

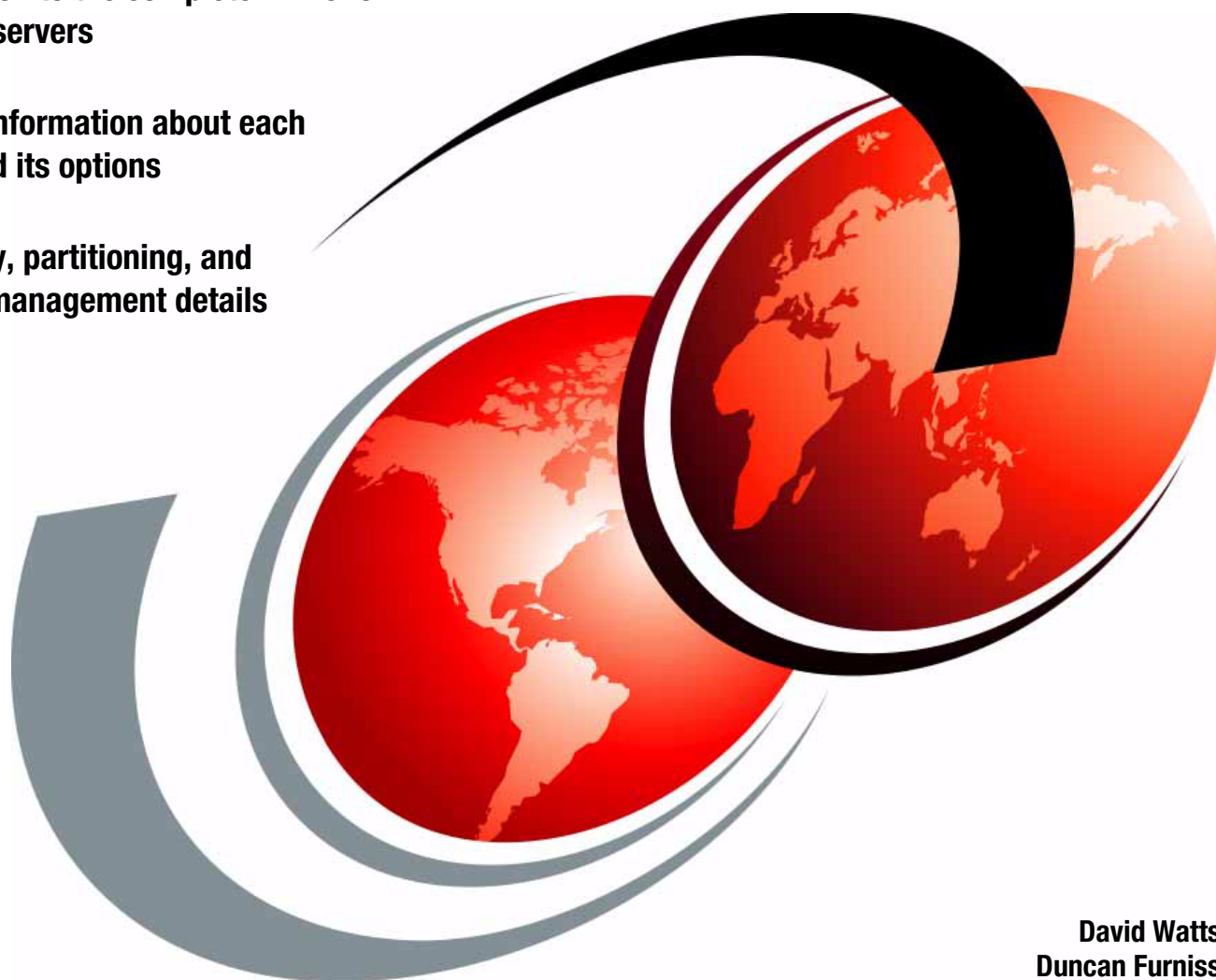
IBM eX5 Portfolio Overview

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

Introduction to the complete IBM eX5 family of servers

Detailed information about each server and its options

Scalability, partitioning, and systems management details



David Watts
Duncan Furniss



International Technical Support Organization

**IBM eX5 Portfolio Overview: IBM System x3850 X5,
x3950 X5, x3690 X5, and BladeCenter HX5**

April 2013

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

Sixth Edition (April 2013)

This edition applies to the following IBM eX5 products:

- ▶ IBM System x3690 X5
- ▶ IBM System x3850 X5
- ▶ IBM System x3950 X5
- ▶ IBM MAX5 for System x
- ▶ IBM BladeCenter HX5
- ▶ IBM MAX5 for BladeCenter

© Copyright International Business Machines Corporation 2010, 2011, 2013. All rights reserved.

Note to U.S. Government Users Restricted Rights -- Use, duplication or disclosure restricted by GSA ADP Schedule Contract with IBM Corp.

Contents

Notices	ix
Trademarks	x
Preface	xi
The team who wrote this paper	xi
Now you can become a published author, too!	xii
Comments welcome	xiii
Stay connected to IBM Redbooks	xiii
Summary of changes	xv
April 2013, Sixth Edition	xv
December 2011, Fifth Edition	xv
October 2010, Fourth Edition	xvi
August 2010, Third Edition	xvii
June 2010, Second Edition	xvii
Chapter 1. Introduction	1
1.1 eX5 systems	2
1.2 Positioning	3
1.2.1 IBM System x3850 X5 and x3950 X5	3
1.2.2 IBM System x3690 X5	5
1.2.3 IBM BladeCenter HX5	6
1.3 Energy efficiency	6
1.4 Services offerings	7
Chapter 2. IBM eX5 technology	9
2.1 eX5 chip set	10
2.2 Intel Xeon processors	10
2.2.1 Intel Xeon E7 processors	10
2.2.2 Intel Advanced Encryption Standard - New Instructions	11
2.2.3 Intel Virtualization Technology	12
2.2.4 Hyper-Threading Technology	12
2.2.5 Turbo Boost Technology	12
2.2.6 QuickPath Interconnect	13
2.2.7 Processor performance in a green world	15
2.3 Memory	16
2.3.1 Memory speed	17
2.3.2 Memory dual inline memory module placement	18
2.3.3 Memory ranking	19
2.3.4 Non-uniform memory access architecture	21
2.3.5 Hemisphere mode	22
2.3.6 Reliability, availability, and serviceability features	23
2.3.7 Scalable memory buffers	26
2.3.8 I/O hubs	26
2.4 MAX5	26
2.5 Scalability	28
2.6 Partitioning	30
2.7 Unified Extensible Firmware Interface system settings	32
2.7.1 System power operating modes	34

2.7.2	System power settings	37
2.8	IBM eXFlash	38
2.8.1	SSD and RAID controllers	39
2.8.2	IBM eXFlash price-performance	40
2.9	Integrated virtualization	41
2.9.1	VMware ESXi and vSphere	41
2.9.2	Red Hat RHEV-H (KVM)	42
2.9.3	Windows 2008 R2, Windows 2012 with Hyper-V	42
Chapter 3.	IBM System x3850 X5 and x3950 X5	43
3.1	Product features	44
3.1.1	IBM System x3850 X5 product features	44
3.1.2	IBM System x3950 X5 product features	47
3.1.3	IBM MAX5 memory expansion unit	47
3.1.4	Comparing the x3850 X5 to the x3850 M2	51
3.2	Target workloads	52
3.3	Models	53
3.3.1	x3850 X5 base models with Intel E7 processors	53
3.3.2	Workload-optimized x3950 X5 models with Intel E7 processors	54
3.4	System architecture	57
3.4.1	System board	57
3.4.2	QPI wrap card	58
3.5	MAX5	60
3.6	Scalability	63
3.6.1	Memory scalability with MAX5	63
3.6.2	Two-node scalability	64
3.6.3	Two-node and MAX5 scalability	66
3.6.4	FlexNode partitioning	67
3.6.5	MAX5 and XceL4v Dynamic Server Cache	68
3.7	Processor options	68
3.7.1	Intel Xeon E7 processor options	68
3.7.2	Population guidelines	69
3.8	Memory	70
3.8.1	Memory cards	70
3.8.2	Memory DIMMs for the x3850 X5	71
3.8.3	Memory DIMMs for MAX5	72
3.8.4	DIMM population sequence	74
3.8.5	Maximizing memory performance	79
3.8.6	Memory mirroring	82
3.8.7	Memory sparing	84
3.8.8	Effect on performance by using mirroring or sparing	84
3.9	Storage	85
3.9.1	Internal disks	85
3.9.2	SAS and SSD 2.5-inch disk support	86
3.9.3	IBM eXFlash and 1.8-inch SSD support	88
3.9.4	SAS and SSD controllers	92
3.9.5	Dedicated controller slot	96
3.9.6	External direct-attach storage connectivity	97
3.10	Optical drives	99
3.11	PCIe slots	99
3.12	I/O cards	101
3.12.1	Emulex 10 GbE Integrated Virtual Fabric Adapter II	101
3.12.2	Optional adapters	104

