the UXSPI acquired the possible updates for the machine, as shown in Figure 9-69. Click **Next**.

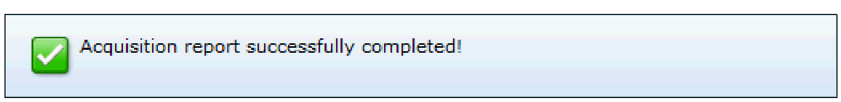


Figure 9-69 Successful completion of acquisition report

10. A message appears showing that the download has completed, as shown in Figure 9-70. Click **Next**.

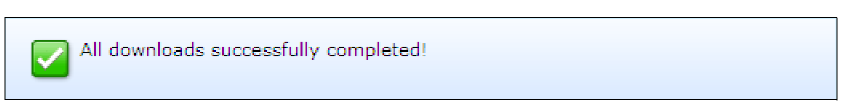


Figure 9-70 Download process completes successfully

11. A component overview shows the components that need updating. UXSPI selects, by default, the components to update. Accept these settings and click **Next**.

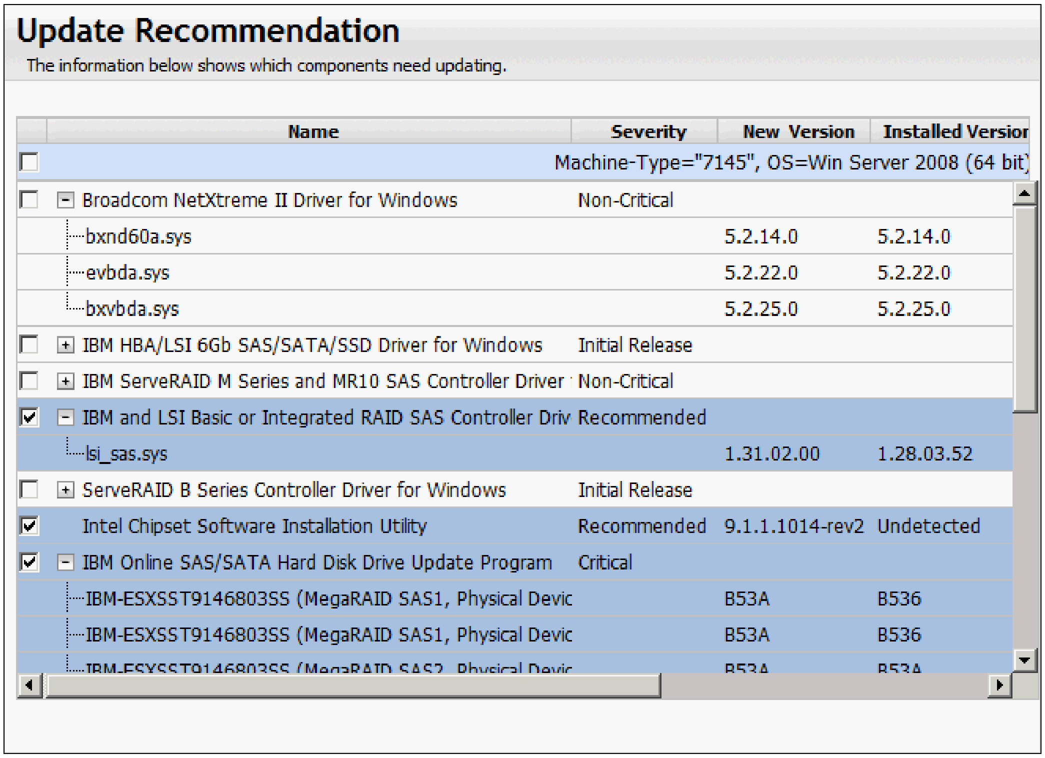


Figure 9-71 Overview of possible updates

12. When the update is finished, a message displays confirming the updates. Click **Next**.

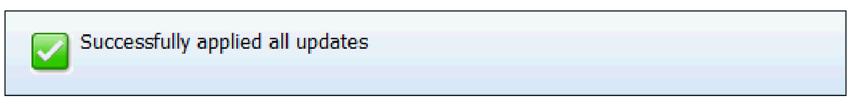


Figure 9-72 Updates are successful

13. Click **Finish** to close the UXSPI.

14. Restart the system to complete the update process.

9.12 Bootable Media Creator

The Bootable Media Creator (BoMC) provides a tool for creating a bootable image for supported media (CD, DVD, ISO image, USB flash drive, or PXE files) to update the system firmware. Because BoMC runs in its own boot environment, you cannot update drivers. BoMC has a graphical and command-line interface. One bootable media image can contain support for multiple systems. The tool uses standard HTTP (port 80) and HTTPS (port 443) to get the updates from IBM. Your firewall must allow these ports.

Supported operating systems

BoMC is supported on Windows, Linux, and VMware operating systems. BoMC supports 32-bit and 64-bit operating systems. The *IBM ToolsCenter Bootable Media Creator Installation and User's Guide* provides a detailed list of supported operating systems at the following website:

http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-BOMC

Creating an update media

Perform these steps to create an update media:

1. Create a folder named `BoMC`.
2. Download the latest version of BoMC from the following website and save it in the `BoMC` folder:

   `http://ibm.com/support/entry/portal/docdisplay?lndocid=TOOL-BOMC`

3. From a command line, enter the command name to start the BOMC. The command name depends on the operating system. Table 9-6 lists the name of the command for each supported operating system.

Table 9-6 Command for each supported operating system

Operating system	Command name
Windows	ibm_utl_bomc_v.r.m_windows_i386.exe
Red Hat Enterprise Linux 3.0	ibm_utl_bomc_v.r.m_rhel3_i386.bin
Red Hat Enterprise Linux 3.0 64-bit	ibm_utl_bomc_v.r.m_rhel3_x86-64.bin
Red Hat Enterprise Linux 4.0	ibm_utl_bomc_v.r.m_rhel4_i386.bin
Red Hat Enterprise Linux 4.0 64-bit	ibm_utl_bomc_v.r.m_rhel4_x86-64.bin
Red Hat Enterprise Linux 5.0	ibm_utl_bomc_v.r.m_rhel5_i386.bin
Red Hat Enterprise Linux 5.0 64-bit	ibm_utl_bomc_v.r.m_rhel5_x86-64.bin
SUSE Linux Enterprise Server 9	ibm_utl_bomc_v.r.m_sles9_i386.bin
SUSE Linux Enterprise Server 9 64-bit	ibm_utl_bomc_v.r.m_sles9_x86-64.bin
SUSE Linux Enterprise Server 10	ibm_utl_bomc_v.r.m_sles10_i386.bin
SUSE Linux Enterprise Server 10 64-bit	ibm_utl_bomc_v.r.m_sles10_x86-64.bin
SUSE Linux Enterprise Server 11	ibm_utl_bomc_v.r.m_sles11_i386.bin
SUSE Linux Enterprise Server 11 64-bit	ibm_utl_bomc_v.r.m_sles11_x86-64.bin

4. Accept the license agreement.
5. The Welcome window opens, as shown in Figure 9-73 on page 516. Click **Next**.

Figure 9-73 Welcome panel

6. Check **Updates** and click **Next**, as shown in Figure 9-74.

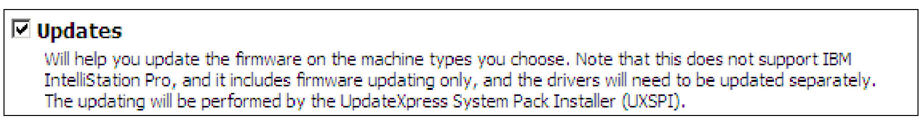


Figure 9-74 Selected update task

7. Select **Latest available individual updates** and click **Next**, as shown in Figure 9-75 on page 517.

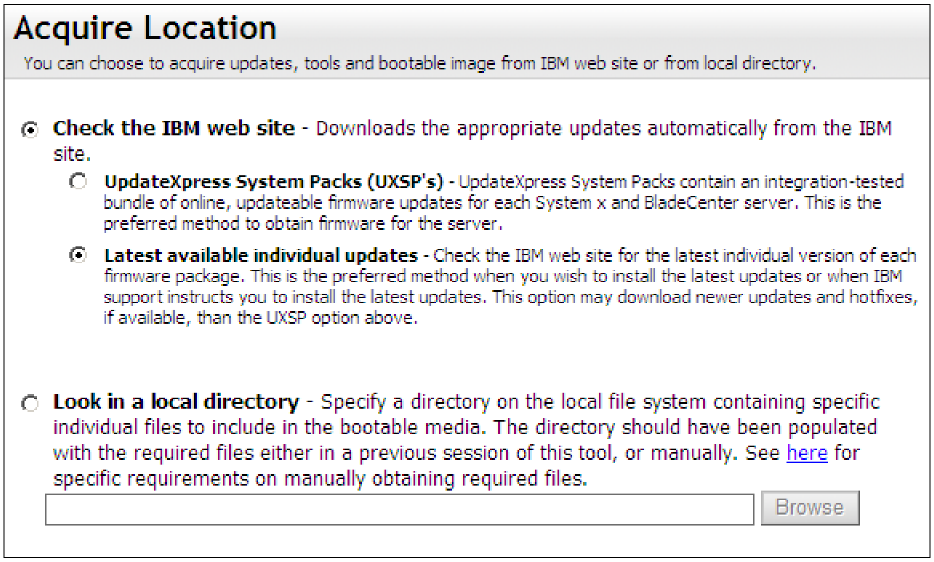


Figure 9-75 Select source of updates

8. Enter the settings for an HTTP proxy server, if necessary, or check **Do not use proxy**, as shown in Figure 9-76. Click **Next**.