3.13	Standard onboard features	107
3.13.1	Onboard Ethernet	107
3.13.2	Environmental data	108
3.13.3	Integrated management module	108
3.13.4	Unified Extensible Firmware Interface	109
3.13.5	Integrated Trusted Platform Module	109
3.13.6	Light path diagnostics	110
3.14	Power supplies and fans of the x3850 X5 and MAX5	110
3.14.1	x3850 X5 power supplies and fans	110
3.14.2	MAX5 power supplies and fans	111
3.15	Integrated virtualization	112
3.16	Operating system support	112
3.17	Rack considerations	113
Chapter 4.	IBM System x3690 X5	115
4.1	Product features	116
4.1.1	System components	117
4.1.2	IBM MAX5 memory expansion unit	119
4.2	Target workloads	121
4.3	Models	122
4.3.1	Base x3690 X5 models with Intel Xeon E7 series processors	122
4.3.2	Workload-optimized x3690 X5 models with Xeon E7 series processors	124
4.4	System architecture	126
4.5	MAX5	126
4.6	Scalability	129
4.7	Processor options	130
4.8	Memory	131
4.8.1	x3690 X5 memory options	133
4.8.2	MAX5 memory options	134
4.8.3	x3690 X5 memory population order	136
4.8.4	MAX5 memory population order	139
4.8.5	Memory balance	140
4.8.6	Mixing DIMMs and the performance effect	141
4.8.7	Memory mirroring	142
4.8.8	Memory sparing	144
4.8.9	Effect on performance when you use mirroring or sparing	145
4.9	Storage	146
4.9.1	2.5-inch SAS drive support	146
4.9.2	IBM eXFlash and SSD 1.8-inch disk support	151
4.9.3	SAS and SSD controller summary	155
4.9.4	Battery backup placement	159
4.9.5	ServeRAID Expansion Adapter	160
4.9.6	Drive combinations	161
4.9.7	External direct-attach serial-attached SCSI storage	166
4.9.8	Optical drives	167
4.10	PCIe slots	167
4.10.1	Riser 1	168
4.10.2	Riser 2	169
4.10.3	Emulex 10 Gb Ethernet Adapter	169
4.10.4	I/O adapters	171
4.11	Standard features	173
4.11.1	Integrated management module	173
4.11.2	Ethernet subsystem	174

4.11.3	USB subsystem	174
4.11.4	Integrated Trusted Platform Module	175
4.11.5	Light path diagnostics	175
4.11.6	Cooling	175
4.12	Power supplies	177
4.12.1	x3690 X5 power subsystem	177
4.12.2	MAX5 power subsystem	178
4.13	Integrated virtualization	178
4.14	Supported operating systems	179
4.15	Rack mounting	180
Chapter 5.	IBM BladeCenter HX5	181
5.1	Introduction	182
5.2	Comparison between HS23 and HX5	184
5.3	Target workloads	185
5.4	Chassis support	186
5.5	HX5 models	187
5.5.1	Base models of machine type 7873	187
5.5.2	Two-node models of machine type 7873	189
5.5.3	Workload optimized models of machine type 7873	190
5.6	System architecture	192
5.7	Speed Burst Card	193
5.8	IBM MAX5 and MAX5 V2 for HX5	194
5.9	Scalability	196
5.9.1	Single HX5 configuration	197
5.9.2	Two-node HX5 configuration	197
5.9.3	HX5 with MAX5	198
5.10	Processor options	199
5.11	Memory	201
5.11.1	Memory options	202
5.11.2	Dual inline memory module population order	204
5.11.3	Memory balance	208
5.11.4	Memory mirroring	209
5.11.5	Memory sparing	210
5.11.6	Mirroring or sparing effect on performance	211
5.12	Storage	212
5.12.1	SSD Expansion Card	212
5.12.2	Solid-state drives for HX5	213
5.12.3	LSI SAS Configuration Utility for HX5	214
5.12.4	Determining which SSD RAID configuration to choose	216
5.12.5	Connecting to external SAS storage devices	216
5.13	BladeCenter PCI Express Gen 2 Expansion Blade II	217
5.13.1	PCIe SSD adapters	218
5.14	I/O expansion cards	219
5.14.1	CIOv	219
5.14.2	CFFh	220
5.15	Standard onboard features	222
5.15.1	Unified Extensible Firmware Interface	222
5.15.2	Onboard network adapters	222
5.15.3	Integrated management module	222
5.15.4	Video controller	223
5.15.5	Trusted Platform Module	223
5.16	Integrated virtualization	223

5.17 Partitioning capabilities	225
5.18 Operating system support	225
Chapter 6. Systems management	227
6.1 Management applications	228
6.2 Embedded firmware	228
6.3 Integrated management module	231
6.4 Firmware levels	232
6.5 UpdateXpress	232
6.6 Deployment tools	233
6.6.1 Bootable Media Creator	233
6.6.2 ServerGuide	234
6.6.3 ServerGuide Scripting Toolkit	234
6.6.4 IBM Start Now Advisor	236
6.7 Configuration utilities	236
6.7.1 MegaRAID Storage Manager	236
6.7.2 Advanced Settings Utility	237
6.7.3 Storage Configuration Manager	237
6.8 IBM Dynamic Systems Analysis	238
6.9 IBM Systems Director	239
6.9.1 Active Energy Manager	240
6.9.2 Tivoli Provisioning Manager for Operating System Deployment	241
Abbreviations and acronyms	243
Related publications	245
IBM Redbooks	245
Other publications	246
Online resources	247
How to get Redbooks	248
Help from IBM	248

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 grant 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.

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

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:

400®	IBM Systems Director Active Energy	RETAIN®
AIX®	Manager™	ServerProven®
BladeCenter®	IBM®	Smarter Planet®
BNT®	PowerPC®	System Storage®
Calibrated Vectors Cooling™	POWER®	System x®
Dynamic Infrastructure®	PureFlex™	System z®
eServer™	RackSwitch™	Tivoli®
Global Technology Services®	Redbooks®	X-Architecture®
GPFS™	Redpaper™	xSeries®
IBM Flex System™	Redbooks (logo)  ®	zEnterprise®

The following terms are trademarks of other companies:

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.

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

Java, and all Java-based trademarks and logos are trademarks or registered trademarks of Oracle and/or its affiliates.

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. 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 portfolio of high-end computing introduces the fifth generation of IBM X-Architecture® technology. The X5 portfolio is 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 IBM Redpaper™ publication introduces the new IBM eX5 portfolio and describes the technical detail behind each server. This document is intended for potential users of eX5 products that are seeking more information about the portfolio.

The team who wrote this paper

This edition of the paper was produced by a team of specialists from around the world working at the International Technical Support Organization (ITSO), Raleigh Center.



David Watts is a Consulting IT Specialist at the IBM ITSO Center in Raleigh. He manages residencies and produces IBM Redbooks® publications on hardware and software topics that are related to IBM Flex System™, IBM System x®, and BladeCenter® servers and associated client platforms. He authored over 200 books, papers, and product guides. He holds a Bachelor of Engineering degree from the University of Queensland (Australia), and has worked for IBM in both the United States and Australia since 1989. David is an IBM Certified IT Specialist, and a member of the IT Specialist Certification Review Board.



Duncan Furniss is a Certified Consulting IT Specialist for IBM in Canada. He provides technical sales support for IBM PureFlex™, System x, BladeCenter, and IBM System Storage® products. He co-authored six previous IBM Redbooks publications, the most recent is *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.

Thanks to the authors of the previous editions:

- ▶ David Watts
- ▶ Duncan Furniss
- ▶ Scott Haddow
- ▶ Jeneea Jervay
- ▶ Eric Kern
- ▶ Cynthia Knight

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

From IBM Marketing:

- ▶ Michelle Brunk
- ▶ Mark Cadiz
- ▶ Mark Chapman
- ▶ Randy Lundin
- ▶ Mike Talplacido
- ▶ David Tareen

From IBM Development:

- ▶ Ralph Begun
- ▶ Jon Bitner
- ▶ Charles Clifton
- ▶ Candice Coletrane-Pagan
- ▶ David Drez
- ▶ Royce Espy
- ▶ Larry Grasso
- ▶ Mark Kapoor
- ▶ Randy Kolvic
- ▶ Chris LeBlanc
- ▶ Greg Sellman
- ▶ Matthew Trzyna

From other IBMers throughout the world:

- ▶ Aaron Belisle, IBM US
- ▶ Randall Davis, IBM Australia
- ▶ Shannon Meier, IBM US
- ▶ Keith Ott, IBM US
- ▶ Andrew Spurgeon, IBM Australia and New Zealand

Now you can become a published author, too!

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

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

ibm.com/redbooks/residencies.html

Comments welcome

Your comments are important to us!

We want our papers to be as helpful as possible. Send us your comments about this paper 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/pages/IBM-Redbooks/178023492563?ref=ts>

- 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>

Summary of changes

This section describes the technical changes that are made in this edition of the paper and in previous editions. This edition might also include minor corrections and editorial changes that are not identified.

Summary of Changes
for IBM eX5 Portfolio Overview: IBM System x3850 X5, x3950 X5, x3690 X5, and
BladeCenter HX5
as created or updated on May 20, 2013.

These revisions reflect the addition, deletion, or modification of new and changed information. Numerous other smaller updates might have occurred that are not listed here.

April 2013, Sixth Edition

- ▶ x3850 X5 and x3950 X5
 - Added new models of machine type 7143
 - Removed withdrawn machine type 7145 with Intel Xeon 6500/7500 series processors
 - Updated supported options tables: memory, drives, adapters, virtualization keys
 - Updated list of supporting operating systems
- ▶ x3690 X5
 - Added new models of machine type 7147
 - Removed withdrawn machine type 7148 with Intel Xeon 6500/7500 series processors
 - Updated supported options tables: memory, drives, adapters, virtualization keys
 - Updated list of supporting operating systems
- ▶ BladeCenter HX5
 - Added new models of machine type 7873
 - Removed withdrawn machine type 7872 with Intel Xeon 6500/7500 series processors
 - Updated supported options tables: memory, drives, adapters, virtualization keys
 - Updated list of supporting operating systems

December 2011, Fifth Edition

New information

- ▶ Technology
 - Intel Xeon processor E7 family (“Westmere EX”)
 - Larger memory capacities with 32 GB dual inline memory modules (DIMMs)
 - New scalability configurations
 - New partitioning configurations
 - New 200 GB solid-state drives
- ▶ x3850 X5 and x3950 X5:
 - New machine type 7143 with the Intel Xeon processor E7 family
 - New MAX5 V2 with support for 1.35 V DIMMs and 32 GB DIMMs
 - MAX5 V2 supported on machine type 7145

- MAX5 (V1) supported on machine type 7143
- MAX5 and MAX5 V2 shipping with both power supplies standard
- Models of type 7143 including Emulex 10 GbE Integrated Virtual Fabric Adapter II
- New standard models
- New workload-optimized models
- Support for two-node plus MAX5 scalability using EXA cabling
- Support for partitioning
- New Intel Xeon E7 processor options
- New memory expansion card for use with systems with E7 processors
- New 1.35 V low-voltage (PC3L) memory options
- New 32 GB memory DIMM option
- New serial-attached SAS drive options
- New solid-state drive (SSD) options
- New integrated virtualization options
- ▶ x3690 X5:
 - New machine type 7147 with Intel Xeon E7 processors
 - New MAX5 V2 with support for 1.35 V DIMMs and 32 GB DIMMs
 - MAX5 V2 supported on machine type 7148
 - MAX5 (V1) supported on machine type 7147
 - MAX5 and MAX5 V2 now ship with both power supplies standard
 - Models of type 7147 include Emulex 10 GbE Integrated Virtual Fabric Adapter II
 - New standard models
 - New workload-optimized models
 - New Intel Xeon E7 processor options
 - New memory mezzanine for use with systems with E7 processors
 - New 1.35 V low-voltage (PC3L) memory options
 - New 32 GB memory DIMM option
 - New SAS drive options
 - New SSD options
 - New integrated virtualization options
- ▶ HX5:
 - New machine type 7873 with Intel Xeon E7 processors
 - New MAX5 V2 with support for low-voltage DIMMs
 - New standard models
 - New workload-optimized models
 - New Intel Xeon E7 processor options
 - New 16 GB memory DIMM option
 - New 1.35 V low-voltage memory options
 - New SSD options including a 200 GB solid-state drive
 - New integrated virtualization options

Changes to existing information

- ▶ Updated lists of supported adapters
- ▶ Updated lists of supported operating systems

October 2010, Fourth Edition

New information

- ▶ IBM eX5 announcements on August 31, 2010

- ▶ x3850 X5 and x3950 X5:
 - New virtualization workload-optimized model of the x3850 X5, 7145-4Dx
 - New memory options for the x3850 X5 and MAX5
 - IBM USB Memory Key for VMware ESXi 4.1 with MAX5, for x3850 X5
- ▶ x3690 X5:
 - MAX5 memory expansion unit
 - New models of the x3690 X5, which include the MAX5
 - New virtualization workload-optimized model of the x3690 X5, 7148-2Dx
 - New memory options for the x3690 X5 and MAX5
 - IBM USB Memory Key for VMware ESXi 4.1 with MAX5, for x3690 X5
 - The use of VMware on a two-processor x3690 X5 requires the memory mezzanine
- ▶ HX5:
 - MAX5 memory expansion blade
 - Additional chassis support information
 - New IBM HX5 MAX5 1-node Scalability Kit
 - New Intel Xeon 130 W processor options
 - New models with MAX5 memory expansion blades standard
 - New model with a Intel Xeon 130 W processor standard
 - New virtualization workload-optimized model
 - HX5+MAX5 system architecture
 - MAX5 memory rules and population order
 - New IBM USB Memory Key for VMware ESXi 4.1 option

Changes to existing information

- ▶ Corrected x3690 X5 physical dimensions
- ▶ For VMware vSphere support, MAX5 requires vSphere 4.1 or later
- ▶ Clarified VMware ESX and ESXi on the HX5

August 2010, Third Edition

New information: IBM System x3690 X5

June 2010, Second Edition

New information

- ▶ IBM eX5 announcements on May 18, 2010
- ▶ MAX5 memory expansion unit product information
- ▶ Models of the x3850 X5 that include MAX5
- ▶ Additional two-node and MAX5 scalability information
- ▶ x3850 X5 memory placement
- ▶ Hemisphere mode
- ▶ MAX5 memory placement
- ▶ x3850 X5 memory performance
- ▶ ServeRAID B5015 SSD controller
- ▶ Support for the ServeRAID M5014 controller
- ▶ ServeRAID M5015 does not include a battery
- ▶ Support for IBM BNT® SFP+ Transceiver, 46C3447
- ▶ MAX5 power supplies and fans



Introduction

The IBM eX5 product portfolio represents the fifth generation of servers that are built upon Enterprise X-Architecture. Enterprise X-Architecture is the culmination of generations of IBM technology and innovation that is derived from our experience in high-end enterprise servers. Now, with eX5, IBM scalable systems technology for Intel processor-based servers has also come 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 IBM 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 that they supersede. Power efficiency and server 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 used, yielding the best return from your application investment. These systems allow your enterprise to grow in processing, input/output (I/O), and memory dimensions. Therefore, you can provision what you need now and expand the system to meet future requirements. System redundancy and availability technologies are more advanced than those previously available in the x86 systems.

The servers in the eX5 product portfolio are based on the Intel Xeon processor E7-8800/4800/2800 product families. With the inclusion of these processors, the eX5 servers are faster, more reliable, and more power-efficient. As with previous generations of IBM Enterprise X-Architecture systems, these servers deliver many class-leading benchmarks, including the highest TPC-E benchmark result for a system of any architecture.