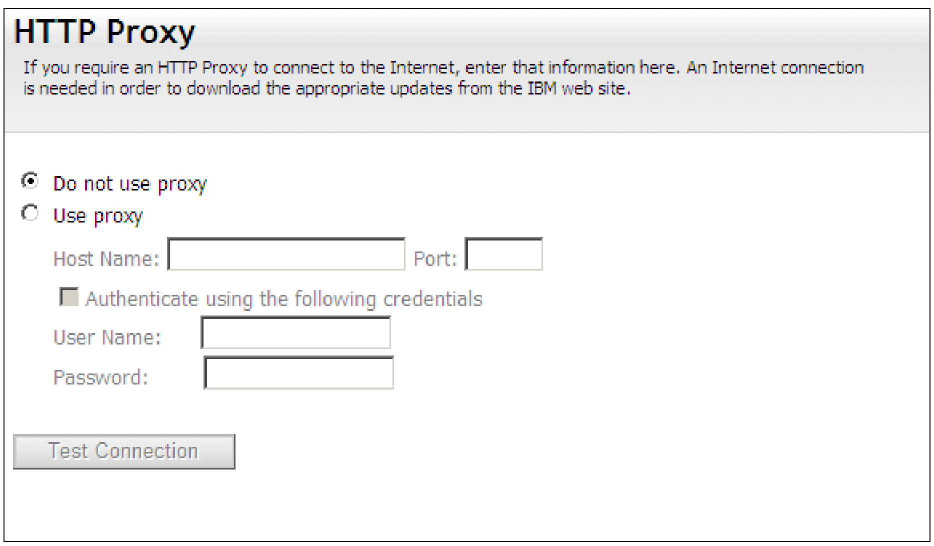


Figure 9-76 HTTP Proxy settings

9. Select one or more machine types that are contained on the bootable media and click **Next**, as shown in Figure 9-77 on page 518.

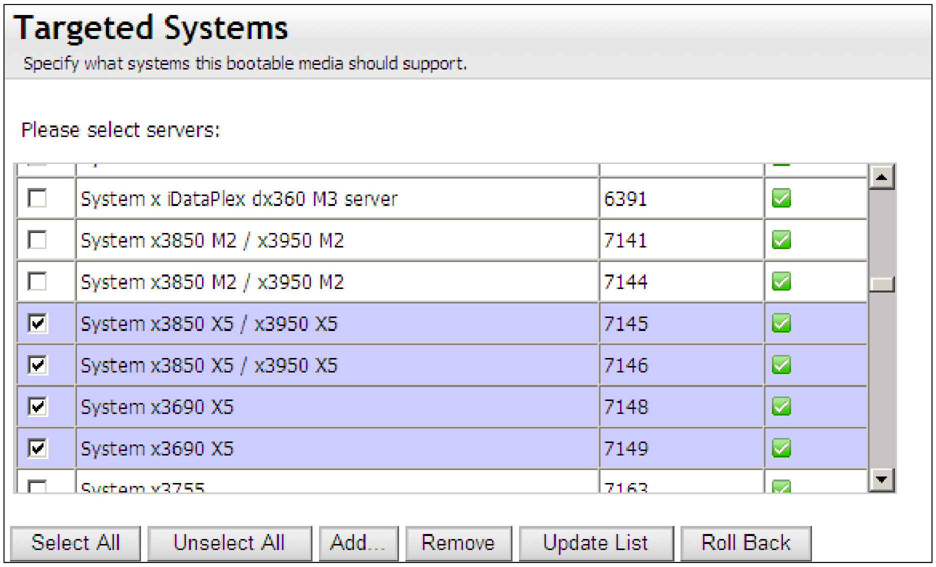


Figure 9-77 The targeted systems panel

10. Select the directory where you want to store the downloaded files, as shown in Figure 9-78. Click **Next**.

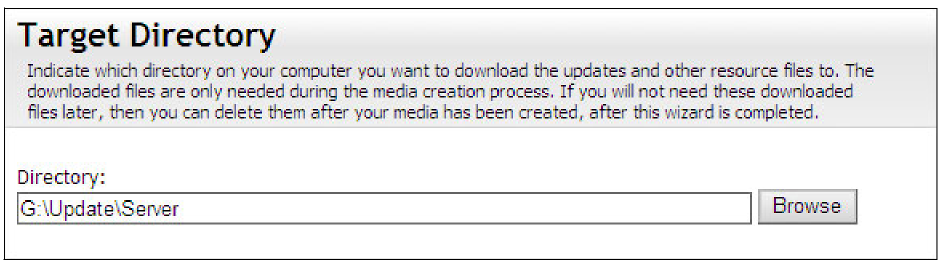


Figure 9-78 Example of a target directory

11. By default, BoMC creates an ISO file, as shown in Figure 9-79. You can choose another medium. Click **Next**.

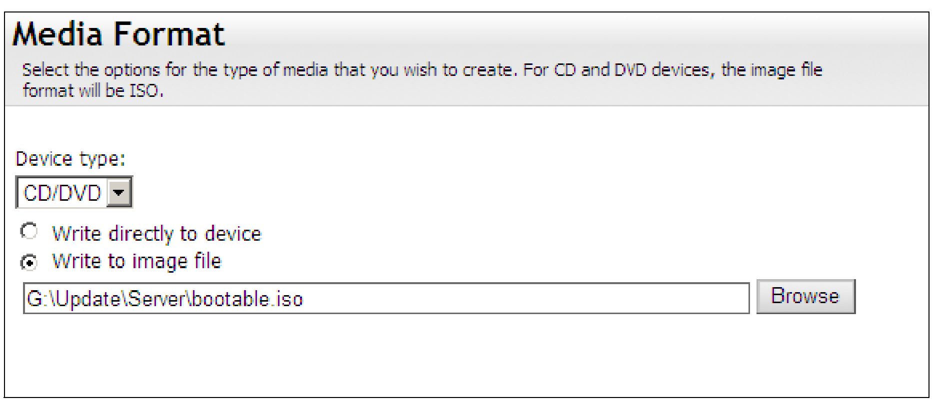


Figure 9-79 Example of a target media

12. Select **Do not use unattended mode** (Figure 9-80 on page 519) and click **Next**.

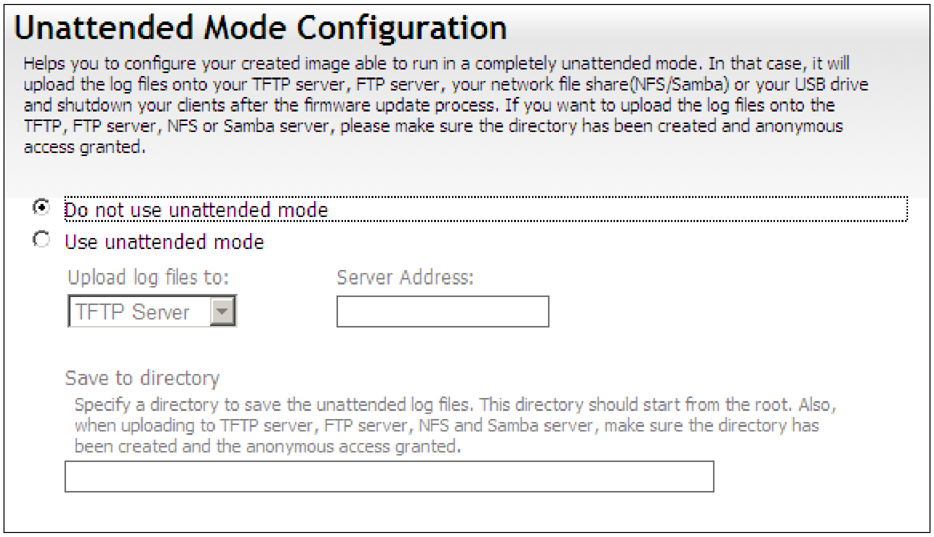


Figure 9-80 Unattended Mode panel

13. Review the selections and confirm that they are correct. Figure 9-81 shows an example. You can click **Save** to save this configuration information to a file. Click **Next**.

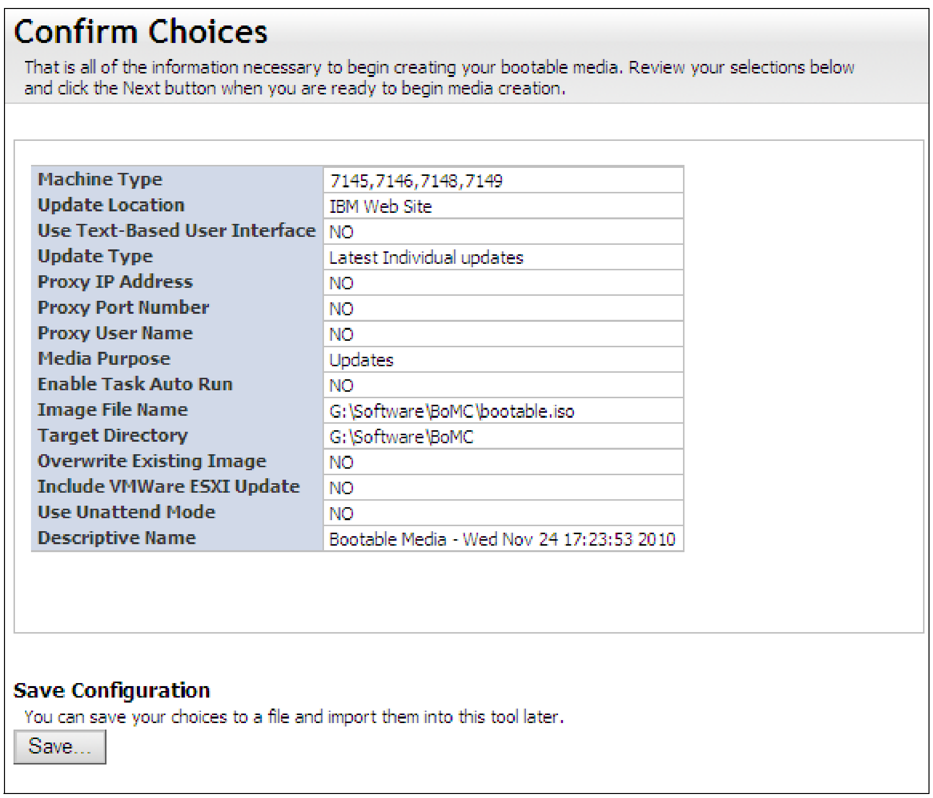


Figure 9-81 Confirm choices panel

14. BoMC acquires the files. In the progress bar, you can see the progress of the updates, as shown in Figure 9-82 on page 520.

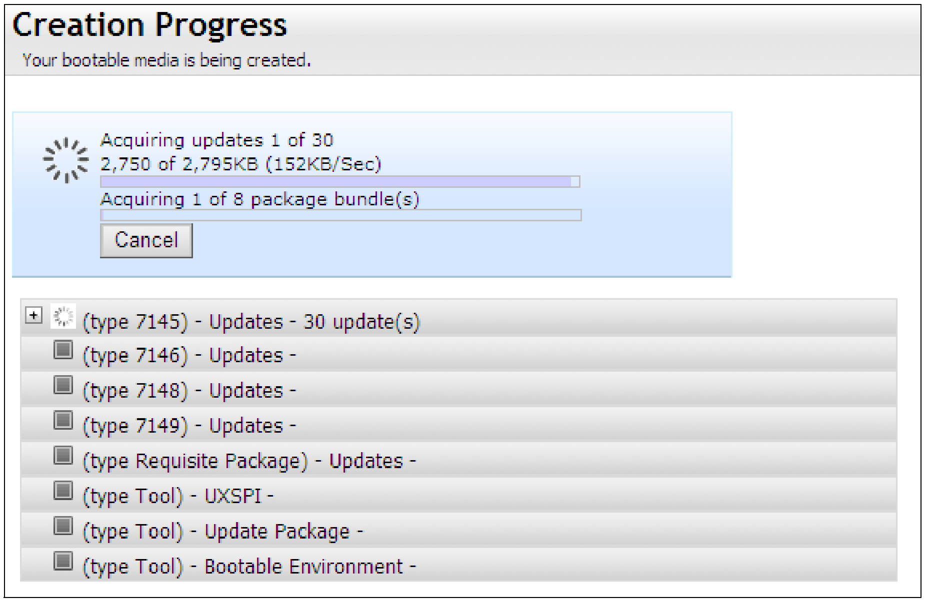


Figure 9-82 Downloading the files

15. After the update completes (Figure 9-83), click **Next**.

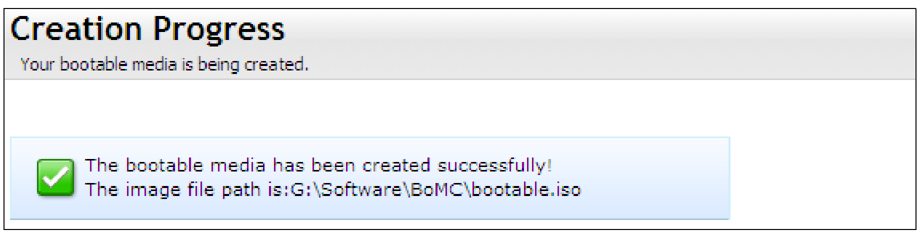


Figure 9-83 Confirmation that the creation process is finished

16. Click **Finish**.

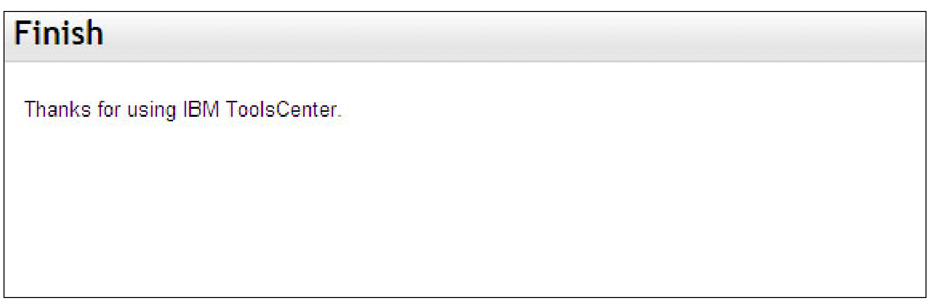


Figure 9-84 Finish panel

17. You now have a bootable image with the updates. You can mount or burn the image to CD and then boot the system with the medium.

9.13 MegaRAID Storage Manager

In this section, we provide an overview of the MegaRAID Storage Manager (MSM) software. For more information, see the *Installation and User's Guide* at the following website:

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

With MSM, you can configure, monitor, and maintain storage configurations on ServeRAID-M controllers. The MegaRAID Storage Manager graphical user interface (GUI) makes it easy for you to create and manage storage configurations. You can use MSM to manage local or remote RAID controllers and configure MSM for remote alert notifications. Also, a command-line interface is available.

The following controllers are supported:

- LSI 1064 SAS controller
- LSI 1064e SAS controller
- LSI 1068e SAS controller
- LSI 1078 SAS controller
- IBM Serial Attached SCSI (SAS) HBA
- IBM 3Gb SAS HBA v2
- IBM SAS Expansion Card (CFFv) for IBM BladeCenter
- IBM SAS Connectivity Card (CFFv) for IBM BladeCenter
- IBM SAS/Serial Advanced Technology Attachment (SATA) RAID Kit
- IBM ServeRAID BR10i SAS controller
- IBM ServeRAID BR10il SAS controller
- IBM MegaRAID 8480 SAS controller
- IBM ServeRAID MR10i SAS controller
- IBM ServeRAID MR10k SAS controller
- IBM ServeRAID MR10M SAS controller
- IBM ServeRAID MR10il SAS controller
- IBM ServeRAID MR10is SAS controller
- IBM ServeRAID MR10ie (CIOv) SAS controller
- IBM ServeRAID M5014 SAS/SATA controller
- IBM ServeRAID M5015 SAS/SATA controller
- IBM ServeRAID M5025 SAS/SATA Controller
- IBM ServeRAID M1015 SAS/SATA controller

9.13.1 Installation

To download the latest MegaRAID Storage Manager software and obtain the *Installation and User's Guide,* select the ServeRAID software matrix website:

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

The MSM supports the following operating systems:

- Microsoft Windows 2000, Microsoft Windows Server 2003, and Microsoft Windows Server 2008
- Red Hat Enterprise Linux Version 4.0 and Version 5.0
- SUSE SLES Version 9, Version 10, and Version 11, with the latest updates and service packs
- VMware 4.0, 4.0U1, and 4.1

To install the MSM software, read the *Installation and User's Guide for MSM* at the following website.

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

You must have administrator or root equivalent operating system privileges to install and fully access the MegaRAID Storage software.