The following topics are covered:

- ▶ 1.1, “eX5 systems” on page 2
- ▶ 1.2, “Positioning” on page 3
- ▶ 1.3, “Energy efficiency” on page 6
- ▶ 1.4, “Services offerings” on page 7

1.1 eX5 systems

The four systems in the eX5 family are the IBM System x3850 X5, x3950 X5, x3690 X5, and the IBM BladeCenter HX5. The eX5 technology is primarily designed around three major workloads: Database servers, server consolidation that uses virtualization services, and Enterprise Resource Planning (application and database) servers. Each system can scale with more 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 servers together to form a single system.

Figure 1-1 shows the IBM eX5 family.



Figure 1-1 eX5 family (top to bottom): BladeCenter HX5 (two-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 *IBM System 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 usage, reliability, and performance for computer-intensive and memory-intensive workloads. These servers can be connected together to form a single system with twice the resources, or to support memory scaling with the attachment of a MAX5. With the Intel Xeon processor E7 family, the x3850 X5 and x3950 X5 can scale to a two-server plus two-MAX5 configuration.

The *IBM System x3690 X5* is a 2U rack-optimized server. This server brings features and performance to the middle tier and 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 two-node system with four processors, or a single-node system with the MAX5 memory expansion blade.

When compared to other servers in the System x portfolio, these systems represent the upper end of the spectrum. These servers are suited for the most demanding x86 tasks, and can handle jobs that previously might have run on other platforms. To assist with selecting the

ideal system for a specified workload, workload-specific models for virtualization and database needs have been designed.

1.2 Positioning

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

Table 1-1 Maximum configurations for the eX5 systems

Maximum configurations		x3850 X5 / x3950 X5	x3690 X5	HX5
Processors	One-node	4	2	2
	Two-node	8	Not available	4
Memory	One-node	2 TB (64 DIMMs) ^a	1 TB (32 DIMMs) ^b	512 GB (16 DIMMs)
	One-node with MAX5	3 TB (96 DIMMs) ^a	2 TB (64 DIMMs) ^b	1.25 TB (40 DIMMs)
	Two-node	4 TB (128 DIMMs) ^a	Not available	1 TB (32 DIMMs)
	Two-node with MAX5	6 TB (192 DIMMs) ^a	Not available	Not available
Disk drives (non-SSD) ^c	One-node	8	16	Not available
	Two-node	16	Not available	Not available
SSDs	One-node	16	24	2
	Two-node	32	Not available	4
Standard 1 Gb Ethernet interfaces	One-node	2	2	2
	Two-node	4	Not available	4
Standard 10 Gb Ethernet interfaces	One-node	2 ^d	2 ^d	0
	Two-node	4	Not available	0

a. Requires full processors 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, extra backplanes might be needed to support these numbers of drives.

d. Standard on most models.

1.2.1 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 2 TB of RAM in a single-node environment.

The x3850 or x3950 X5 with the MAX5 memory expansion unit attached, as shown in Figure 1-2, can add up to an extra 1 TB of RAM for a total of 3 TB of memory.



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

Two x3850 or x3950 X5 servers with 2 MAX5 memory expansion units can be connected for a single system image with eight processors and 6 TB of RAM.

Table 1-2 compares the number sockets, cores, and memory capacity of the mid-range, Intel Xeon E5-4600 processor-based four socket server and the eX5 systems.

Table 1-2 Comparing the x3750 M4 with the eX5 servers

	Processor sockets	Processor cores	Maximum memory
Mid-range four socket server with Intel Xeon E5-4600 processors			
x3750 M4	4	32	1.5 TB
Next generation server (eX5) with Intel Xeon processor E7 family			
x3850 and x3950 X5	4	40	2 TB
x3850 and x3950 X5 two-node	8	80	4 TB
x3850 and x3950 X5 with MAX5	4	40	3 TB
x3850 and x3950 X5 two-node with MAX5	8	80	6 TB

1.2.2 IBM System x3690 X5

The x3690 X5, as shown in Figure 1-3, is a two-processor server that exceeds the capabilities of the current mid-tier server, the x3650 M4. 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 1 TB of RAM, whereas the x3650 M3 has a maximum memory capacity of 768 GB.



Figure 1-3 x3690 X5

Table 1-3 compares the processing and memory capacities of the x3650 M4 and the x3690 X5.

Table 1-3 x3650 M4 compared to x3690 X5

	Processor sockets	Processor cores	Maximum memory
Mid-tier server with Intel Xeon E5-2600 processors			
x3650 M4	2	16	768 GB
High-end (eX5) server with Intel Xeon E7 family processor			
x3690 X5	2	20	1 TB ^a
x3690 with MAX5	2	20	2 TB ^a

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

1.2.3 IBM BladeCenter HX5

The IBM BladeCenter HX5 is shown in Figure 1-4 in a two node configuration. HX5 is a blade that exceeds the capabilities of the Intel Xeon E5 based system, the HS23.



Figure 1-4 Blade HX5 dual scaled

Table 1-4 compares these blades.

Table 1-4 HS23 and HX5 compared

	Processor sockets	Processor cores	Maximum memory
BladeCenter server with Intel Xeon E5-2600 processors			
HS23 (30 mm)	2	16	512 GB
IBM eX5 blade server with Intel Xeon processor E7 family			
HX5 (30 mm)	2	20	512 GB
HX5 two-node (60 mm)	4	40	1 TB
HX5 with MAX5 (60 mm)	2	20	1.25 TB

1.3 Energy efficiency

IBM put extensive engineering effort into keeping your energy bills low, from high-efficiency power supplies and fans to lower-draw processors, memory, and solid-state drives (SSDs). IBM strives to reduce the power that is consumed by the systems to the extent that we include *altimeters*. These altimeters can measure the density of the atmosphere in the servers and then adjust the fan speeds accordingly for optimal cooling efficiency.

Technologies like these altimeters, along with the Intel Xeon processor E7 family that intelligently adjust their voltage and frequency, help take costs out of IT in several ways:

- Eight-core processors that are 95 W use 27% less energy than 130 W processors. The Intel Xeon E7 processors provide 25% more cores, threads, and last-level cache with the same thermal design profile (TDP) as the 6500 and 7500 series. The new processor cores can independently shut down to 0 W when idle, and the entire processor can reach near 0 W at idle.

- ▶ DDR3 DIMMs that are 1.5 V use 10 - 15% less energy than the DDR2 DIMMs that were used in older servers. The Intel Xeon processor E7 family supports low voltage (1.35 V) DIMMs, using 10% less power than 1.5 V DDR3 DIMMs. The memory buffers for the newer processors draw 1.3 - 3 W less power, depending on load.
- ▶ SSDs use up to 80% less energy than 2.5-inch HDDs and up to 88% less energy than 3.5-inch HDDs.
- ▶ If there is 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, 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.

1.4 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 offers six products 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. The consultants also 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 an 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>



IBM eX5 technology

This chapter describes the technology that IBM brings to the IBM eX5 portfolio of servers.

The fifth generation of IBM Enterprise X-Architecture (EXA) chip sets, called *eX5*, is described. 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 Intel Xeon processors that are used in the eX5 servers, the Intel Xeon processor E7 product families (“Westmere EX”), are described next.

We then describe the 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. Integrated virtualization is also described.

The following topics are covered:

- ▶ 2.1, “eX5 chip set” on page 10
- ▶ 2.2, “Intel Xeon processors” on page 10
- ▶ 2.3, “Memory” on page 16
- ▶ 2.4, “MAX5” on page 26
- ▶ 2.5, “Scalability” on page 28
- ▶ 2.6, “Partitioning” on page 30
- ▶ 2.7, “Unified Extensible Firmware Interface system settings” on page 32
- ▶ 2.8, “IBM eXFlash” on page 38
- ▶ 2.9, “Integrated virtualization” on page 41

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 use the increased computational ability of the new generation of Intel processors, eX5 provides more memory capacity and more scalable memory interconnects (SMIs), increasing bandwidth to memory. eX5 also provides the following reliability, availability, and serviceability (RAS) capabilities for memory:

- ▶ Chipkill
- ▶ Memory ProteXion
- ▶ Full Array Memory Mirroring

QPI uses a source *snoop protocol*. This technique means that even if a CPU knows that another processor has a cache line that it wants (the cache line address is in the *snoop filter* and in the shared state), it must request a copy of the cache line. The CPU has to 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 snooping and the source snoop protocol, see 2.2.6, “QuickPath Interconnect” on page 13.