There are four setup options:

- Complete: This option installs all program features.
- Client: This option installs the required components to remotely view and configure servers.
- Server: This option only installs the required components for remote server management.
- Stand Alone: This option only installs the required components for local server management.

9.13.2 Drive states

A *drive group* is one or more drives that are controlled by the RAID controller.

There are multiple drive states. The following list describes all of the possible drive states:

- Online

  A drive that can be accessed by the RAID controller and is part of the virtual drive.
- Unconfigured Good

  A drive that is functioning normally but is not configured.
- Hot Spare

  A drive that is powered up and ready for use as a spare in case an online drive fails. There are two Hot Spare drive states: Dedicated and Global.
- Failed

  A drive that was originally configured as Online or Hot Spare, but on which the firmware detects an unrecoverable error.
- Rebuild

  A drive to which data is being written to restore full redundancy for a virtual drive.
- Unconfigured Bad

  A drive on which the firmware detects an unrecoverable error.
- Missing

  A drive that was Online but which has been removed from its location.
- Offline

  A drive that is part of a virtual drive but which has invalid data as far as the RAID configuration is concerned.

9.13.3 Virtual drive states

A *virtual drive* is a partition in a drive group that is made up of contiguous data segments on the drives.

There are multiple virtual drive states. The following list describes all of the possible virtual drive states:

- Optimal

  The virtual drive operating condition is good. All configured drives are online.

- Degraded

  The virtual drive operating condition is not optimal. One of the configured drives has failed or is offline.

- Partially Degraded

  The operating condition in a RAID-6 virtual drive is not optimal. One of the configured drives has failed or is offline.

- Failed

  The virtual drive has failed.

- Offline

  The virtual drive is not available to the RAID controller.

9.13.4 MegaCLI utility for storage management

In this section, we provide an overview of the MegaCLI utility. For more information, see the *Installation and User's Guide* at the following website:

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

The MegaCLI utility is a command-line interface application. You can use this utility to configure, monitor, and maintain ServeRAID SAS RAID controllers and the devices that connect to them.

Creating a virtual drive with command-line interface (CLI)

In this example, we have two hard drives in slots one and two. Both hard drives must be Unconfigured Good, as shown in Figure 9-85.

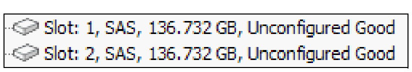


Figure 9-85 Two Unconfigured Good hard drives

Follow these steps to create a virtual drive with CLI:

1. Use the following command to locate the Enclosure Device ID and the Slot Number of both hard drives:

   ```
   MegaCli -PDList -aAll
   ```

   Example 9-15 shows the resulting output.

   Example 9-15 Output of the MegaCli -PDList -aAll command

   ```
   Enclosure Device ID: 252
   Slot Number: 1
   Device Id: 8
   ```

```
Sequence Number: 4
Media Error Count: 0
Other Error Count: 0
Predictive Failure Count: 0
Last Predictive Failure Event Seq Number: 0
PD Type: SAS
Raw Size: 136.731 GB [0x11176d60 Sectors]
Non Coerced Size: 136.231 GB [0x11076d60 Sectors]
Coerced Size: 135.972 GB [0x10ff2000 Sectors]
Firmware state: Unconfigured(good), Spun Up
SAS Address(0): 0x5000c50016f81501
SAS Address(1): 0x0
Connected Port Number: 1(path0)
Inquiry Data: IBM-ESXSST9146803SS     B53B3SD1GZCW0825B53B
IBM FRU/CRU: 42D0422
FDE Capable: Not Capable
FDE Enable: Disable
Secured: Unsecured
Locked: Unlocked
Foreign State: None
Device Speed: 6.0Gb/s
Link Speed: 6.0Gb/s
Media Type: Hard Disk Device
Drive:  Not Certified

Enclosure Device ID: 252
Slot Number: 2
Device Id: 16
Sequence Number: 3
Media Error Count: 0
Other Error Count: 0
Predictive Failure Count: 0
Last Predictive Failure Event Seq Number: 0
PD Type: SAS
Raw Size: 136.731 GB [0x11176d60 Sectors]
Non Coerced Size: 136.231 GB [0x11076d60 Sectors]
Coerced Size: 135.972 GB [0x10ff2000 Sectors]
Firmware state: Unconfigured(good), Spun Up
SAS Address(0): 0x5000c5001cf9c455
SAS Address(1): 0x0
Connected Port Number: 0(path0)
Inquiry Data: IBM-ESXSST9146803SS     B53B3SD2FM090825B53B
IBM FRU/CRU: 42D0422
FDE Capable: Not Capable
FDE Enable: Disable
Secured: Unsecured
Locked: Unlocked
Foreign State: None
Device Speed: 6.0Gb/s
Link Speed: 6.0Gb/s
Media Type: Hard Disk Device
Drive:  Not Certified
```

2. Now we can create the virtual drive. In our example, we issue the following command to create a RAID-1:

   `MegaCli -CfgLDAdd -R1[252:1,252:2] -a0`

   Example 9-16 on page 525 shows the resulting output.

Example 9-16 Output from command MegaCli -CfgLDAdd -R1[252:1,252:2] -a0

```
Adapter 0: Created VD 1

Adapter 0: Configured the Adapter!!

Exit Code: 0x00
```

3. The virtual drive is successfully created.

Additional command examples

The following command examples use the MegaCli command:

- Display help for MegaCLI

  `MegaCli –h|–Help|?`
- Display controller properties for all installed adapters

  `MegaCli -AdpAllinfo -aALL`
- Save configuration on the controller

  `MegaCli -CfgSave -f c:\saveconfig.txt -a0`
- Restore configuration data from file

  `MegaCli -CfgRestore -f c:\saveconfig.txt -a0`
- Display virtual drive information for all VD on all adapters

  `MegaCli -LDInfo -Lall -aALL`
- Display virtual drive and physical drive information for all adapters

  `MegaCli -LDPDInfo -aAll`
- Display number of virtual drives for all adapters

  `MegaCli –LDGetNum -aALL`
- Display list of physical devices for all adapters

  `MegaCli –PDList -aAll`

9.14 Serial over LAN

Serial over LAN (SoL) is a mechanism that enables the input and output of the serial port of a managed system to be redirected on the network over TCP/IP. SoL provides a means to manage servers remotely by using a command-line interface (CLI) over a Telnet or Secure Shell (SSH) connection.

SoL can give you remote access to the UEFI/BIOS and power-on self test (POST) messages. Using SoL, you can log in to the machine remotely. It can give you access to special operating system functions during boot.

In the x3850 X5 and x3690 X5, the serial port is shared with the integrated management module (IMM). The IMM can take control of the shared serial port to perform text console redirection and to redirect serial traffic, using Serial over LAN (SoL).

In the HX5, the Advanced Management Module (AMM) CLI provides access to the text console command prompt through the SoL connection.

In this section, we describe these topics:

9.14.1 Enabling SoL in UEFI

To enable SoL from the UEFI interface at boot time, press F1 when given the option and select **System Settings** → **Devices and I/O Ports** → **Console Redirection Settings**. In Table 9-7, you can find the UEFI settings that need to be set.

Table 9-7 Settings in UEFI for SoL

Value	HX5	x3850 X5 and x3690 X5
General settings		
COM Port 1	Disable	Enable
COM Port 2	Enable	Enable
Remote Console	Enable	Enable
Serial Port Sharing	Not available	Enable
Serial Port Access Mode	Not available	Shared
SP Redirection	Not available	Enable
Legacy Option ROM Display	COM Port 2	COM Port 2
COM settings		
Baud Rate	115200	115200
Data Bits	8	8
Parity	None	None
Stop Bits	1	1
Terminal Emulation	VT100	VT100
Active After Boot	Enable	Enable
Flow Control	Hardware	Hardware

Settings in UEFI for SoL:

- COM Port 1 has to be enabled only if the HX5 is used within a BC-H or BC-S where the Serial Pass-thru Module is implemented.
- Terminal Emulation can be set either to VT100 or ANSI; however, when configuring Linux operating systems, make sure that the OS settings match the terminal emulation that is selected in the hardware.

9.14.2 BladeCenter requirements

The BladeCenter chassis must be correctly configured before you can use the CLI and SoL. AMM is able to communicate and manage the blades effectively using management virtual LAN (VLAN) 4095 through switches in bays 1 and 2. By default, all internal ports of all BladeCenter switches are members of VLAN 4095.

Network switch requirements

You must verify a few settings to ensure that SoL will work on the chassis and the HX5. The network has these prerequisites:

- One Ethernet switch (any vendor) in Bay 1, regardless of the chassis type.
- Confirm that the HX5 is a member of VLAN 4095.

Cisco switch: If you use a Cisco switch, make sure that the firmware is at Version 12.1(22)EA6a or newer.

Enabling SoL in the AMM

An HX5 requires the AMM to function. You must set and adjust the following settings in the AMM:

1. In the AMM, ensure that the management network VLAN ID is set to `4095`, which is the default setting. To check the setting, select **Blade Tasks** → **Configuration** → **Management Network** (Figure 9-86).

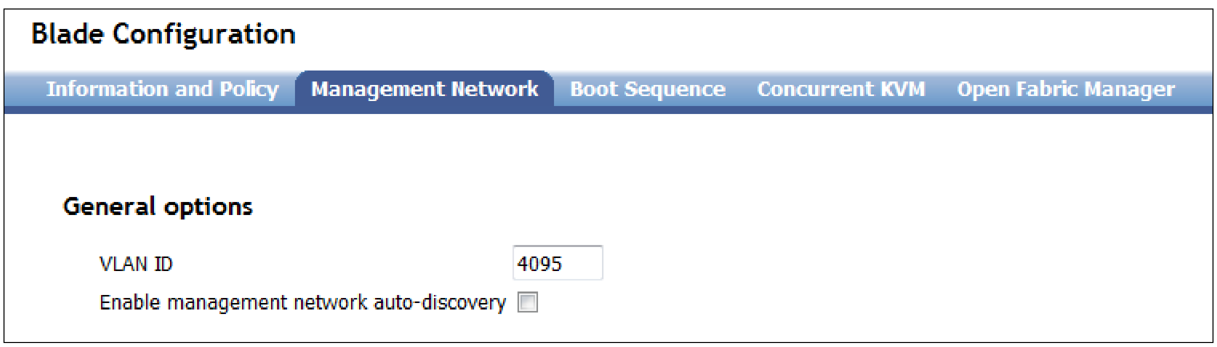
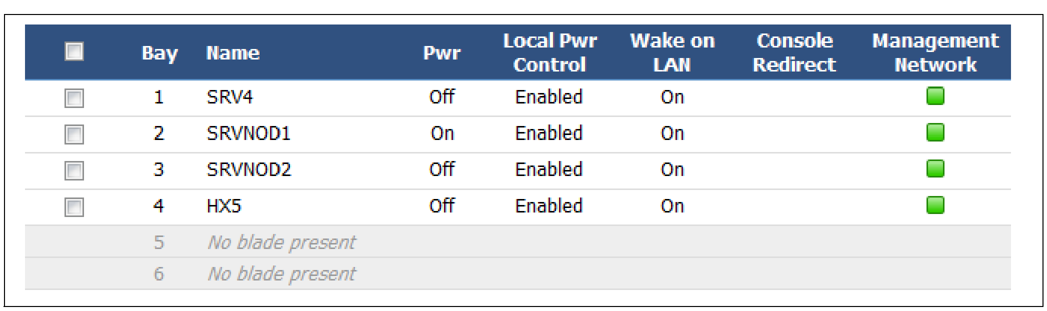


Figure 9-86 Management Network VLAN ID setting

2. If the HX5 works properly with the management VLAN ID 4095, select **Blade Tasks** → **Power/Restart** to verify that the Management Network icon is green, as shown in Figure 9-87.



	Bay	Name	Pwr	Local Pwr Control	Wake on LAN	Console Redirect	Management Network
	1	SRV4	Off	Enabled	On		
	2	SRVNOD1	On	Enabled	On		
	3	SRVNOD2	Off	Enabled	On		
	4	HX5	Off	Enabled	On		
	5	*No blade present*					
	6	*No blade present*					

Figure 9-87 Overview of Blade Status

3. Ensure that the SoL is enabled for the chassis. To verify, from the AMM, select **Blade Tasks** → **Serial Over LAN** → **Serial Over LAN Configuration**. Verify that Serial over LAN is **Enabled**, as shown in Figure 9-88.

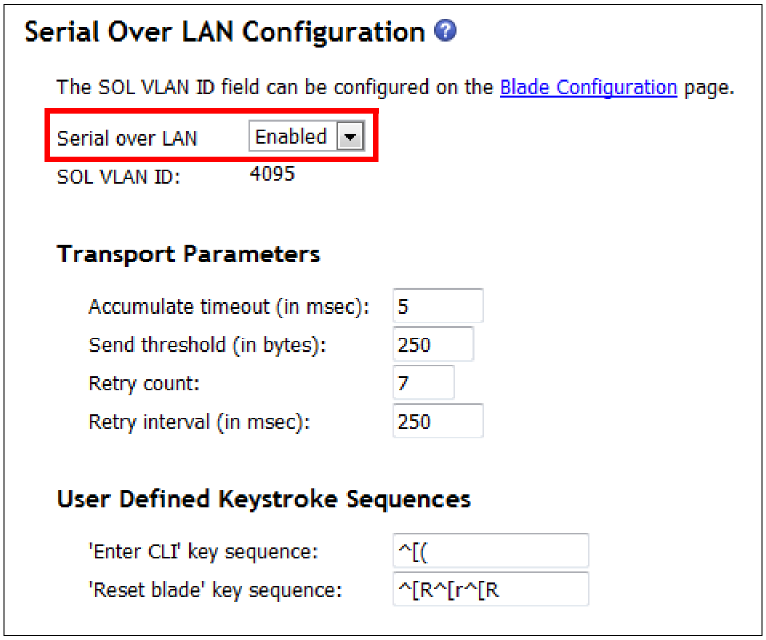


Figure 9-88 Serial Over LAN Configuration page

4. Finally, check if the SoL for the HX5 is enabled. Select **Blade Tasks** → **Serial Over LAN** → **Serial Over LAN Configuration** and scroll down to the Serial Over LAN Status section. Ensure that the SOL Status is green (Figure 9-89).

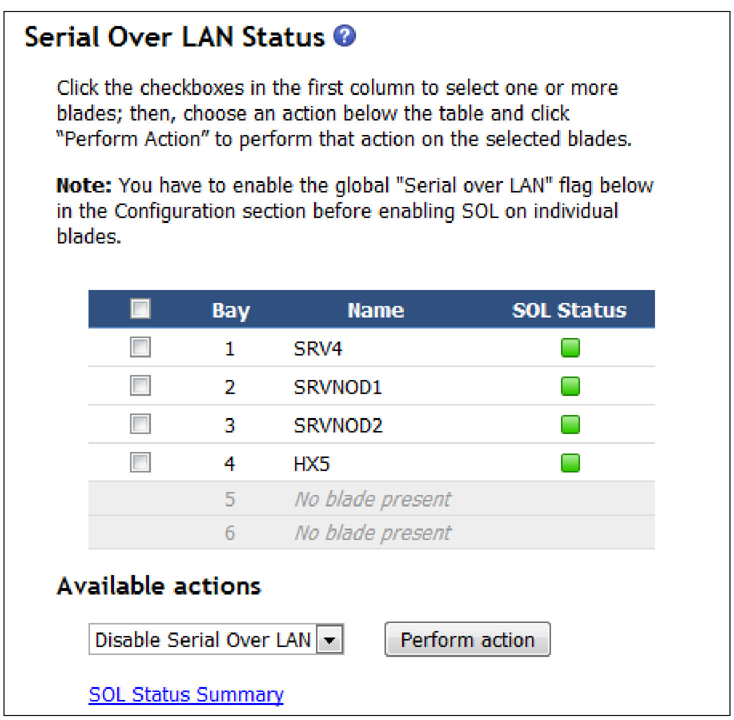


Figure 9-89 Serial Over LAN Status overview

Red circle icon: If you see a red circle icon next to the blade server, a requirement might not be satisfied. Or, you might have an issue with the Broadcom NIC drivers.

9.14.3 Enabling SoL in the operating system

In this section, we describe the settings to enable SoL in these operating systems:

- "Windows Server 2008"
- "Windows Server 2003" on page 531
- "Linux" on page 533

The onboard Broadcom NIC driver is required to be at the latest version. Certain older versions are known to impede SoL traffic. Download the latest version from the IBM FixCentral website:

http://ibm.com/support/fixcentral/

Windows Server 2008

To enable the Microsoft Emergency Messaging Service (EMS) and the Special Administration Console (SAC), use the following procedure. You must have administrator privileges. Perform these steps:

1. Start a command prompt: **Start** → **Run** → **cmd**.
2. Enter the command **bcdedit**. Example 9-17 shows the output from the **bcdedit** command.