Memory that is directly controlled by a processor can be accessed more quickly than through the eX5 chip set. However, it is connected to all processors and introduces less delay than accesses to memory controlled by another processor in the system.

The eX5 chip set also has, as with previous generations, connectors to allow systems to scale beyond the capabilities that are provided by the Intel chip sets. We call this scaling, *Enterprise X-Architecture (EXA) scaling*. With EXA scaling, you can connect two x3850 X5 servers and two MAX5 memory expansion units together to form a single system image with up to eight Intel Xeon E7 processors and up to 6 TB of RAM. We introduce MAX5 in 2.4, “MAX5” on page 26.

2.2 Intel Xeon processors

The current models of the eX5 systems use Intel Xeon E7 processors. Earlier models used Intel Xeon 7500 or 6500 series processors. The processor families are now introduced. The main features of the processors are then described.

2.2.1 Intel Xeon E7 processors

The Intel Xeon processor E7 family that is used in the eX5 systems are follow-ons to the Intel Xeon processor 7500 series and 6500 series. Although the processor architecture is largely unchanged, the lithography size was reduced from 45 nm to 32 nm. This change allows for more cores (and thus more threads with Hyper-Threading Technology) and more last-level

cache, although staying within the same thermal design profile (TDP) and physical package size.

There are three groups of the Intel Xeon processor E7 family that support scaling to separate levels:

- ▶ The Intel Xeon processor E7-2800 product family is used in the x3690 X5 and BladeCenter HX5. This series supports only two processor configurations, so it cannot be used in a two-node HX5 configuration. Most processors in this family support connection to a MAX5.
- ▶ The Intel Xeon processor E7-4800 product family is primarily used in the x3850 X5 and the HX5. This series supports four-processor configurations, so it can be used for two-node HX5s. All of the E7-4800 family support connection to a MAX5 and can also be used for two-node x3850 X5s with MAX5 configurations. Such configurations use EXA scaling, which the E7-4800 processors support.

However, two-node x3850 X5 configurations *without* MAX5 cannot use E7-4800 processors because such configurations require QPI scaling, which E7-4800 processors do not support. A 4800 family processor is available for the x3690 X5 because of its low-power rating.

- ▶ The Intel Xeon processor E7-8800 product family is used in the x3850 X5 to scale to two nodes without MAX5s. Specific high-frequency and low-power models of this processor are available for the x3690 X5 and HX5 as well.

These scalability capabilities are summarized in Table 2-1.

Table 2-1 Comparing the scalability features of the Intel Xeon processor E7 family

	E7-2800	E7-4800	E7-8800
x3690	Yes	Yes	Yes
x3690 X5 with MAX5	Yes ^a	Yes	Yes
HX5	Yes	Yes	Yes
HX5 with MAX5	Yes ^a	Yes	Yes
HX5 two-node	Not supported	Yes	Yes
x3850 X5	Not supported	Yes	Yes
x3850 X5 with MAX5	Not supported	Yes	Yes
x3850 X5 two-node without MAX5	Not supported	Not supported	Yes
x3850 X5 two-node with MAX5	Not supported	Yes (EXA scaling)	Yes (EXA scaling)

a. E7-2803 and E7-2820 processors do not support MAX5

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

- ▶ IBM System x3850 X5: 3.7, “Processor options” on page 68
- ▶ IBM System x3690 X5: 4.7, “Processor options” on page 130
- ▶ IBM BladeCenter HX5: 5.10, “Processor options” on page 199

2.2.2 Intel Advanced Encryption Standard - New Instructions

Advanced Encryption Standard (AES) is an encryption standard that is widely used to protect network traffic and sensitive data. Advanced Encryption Standard - New Instructions

(AES-NI), available with the E7 processors, implements certain complex and performance intensive steps of the AES algorithm by using processor hardware. AES-NI can accelerate the performance and improve the security of an implementation of AES over an implementation that is completely performed by software.

For more information about Intel AES-NI, visit the following web page:

<http://software.intel.com/en-us/articles/intel-advanced-encryption-standard-instructions-aes-ni>

2.2.3 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. Enhancements include 64-bit guest OS support.

Intel VT Flex Priority optimizes virtualization software efficiency by improving interrupt handling. Intel VT Flex migration enables the eX5 servers to be added to existing virtualization pools with single, two, four, or eight-socket servers.

For more information about Intel Virtualization Technology, visit the following web page:

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

2.2.4 Hyper-Threading Technology

Intel Hyper-Threading Technology enables a single physical processor to run two separate code streams (threads) concurrently. To the operating system, a processor core with Hyper-Threading is seen as two logical processors. Each processor 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 run 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. This process is done by using 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. Application types that make the best use of Hyper-Threading are virtualization, databases, email, Java, and web servers.

For more information about Hyper-Threading Technology, visit the following web page:

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

2.2.5 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, a 2.26 GHz eight-core processor can run with two cores that are shut off and six cores active 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 this feature. Turbo Boost Technology can be engaged with any number of cores that are 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 eight-core processors, the number of frequency increments is expressed as four numbers that are separated by slashes. The first digit is for when seven or eight cores are active, the next is for when five or six cores are active, the next is for when three or four cores are active, and the last is 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 reevaluates. At any specified time, all active cores run at the same frequency.

For more information about Turbo Boost Technology, visit the following web page:

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

2.2.6 QuickPath Interconnect

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

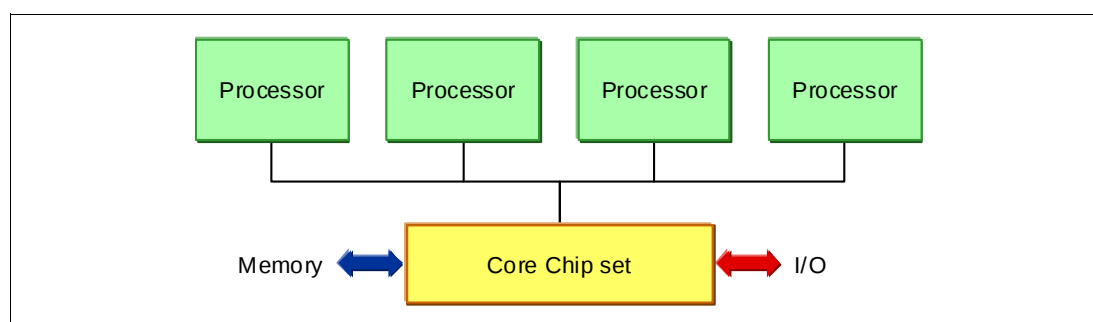


Figure 2-1 Shared front-side bus in the IBM x360 and x440 with snoop filter in the x365 and x445

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 that is stored in its cache. This process is described as *snooping* the other processor's caches, and it puts much 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 tracks 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, along with 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).

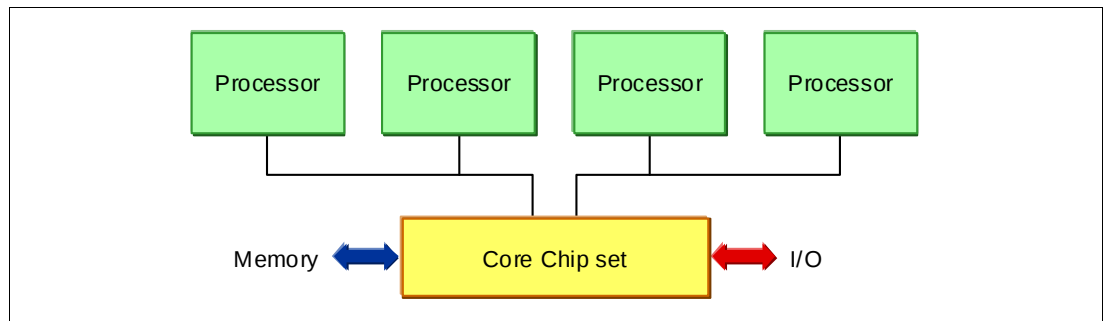


Figure 2-2 Dual independent buses, as in the x366 and x460 (later called the x3850 and x3950)

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, which are shown in Figure 2-3. Servers implementing this design included the IBM System x3850 M2 and x3950 M2.