Example 9-17 Output from bcdedit command

```
C:\Users\Administrator>bcdedit

Windows Boot Manager
--------------------
identifier              {bootmgr}
device                  partition=\Device\HarddiskVolume1
path                    \EFI\Microsoft\Boot\bootmgfw.efi
description             Windows Boot Manager
locale                  en-US
inherit                 {globalsettings}
default                 {current}
resumeobject            {87209f03-3477-11e0-a416-a69aee999ac5}
displayorder            {current}
toolsdisplayorder       {memdiag}
timeout                 30

Windows Boot Loader
-------------------
identifier              {current}
device                  partition=C:
path                    \Windows\system32\winload.efi
description             Windows Server 2008 R2
locale                  en-US
inherit                 {bootloadersettings}
recoverysequence        {87209f05-3477-11e0-a416-a69aee999ac5}
recoveryenabled         Yes
osdevice                partition=C:
systemroot              \Windows
resumeobject            {87209f03-3477-11e0-a416-a69aee999ac5}
nx                      OptOut
```

```
C:\Users\Administrator>
```

3. Enter the command **bcdedit /ems on**. Example 9-18 shows the output from this command.

Example 9-18 Output of bcdedit /ems on

```
C:\Users\Administrator>bcdedit /ems on
The operation completed successfully.
```

4. Modify the EMS settings to match the parameters that were configured at the hardware level with the following command:

   bcdedit /emssettings emsport:2 emsbaudrate:115200

Example 9-19 Output of the bcedit /emssettings emsport:2 emsbaudrate:115200 command

```
C:\Users\Administrator>bcdedit /emssettings emsport:2 emsbaudrate:115200
The operation completed successfully.
```

5. Enter **bcdedit** again to verify that EMS is activated.

Example 9-20 Output of the bcdedit command

```
C:\Users\Administrator>bcdedit

Windows Boot Manager
--------------------
identifier              {bootmgr}
device                  partition=\Device\HarddiskVolume1
path                    \EFI\Microsoft\Boot\bootmgfw.efi
description             Windows Boot Manager
locale                  en-US
inherit                 {globalsettings}
default                 {current}
resumeobject            {87209f03-3477-11e0-a416-a69aee999ac5}
displayorder            {current}
toolsdisplayorder       {memdiag}
timeout                 30

Windows Boot Loader
-------------------
identifier              {current}
device                  partition=C:
path                    \Windows\system32\winload.efi
description             Windows Server 2008 R2
locale                  en-US
inherit                 {bootloadersettings}
recoverysequence        {87209f05-3477-11e0-a416-a69aee999ac5}
recoveryenabled         Yes
osdevice                partition=C:
systemroot              \Windows
resumeobject            {87209f03-3477-11e0-a416-a69aee999ac5}
nx                      OptOut
ems                     Yes

C:\Users\Administrator>
```

6. Reboot the server to make the changes effective.

IMM setting

You can change the CLI mode for the COM port for EMS. Use the following procedure:

1. Log in to the web interface of IMM.
2. Navigate to **IMM Control** → **Serial Port**.
3. Change the CLI mode to **CLI with EMS compatible keystroke sequences** (Figure 9-90).

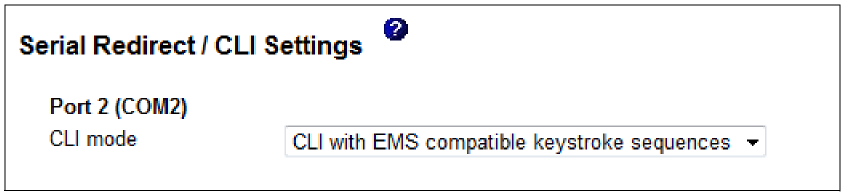


Figure 9-90 Serial Redirect/CLI Settings

4. Click **Save** to save the changes.

For more information about Microsoft Emergency Messaging Service and the Special Administration Console, see the following documents:

- *Boot Parameters to Enable EMS Redirection*

  http://msdn.microsoft.com/en-us/library/ff542282.aspx

- *Special Administration Console (SAC) and SAC commands*

  http://msdn.microsoft.com/en-us/library/cc785873

Windows Server 2003

Use the following procedure to enable the Microsoft Emergency Messaging Service and the Special Administration Console. You must have administrator privileges.

1. Start a command prompt **Start** → **Run** → **cmd**.
2. Enter the command **`bootcfg`**.

Example 9-21 Output of the bootcfg command

```
C:\>bootcfg

Boot Loader Settings
--------------------
timeout:30
default:multi(0)disk(0)rdisk(0)partition(1)\WINDOWS

Boot Entries
------------
Boot entry ID:    1
OS Friendly Name: Windows Server 2003, Enterprise
Path:             multi(0)disk(0)rdisk(0)partition(1)\WINDOWS
OS Load Options:  /noexecute=optout /fastdetect

C:\>
```

3. Examine the output. If more than one boot entry exist, you must determine the default entry.
4. Enable EMS with the **`bootcfg /ems on /port com2 /baud 115200 /id 1`** command. In our example, the default boot entry has the ID `1` (Example 9-22 on page 532).

Example 9-22 Output of the bootcfg /ems on /port com2 /baud 115200 /id 1 command

```
C:\>bootcfg /ems on /port com2 /baud 115200 /id 1
SUCCESS: Changed the redirection port in boot loader section.
SUCCESS: Changed the redirection baudrate in boot loader section.
SUCCESS: Changed the OS entry switches for line "1" in the BOOT.INI file.
```

5. Enter **bootcfg** again to verify that the EMS is activated.

Example 9-23 Output of the bootcfg command

```
C:\>bootcfg

Boot Loader Settings
--------------------
timeout:         30
default:         multi(0)disk(0)rdisk(0)partition(1)\WINDOWS
redirect:        COM2
redirectbaudrate:115200

Boot Entries
------------
Boot entry ID:    1
OS Friendly Name: Windows Server 2003, Enterprise
Path:             multi(0)disk(0)rdisk(0)partition(1)\WINDOWS
OS Load Options:  /noexecute=optout /fastdetect /redirect

C:\>
```

6. Reboot the server to make the changes effective.

IMM Setting

To change the CLI mode for the COM port for use with EMS, use the following procedure:

1. Log in to the web interface of IMM.
2. Navigate to **IMM Control** → **Serial Port**.
3. Change the CLI mode to **CLI with EMS compatible keystroke sequences**.

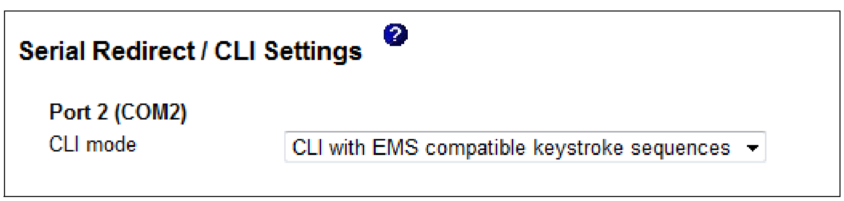


Figure 9-91 Serial Redirect/CLI Settings

4. Click **Save** to save the changes.

For more information about Microsoft Emergency Messaging Service and the Special Administration Console, see the following documents:

- *Boot Parameters to Enable EMS Redirection*

  http://msdn.microsoft.com/en-us/library/ff542282.aspx

- *Special Administration Console (SAC) and SAC commands*

  http://msdn.microsoft.com/en-us/library/cc785873

Linux

You must edit two files in Linux to ensure that the console redirection still works after the operating system has loaded. The same files are changed for Red Hat Linux (RHEL) and SUSE Linux.

Edit these files:

- `/boot/grub/menu.lst`
- `/etc/inittab`

> **RHEL 6:** If you have installed RHEL 6 in UEFI mode, you must edit the `/boot/efi/EFI/redhat/grub.conf` file instead of the `/boot/grub/menu.lst` file.

Menu.lst or grub.conf

Add the parameter that is highlighted in bold in the file that is shown in Example 9-24.

Example 9-24 Example of the grub.conf

```
# grub.conf generated by anaconda
#
# Note that you do not have to rerun grub after making changes to this file
# NOTICE:  You have a /boot partition.  This means that
#          all kernel and initrd paths are relative to /boot/, eg.
#          root (hd0,1)
#          kernel /vmlinuz-version ro root=/dev/mapper/VolGroup-lv_root
#          initrd /initrd-[generic-]version.img
#boot=/dev/sda1
device (hd0) HD(1,800,64000,699900f5-c584-4061-a99f-d84c796d5c72)
default=0
timeout=5
splashimage=(hd0,1)/grub/splash.xpm.gz
hiddenmenu
title Red Hat Enterprise Linux (2.6.32-71.el6.x86_64)
        root (hd0,1)
        kernel /vmlinuz-2.6.32-71.el6.x86_64 ro root=/dev/mapper/VolGroup-lv_roo
t rd_LVM_LV=VolGroup/lv_root rd_LVM_LV=VolGroup/lv_swap rd_NO_LUKS rd_NO_MD rd_N
O_DM LANG=en_US.UTF-8 SYSFONT=latarcyrheb-sun16 KEYBOARDTYPE=pc KEYTABLE=de cras
hkernel=auto console=ttyS1,115200n8 rhgb quiet
        initrd /initramfs-2.6.32-71.el6.x86_64.img
[root@localhost redhat]#
```

/etc/inittab

Add the parameter that is highlighted in bold at the end of the `/etc/inittab` file, as shown in Example 9-25.

Example 9-25 The /etc/inittab file

```
id:3:initdefault:
co:2345:respawn:/sbin/agetty ttyS1 115200 vt100-nav
[root@localhost etc]#
```

9.14.4 How to start a SoL connection

In this section, we explain how to connect to the HX5, x3850 X5, or x3690 X5 through SoL.

Connecting to the x3690 X5 or x3850 X5

You can use Telnet or SSH to connect to the x3690 X5 or x3850 X5 servers. Follow these steps:

1. Start a Telnet/SSH session to the IMM IP address.
2. Log in to the IMM. The default user id is `USERID` and the default password is `PASSW0RD`, where the 0 is a zero.
3. The IMM CLI main page appears (Figure 9-92).

```
                              Legacy CLI Application

system>
```

Figure 9-92 IMM CLI main page

4. Start SoL with the `console 1` command.

The SoL console starts and you see whatever is transmitted over the SoL connection (for example, the UEFI setup windows). See "Sample output" for examples of what you see over an SoL connection when the system boots.

Connecting to the HX5

You can use Telnet or SSH for the connection to the BladeCenter HX5. Follow these steps:

1. Start a Telnet/SSH session to the AMM IP address.
2. Log in to the AMM. The default user id is `USERID` and default password is `PASSW0RD`, where the 0 is a zero.
3. The AMM CLI main page appears (Figure 9-93).

```
Hostname:             MM00145EDF234C
Static IP address:    10.0.0.125
Burned-in MAC address: 00:14:5E:DF:23:4C
DHCP:                 Disabled - Use static IP configuration.
Last login: Saturday January 1 2000 1:40 from 10.0.0.100 (Web)

system>
```

Figure 9-93 AMM CLI main page

4. Start SoL with the **`console -T system:blade[5]`** command.

> **Determining the slot number:** In this example, we started a SoL session with the blade in slot 5. To determine the slot number of your blade, select **Blade Tasks** → **Power/Restart**.

The SoL console starts and you see whatever is transmitted over the SoL connection (for example, the UEFI setup windows). See "Sample output" for examples of what you see over an SoL connection when the system boots.

Sample output

After you connect to the system using SoL and power on the server, you see the output in the telnet/SSH window. Figure 9-94 shows the x3850 X5 booting.

```
Platform Initization Complete

System x3850 X5

UEFI Builde Ver: 1.40    IMM Build Ver: 1.24   Diagnostics Build Ver: 3.30

2 CPU Packages Available at 2.00 GHz Link Speed
16 GB Memory Available at 1067 MHz in Lockstep Mode




Connecting Boot Devices and Adapters...
```

Figure 9-94 Booting the x3850 X5

Figure 9-95 shows the logo window via SoL. The IBM logo is not displayed.

```
<F1> Setup          <F2> Diagnostics       <F12> Select Boot Device
```

Figure 9-95 Logo boot panel

Figure 9-96 shows Windows starting.

```
                              Starting Windows...

For troubleshooting and advanced startup options for Windows, press F8.
```

Figure 9-96 Windows booting as seen via SoL

Figure 9-97 shows the Windows SAC console.

```
Computer is booting, SAC started and initialized.

Use the “ch -?” command for information about using channels
Use the “?” command for general help.

SAC>
```

Figure 9-97 Microsoft Windows Special Administration Console (Windows only)

Figure 9-98 shows Linux booting.

```
Enabling /etc/fstab swaps:  [ OK ]
Entering non-interactive startup
Applying Intel CPU microcode update: Calling the system activity data collector (sadc):
Starting monitoring for VG VolGroup:  3 logical volume(s) in volume group “VolGroup”
monitored
[ OK ]
ip6tables: Applying firewall rules: [ OK ]
iptables: Applying firewall rules: [ OK ]
Bringing up loopback interface: [ OK ]
Starting auditd: [ OK ]
Starting system logger: [ OK ]
Enabling ondemand cpu frequency scaling: [ OK ]
Starting irqbalance: [ OK ]
Starting rpcbind: [ OK ]
Starting NFS statd: [ OK ]
Starting mdmonitor: [ OK ]
Starting RPC idmapd: [ OK ]
Your running kernel is usuing more than 70% of the amount of space you reserved for
kdump, you should consider increasing your crashkernel reservation[WARNING]
Starting kdump: [ OK ]
Starting system message bus: [ OK ]
Mounting other filesystems: [ OK ]
Starting acpi daemon: [ OK ]
Starting HAL daemon:
```

Figure 9-98 Linux boot sequence as seen in SoL

Abbreviations and acronyms

ac	alternating current
AES	Advanced Encryption Standard
AIK	Automated Installation Kit
AMM	Advanced Management Module
API	application programming interface
APIC	Advanced Programmable Interrupt Controller
ASU	Advanced Settings Utility
BC	BladeCenter
BCD	Boot Configuration Database
BIOS	basic input output system
BMC	Baseboard Management Controller
BoMC	Bootable Media Creator
BS	Blue Screen
CAS	column address strobe
CD	compact disc
CIM	Common Information Model
CKE	Clock Enable
CKVM	Concurrent KVM
CLI	command-line interface
CMOS	complementary metal oxide semiconductor
CNA	Converged Network Adapter
COD	configure on disk
COG	configuration and option guide
COM	Component Object Model
CPU	central processing unit
CRC	cyclic redundancy check
CRT	Cathode Ray Tube
CRU	customer-replaceable units
CTO	configure-to-order
DAU	demand acceleration unit
DB	database
DDF	Disk Data Format
DHCP	Dynamic Host Configuration Protocol
DIMM	dual inline memory module
DMA	direct memory access
DNS	Domain Name System
DPC	deferred procedure call
DRAM	dynamic random access memory
DSA	Dynamic System Analysis
ECC	error checking and correcting
EIA	Electronic Industries Alliance
EMS	Emergency Messaging Service
ER	enterprise rack
ESA	Electronic Service Agent
ESD	electrostatic discharge
ETS	Enhanced Technical Support
EXA	Enterprise X-Architecture
FAMM	Full Array Memory Mirroring
FC	Fibre Channel
FCP	Fibre Channel Protocol
FPGA	Field Programmable Gate Array
FRU	field-replaceable unit
GB	gigabyte
GPT	GUID Partition Table
GPU	Graphics Processing Unit
GRUB	Grand Unified Bootloader
GT	Gigatransfers
GUI	graphical user interface
HBA	host bus adapter
HDD	hard disk drive
HE	high end
HPC	high performance computing
HPCBP	High Performance Computing Basic Profile
HPET	High Precision Event Timer
HS	hot swap
HT	Hyper-Threading
HTTP	Hypertext Transfer Protocol
I/O	input/output
IBM	International Business Machines
ID	identifier
IEC	International Electrotechnical Commission
IEEE	Institute of Electrical and Electronics Engineers
IM	instant messaging
IME	Integrated Mirroring Enhanced
IMM	Integrated Management Module
IOPS	I/O operations per second

IP	Internet Protocol
IPMB	Intelligent Platform Management Bus
IPMI	Intelligent Platform Management Interface
IS	information store
ISO	International Organization for Standards
IT	information technology
ITSO	International Technical Support Organization
JBOD	just a bunch of disks
KB	kilobyte
KCS	keyboard console style
KVM	keyboard video mouse
LAN	local area network
LDAP	Lightweight Directory Access Protocol
LED	light emitting diode
LGA	land grid array
LPD	light path diagnostic
LUN	logical unit number
MAC	media access control
MB	megabyte
MCA	Machine Check Architecture
MDIX	medium-dependent interface crossover
MESI	modified exclusive shared invalid
MIPS	millions of instructions per second
MM	Management Module
MSM	MegaRAID Storage Manager
NAS	network-attached storage
NHS	non-hot-swap
NIC	network interface card
NMI	non-maskable interrupt
NOS	network operating system
NTLM	NT LAN Manager
NUMA	Nonuniform memory access
OGF	Open Grid Forum
OS	operating system
PCI	Peripheral Component Interconnect
PD	problem determination
PDSG	Problem Determination and Service Guide
PE	Preinstallation Environment
PFA	Predictive Failure Analysis
PMI	Project Management Institute
POST	power-on self test
PS	Personal System
PXE	Preboot eXecution Environment
QPI	QuickPath Interconnect
RAID	redundant array of independent disks
RAM	random access memory
RAS	remote access services; row address strobe; reliability, availability, and serviceability
RDMA	Remote Direct Memory Access
RETAIN	Remote Electronic Technical Assistance Information Network
RHEL	Red Hat Enterprise Linux
RISC	reduced instruction set computer
ROC	RAID-on-card
ROM	read-only memory
RPM	revolutions per minute
RSA	Remote Supervisor Adapter
RSS	Receive-side scaling
RTC	real-time clock
RTS	request to send
SAC	Special Administration Console
SAN	storage area network
SAS	Serial Attached SCSI
SATA	Serial Advanced Technology Attachment
SCSI	Small Computer System Interface
SDRAM	static dynamic RAM
SED	self-encrypting drive
SFP	small form-factor pluggable
SLC	Single Level Cell
SLES	SUSE Linux Enterprise Server
SLP	Service Location Protocol
SMASH	Systems Management Architecture for Server Hardware
SMI	scalable memory interconnect
SMP	symmetric multiprocessing
SNMP	Simple Network Management Protocol
SOA	service-oriented architecture
SOL	Serial over LAN
SPORE	ServerProven Opportunity Request for Evaluation
SR	short range