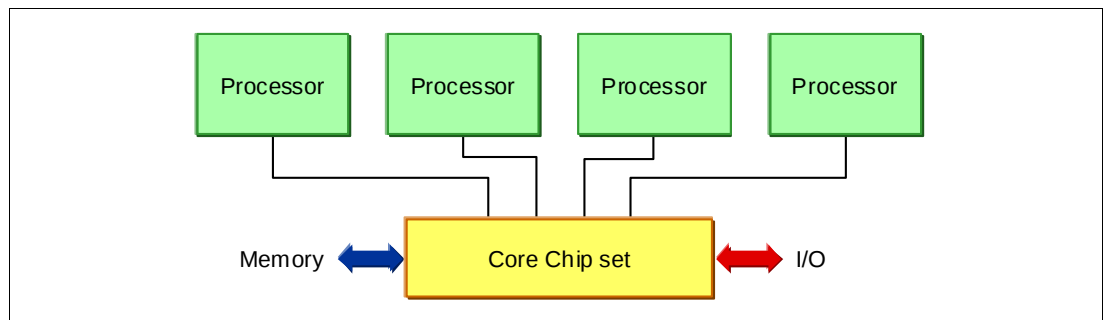


Figure 2-3 Independent processor buses, as in the x3850 M2 and x3950 M2

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 that are 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 configuration.

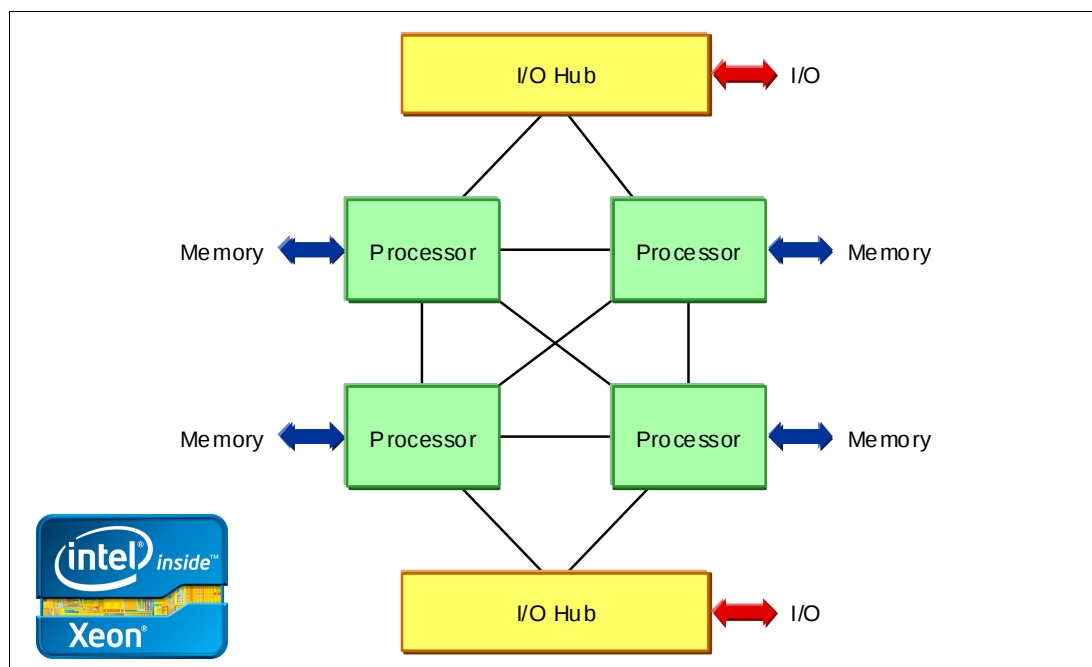


Figure 2-4 Figure 2-4 QPI, as used in the eX5 portfolio

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. This system is where part of the memory is directly connected to the processor where a specified 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 modified the MESI cache coherence protocol to include a forwarding state. Therefore, when a processor asks to copy a shared cache line, only one other processor responds.

For more information about QPI, visit the following web page:

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

2.2.7 Processor performance in a green world

All eX5 servers from the factory are designed to use power by the most efficient means possible. Controlling how much power the server is going to use is managed by controlling the core frequency and power that is applied to the processors. This process controls the frequency and power that is applied to the memory and reduces 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 attempt to manage power consumption at the operating system level. In this environment, the operating system decides 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 about moving virtual servers from one host to another so that it can 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. This practice is 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 uses more power.

The solution for virtual server clusters is to turn off the power management features of the host servers. To change the hardware-controlled power management on the F1-Setup page during power-on self-test (POST), 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.

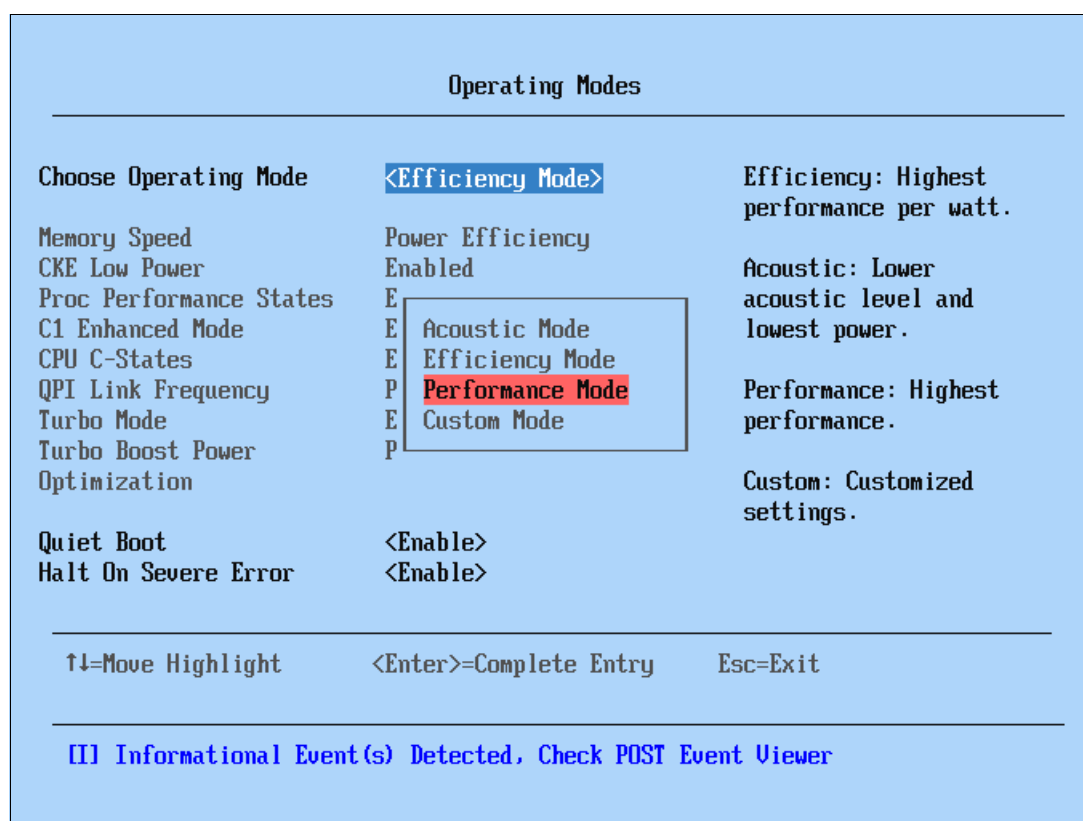


Figure 2-5 Setup (F1) → System Settings → Operating Modes to set Performance Mode

2.3 Memory

The major features of the memory subsystem in eX5 systems are now described. The following topics are covered:

- ▶ 2.3.1, “Memory speed” on page 17
- ▶ 2.3.2, “Memory dual inline memory module placement” on page 18
- ▶ 2.3.3, “Memory ranking” on page 19
- ▶ 2.3.4, “Non-uniform memory access architecture” on page 21
- ▶ 2.3.5, “Hemisphere mode” on page 22
- ▶ 2.3.6, “Reliability, availability, and serviceability features” on page 23

- ▶ 2.3.7, “Scalable memory buffers” on page 26
- ▶ 2.3.8, “I/O hubs” on page 26

2.3.1 Memory speed

The speed at which the memory in the eX5 servers runs depends on the capabilities of the specific processors selected. With these servers, 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.

SMI link speed

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

- ▶ QPI link speed of 6.4 gigatransfers per second (GT/s) can run memory speeds up to 1066 MHz
- ▶ QPI link speed of 5.86 GT/s can run memory speeds up to 978 MHz
- ▶ QPI link speed of 4.8 GT/s can run 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 specific 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 be used only 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 that is associated with that CPU slot.

Memory bus speed

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 something 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 in turn 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 in turn 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.

Maximum memory speed: The maximum memory speed that is supported by the processors that are used in the eX5 systems is 1066 MHz. Although DIMMs rated for 1333 MHz are supported, they operate at a maximum speed of 1066 MHz in the eX5 servers.

Memory performance test on various memory speeds

Based on benchmarks by using an IBM internal load generator that is run on an x3850 X5 system that is 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 dual inline memory module placement

The eX5 servers support various ways to install memory DIMMs, which are described 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 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).
- ▶ Populating only one memory card per socket can result in approximately a 50% performance degradation (compare configurations 1 and 5).
- ▶ Memory performance is better if you install DIMMs on all memory channels than if you leave any memory channels empty (compare configurations 2 and 3).
- ▶ Two DIMMs per channel result in better performance than one DIMM per channel (compare configurations 1 and 2, and 5 and 6).

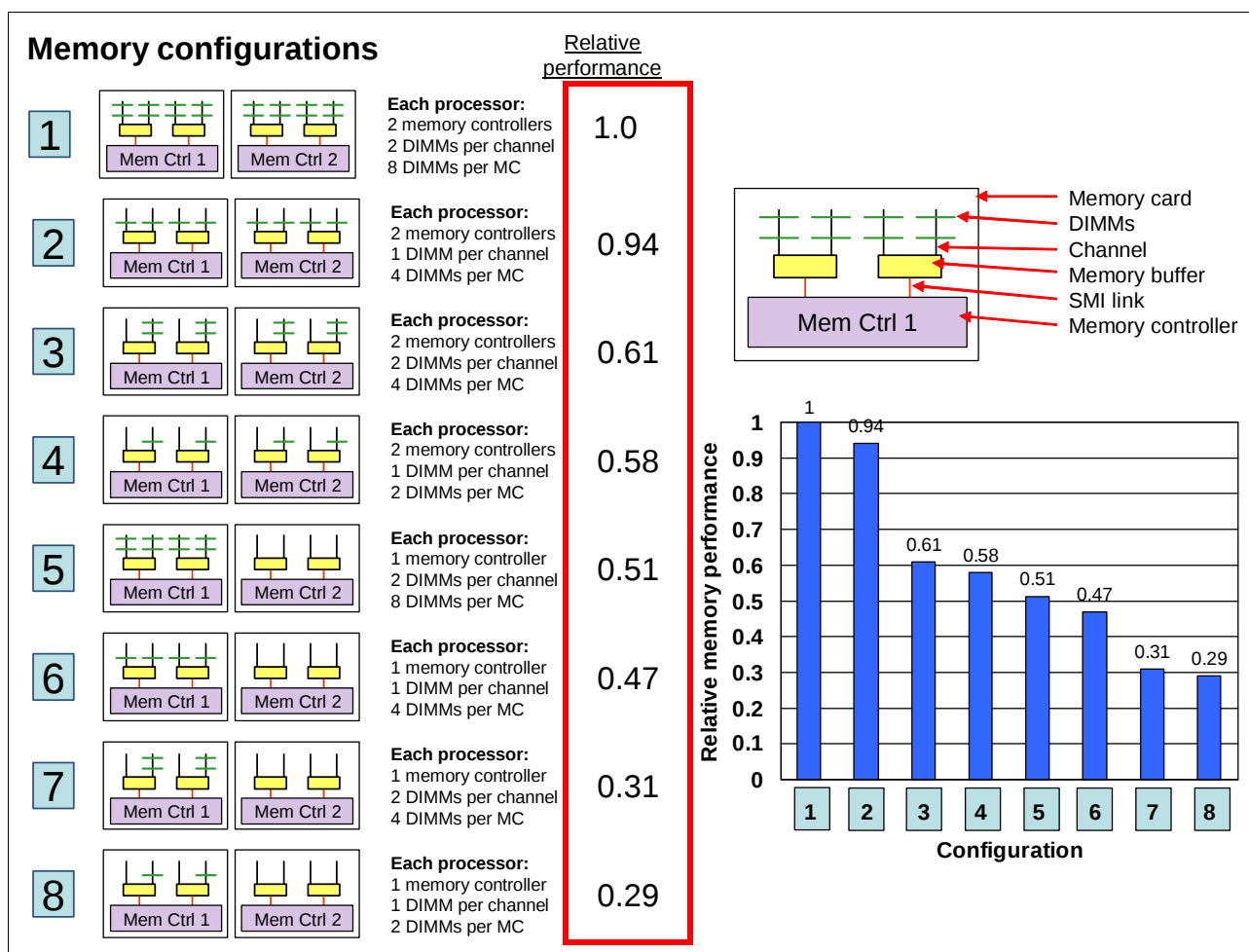


Figure 2-6 Relative memory performance based on DIMM placement (one processor and two memory cards shown)

2.3.3 Memory ranking

The underlying speed of the memory as measured in MHz is *not* sensitive to memory population. (In Intel Xeon 5600 processor-based systems, such as the x3650 M3, if rules regarding optimal memory population are not followed, the system basic input/output system (BIOS) clocks the memory subsystem down to a slower speed. This scenario is not the case with the x3850 X5.)

Number of ranks

Unlike Intel 5600 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 5600-based systems, such as the x3650 M3, 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 eX5 server 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, making the pages effectively larger and therefore creating more page-hit cycles.

Three types of memory DIMMs were used for this analysis:

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

The following memory configurations were used:

- ▶ 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 3-18 on page 70)
- ▶ 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 that were gathered by using the industry-standard STREAM benchmark, as shown in Figure 2-7.

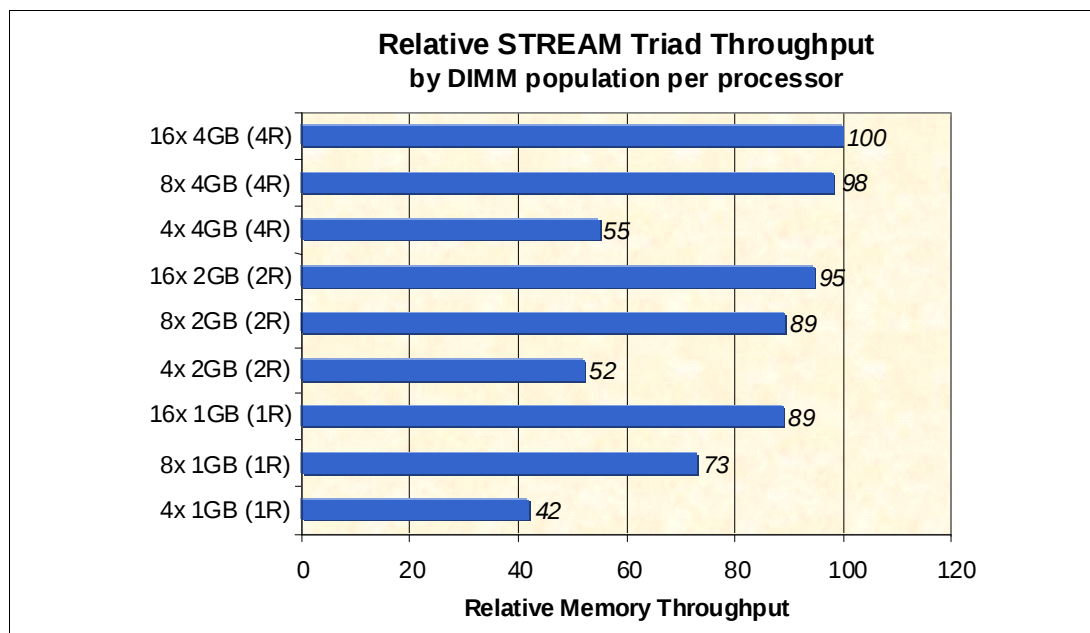


Figure 2-7 Comparing the performance of memory DIMM configurations by using STREAM

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, performance drops to 89% of the top performance with 16x 1 GB single-rank DIMMs.