SSCT	Standalone Solution Configuration Tool
SSD	solid-state drive
SSH	Secure Shell
SSIC	System Storage Interoperation Center
SSL	Secure Sockets Layer
STG	Server and Technology Group
TB	terabyte
TCG	Trusted Computing Group
TCO	total cost of ownership
TCP	Transmission Control Protocol
TCP/IP	Transmission Control Protocol/Internet Protocol
TDP	thermal design power
TFTP	Trivial File Transfer Protocol
TOE	TCP offload engine
TPM	Trusted Platform Module
UDP	user datagram protocol
UE	Unrecoverable Error
UEFI	Unified Extensible Firmware Interface
UPS	uninterruptible power supply
URL	Uniform Resource Locator
USB	universal serial bus
UXSP	UpdateXpress System Packs
UXSPI	UpdateXpress System Packs Installer
VFA	Virtual Fabric Adapter
VLAN	virtual LAN
VLP	very low profile
VMFS	virtual machine file system
VPD	vital product data
VRM	voltage regulator module
VSMP	Virtual SMP
VT	Virtualization Technology

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this book.

IBM Redbooks publications

The following IBM Redbooks publications provide additional information about the topic in this document. Note that several publications referenced in this list might be available in softcopy only.

- *Architecting a Highly Efficient Image Management System with Tivoli Provisioning Manager for OS Deployment*, REDP-4294
- *Deploying Linux Systems with Tivoli Provisioning Manager for OS Deployment*, REDP-4323
- *Deployment Guide Series: Tivoli Provisioning Manager for OS Deployment V5.1*, SG24-7397
- *Emulex 10Gb Virtual Fabric Adapter for IBM System x*, TIPS0762
- *High availability virtualization on the IBM System x3850 X5*, TIPS0771
- *IBM 6Gb SSD Host Bus Adapter for IBM System x*, TIPS0744
- *IBM eX5 Portfolio Overview: IBM System x3850 X5, x3950 X5, x3690 X5, and BladeCenter HX5*, REDP-4650
- *IBM Midrange System Storage Hardware Guide*, SG24-7676
- *IBM Midrange System Storage Implementation and Best Practices Guide*, SG24-6363
- *IBM ServeRAID Adapter Quick Reference*, TIPS0054
- *IBM System Storage DS3000: Introduction and Implementation Guide*, SG24-7065
- *IBM System Storage Solutions Handbook*, SG24-5250
- *IBM XIV Storage System: Architecture, Implementation, and Usage*, SG24-7659
- *Implementing an IBM System x iDataPlex Solution*, SG24-7629
- *Implementing an IBM/Brocade SAN with 8 Gbps Directors and Switches*, SG24-6116
- *Implementing an IBM/Cisco SAN*, SG24-7545
- *Implementing an Image Management System with Tivoli Provisioning Manager for OS Deployment: Case Studies and Business Benefits*, REDP-4513
- *Implementing IBM Systems Director 6.1*, SG24-7694
- *Implementing IBM Systems Director Active Energy Manager 4.1.1*, SG24-7780
- *ServeRAID B5015 SSD Controller*, TIPS0763
- *ServeRAID M1015 SAS/SATA Controller for System x*, TIPS0740
- *ServeRAID M5015 and M5014 SAS/SATA Controllers for IBM System x*, TIPS0738
- *ServeRAID M5025 SAS/SATA Controller for IBM System x*, TIPS0739
- *ServeRAID-BR10il SAS/SATA Controller v2 for IBM System x*, TIPS0741
- *Tivoli Provisioning Manager for OS Deployment in a Retail Environment*, REDP-4372

- *Vista Deployment Using Tivoli Provisioning Manager for OS Deployment*, REDP-4295

You can search for, view, or download IBM Redbooks, Redpapers, Webdocs, draft publications and additional materials, as well as order hardcopy IBM Redbooks publications, at this website:

ibm.com/redbooks

Other publications

Publications listed in this section are also relevant as further information sources.

IBM System x3850 X5 and x3950 X5

See the following publications:

- *Installation and User's Guide - IBM System x3850 X5 and x3950 X5*

  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5084846
- *Problem Determination and Service Guide - IBM System x3850 X5 and x3950 X5*

  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5084848
- *Rack Installation Instructions - IBM System x3850 X5 and x3950 X5*

  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5083419
- *Installation Instructions for the IBM 2-Node x3850 X5 and x3950 X5 Scalability Kit*

  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5084859
- *Installation Instructions for the IBM eX5 MAX5 to x3850 X5 and x3950 X5 QPI Cable Kit and IBM eX5 MAX5 2-Node EXA Scalability Kit*

  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5084861

IBM System x3690 X5

See the following publications:

- *Installation and User's Guide - IBM System x3690 X5*

  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5085206
- *Problem Determination and Service Guide - IBM System x3690 X5*

  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5085205
- *IBM eX5 MAX5 to x3690 X5 QPI cable kit and IBM eX5 MAX5 2-node EXA scalability kit installation instructions - IBM System x3690 X5*

  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5085207

IBM BladeCenter HX5

See the following publications:

- *IBM BladeCenter Information Center*

  http://publib.boulder.ibm.com/infocenter/bladectr/documentation
- *Installation and User's Guide - IBM BladeCenter HX5*

  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5084612
- *Problem Determination and Service Guide - IBM BladeCenter HX5*

  http://www.ibm.com/support/docview.wss?uid=psg1MIGR-5084529

Online resources

The following websites are also relevant as further information sources:

- IBM eX5 home page

  http://ibm.com/systems/ex5
- IBM System x3850 X5 home page

  http://ibm.com/systems/x/hardware/enterprise/x3850x5
- IBM BladeCenter HX5 home page

  http://ibm.com/systems/bladecenter/hardware/servers/hx5
- IBM System x3690 X5 home page

  http://ibm.com/systems/x/hardware/enterprise/x3690x5
- IBM System x Configuration and Options Guide

  http://ibm.com/support/docview.wss?uid=psg1SCOD-3ZVQ5W
- IBM ServerProven

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

Help from IBM

IBM support and downloads

ibm.com/support

IBM Global Services

ibm.com/services

Index

Numerics

A

B

C

D

I

J

K

L

M

V

W

X

IBM eX5 Implementation Guide

(1.0” spine)
0.875”<->1.498”
460 <-> 788 pages

IBM eX5 Implementation Guide

Covers the IBM System x3950 X5, x3850 X5, x3690 X5, and the IBM BladeCenter HX5

Details technical information about each server and option

Describes how to implement two-node configurations

High-end workloads drive ever-increasing and ever-changing constraints. In addition to requiring greater memory capacity, these workloads challenge you to do more with less and to find new ways to simplify deployment and ownership. And although higher system availability and comprehensive systems management have always been critical, they have become even more important in recent years.

Difficult challenges, such as these, create new opportunities for innovation. The IBM eX5 portfolio delivers this innovation. This family of high-end computing introduces the fifth generation of IBM X-Architecture technology. The family includes the IBM System x3850 X5, x3690 X5, and the IBM BladeCenter HX5. These servers are the culmination of more than a decade of x86 innovation and firsts that have changed the expectations of the industry. With this latest generation, eX5 is again leading the way as the shift toward virtualization, platform management, and energy efficiency accelerates.

This book is divided into two parts. In the first part, we provide detailed technical information about the servers in the eX5 portfolio. This information is most useful in designing, configuring, and planning to order a server solution. In the second part of the book, we provide detailed configuration and setup information to get your servers operational. We focus particularly on setting up MAX5 configurations of all three eX5 servers as well as 2-node configurations of the x3850 X5 and HX5.

This book is aimed at clients, IBM Business Partners, and IBM employees that want to understand the features and capabilities of the IBM eX5 portfolio of servers and want to learn how to install and configure the servers for use in production.

INTERNATIONAL TECHNICAL SUPPORT ORGANIZATION

BUILDING TECHNICAL INFORMATION BASED ON PRACTICAL EXPERIENCE

IBM Redbooks are developed by the IBM International Technical Support Organization. Experts from IBM, Customers and Partners from around the world create timely technical information based on realistic scenarios. Specific recommendations are provided to help you implement IT solutions more effectively in your environment.

For more information:
ibm.com/redbooks

SG24-7909-00 ISBN 0738435643