You can see similar effects across the three configurations that are based on eight DIMMs per processor and four DIMMs per processor. These results also emphasize the same effect that is shown in 3.8.5, “Maximizing memory performance” on page 79 for the x3850 X5. This is where performance drops away dramatically when all eight memory channels per CPU are not used.

Additional ranks: Additional ranks increase the memory bus loading, which is why on Xeon 5600 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 processors avoids this slowdown.

2.3.4 Non-uniform memory access architecture

Non-uniform memory access (NUMA) architecture is an important consideration when you configure 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. This is true 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: 3.8.4, “DIMM population sequence” on page 74
- ▶ IBM System x3690 X5: 4.8.3, “x3690 X5 memory population order” on page 136
- ▶ IBM BladeCenter HX5: 5.11.2, “Dual inline memory module population order” on page 204

2.3.5 Hemisphere mode

Hemisphere mode is an important performance optimization of the Xeon E7 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. Hemisphere mode 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 two-node configurations on x3850 X5.

Hemisphere mode is a global parameter that is set at the system level. This configuration means that if even one processor’s memory is incorrectly configured, the entire system loses 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. The eX5 server 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). Therefore, 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 must 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 eight 4 GB DIMMs and processor 1 is configured with eight 2 GB DIMMs. Depending on the application characteristics, however, an unbalanced memory configuration can cause reduced performance. This outcome is because it forces a larger number of remote memory requests over the inter-CPU QPI links to the processors with more memory.

In summary:

- ▶ To enable hemisphere mode, each memory channel must contain at least one DIMM.
- ▶ On an x3850 X5 or x3690 X5, this means that 8 or 16 DIMMs must be installed for each processor.

- On a BladeCenter HX5, this means that exactly 8 DIMMs must be installed for each processor.

Industry-standard tests that are run on one processor with various memory configurations show 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.

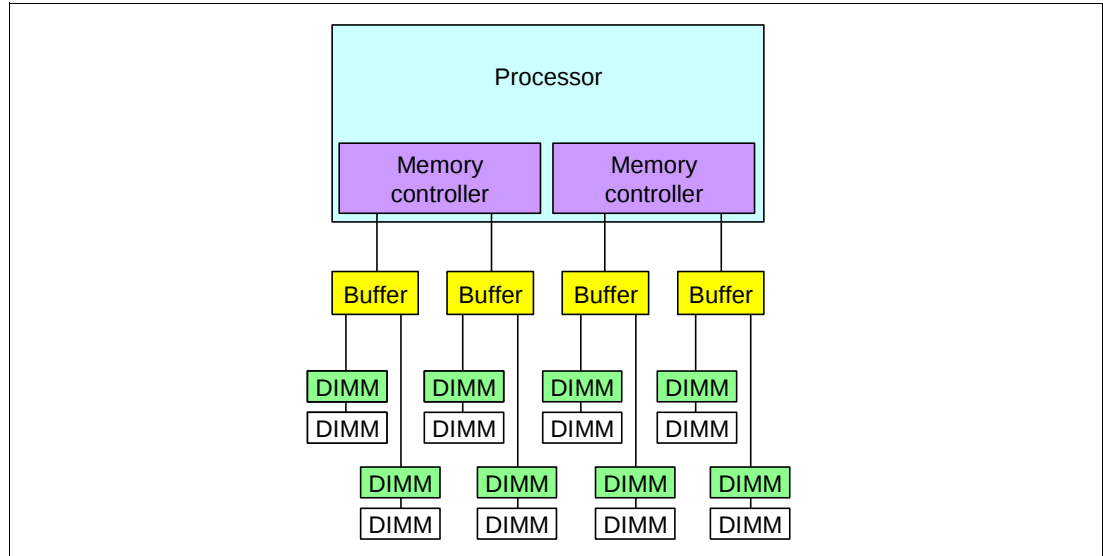


Figure 2-8 Example memory configuration

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

- IBM System x3850 X5: 3.8.4, “DIMM population sequence” on page 74
- IBM System x3690 X5: 4.8.3, “x3690 X5 memory population order” on page 136
- IBM BladeCenter HX5: 5.11.2, “Dual inline memory module population order” on page 204

2.3.6 Reliability, availability, and serviceability features

In addition to hemisphere mode, DIMM balance, and memory size, memory performance is also affected by the various memory 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 eX5 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 processors that are used in these systems. Instead, it provides a brief introduction to each mode and its corresponding performance effects.

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

For more information, see *Reliability, Availability, and Serviceability Features of the IBM eX5 Portfolio*, REDP-4864, available from:

<http://www.redbooks.ibm.com/abstracts/redp4864.html>

Memory mirroring

To improve memory reliability and availability beyond error correction code (ECC) and Chipkill (see “Chipkill” on page 25), the chip set can mirror memory data on two memory channels. 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 channels on every memory-write cycle. Two copies of the data are kept, similar to the way a Redundant Array of Independent Disks mirror (RAID-1) writes to disk. Reads are interleaved between memory channels. The system automatically uses the most reliable memory channel as determined by error logging and monitoring.

If errors occur, only the alternate memory channel is used until bad memory is replaced. Because a redundant copy is kept, mirroring results in only half the installed memory being 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 32 GB of identical memory equally and symmetrically across the two memory channels to achieve 16 GB of mirrored memory. FAMM enables other enhanced memory features, such as unrecoverable error (UE) recovery. Memory mirroring is independent of the operating system.

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

- ▶ x3850 X5: 3.8.6, “Memory mirroring” on page 82
- ▶ x3690 X5: 4.8.7, “Memory mirroring” on page 142
- ▶ BladeCenter HX5: 5.11.4, “Memory mirroring” on page 209

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, and the spare counterpart is 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 must 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.

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

- ▶ IBM System x3850 X5: 3.8.7, “Memory sparing” on page 84

- ▶ IBM System x3690 X5: 4.8.8, “Memory sparing” on page 144
- ▶ IBM BladeCenter HX5: 5.11.5, “Memory sparing” on page 210

Chipkill

Chipkill memory technology, an advanced form of ECC from IBM, is available for the eX5 servers. 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 and double device data correction

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 fails and when to route the data around the failed chip.

The eX5 servers with the E7 processors support the Intel implementation of RBS, which they call *double device data correction* (DDDC). RBS is automatically enabled in the MAX5 memory port if all DIMMs installed to that memory port are x4 DIMMs. The x8 DIMMs do not support RBS.

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. If 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 that use 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 are generated on that DIMM.

Lock step

IBM eX5 memory can operate in *lock step* mode. Lock step is a memory protection feature that involves the pairing of two memory DIMMs. The paired DIMMs can do the same operations, and the results are compared. If any discrepancies exist between the results, a memory error is signaled. As an example, 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 by using 8 GB DIMMs.

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

Machine Check Architecture

Machine Check Architecture (MCA) is a RAS feature that previously was only available for other processor architectures, such as Intel Itanium, IBM POWER®, and mainframes. Implementation of the MCA requires hardware support, firmware support, such as 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 announced support for the Intel MCA on the Xeon E7 processors.

2.3.7 Scalable memory buffers

Unlike the Xeon E5 series processors, which use unbuffered memory channels, the processors in the eX5 systems use scalable memory buffers (SMBs) in the systems design. This approach reflects the various workloads for which these processors were intended. These processors are designed for workloads that require more memory, such as virtualization and databases. The use of SMBs allows more memory per processor and prevents memory bandwidth reductions when more memory is added per processor.

The SMBs for the E7 processor family enable support for 32 GB DIMMs and low voltage (1.35 V) DIMMs. These SMBs are more power-efficient, which means that all eX5 systems can operate memory at the maximum speed as dictated by the processor.

2.3.8 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. The hubs then connect to the processors through QPI links.

Figure 2-4 on page 15 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 about 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 101.
- ▶ IBM System x3690 X5: 4.10.4, “I/O adapters” on page 171.
- ▶ IBM BladeCenter HX5: 5.14, “I/O expansion cards” on page 219.

2.4 MAX5

Memory Access for eX5 (MAX5) is the name given 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) takes 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 blade bay).

For the E7 processor-based systems, there is a new version of the MAX5 called *MAX5 V2*. MAX5 V2 has the newer scalable memory buffers, so it supports higher-density DIMMs and low voltage memory.

Figure 2-9 shows an HX5 with a 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 (1U unit beneath the main system)

Figure 2-11 shows the MAX5 for System x removed from the housing.

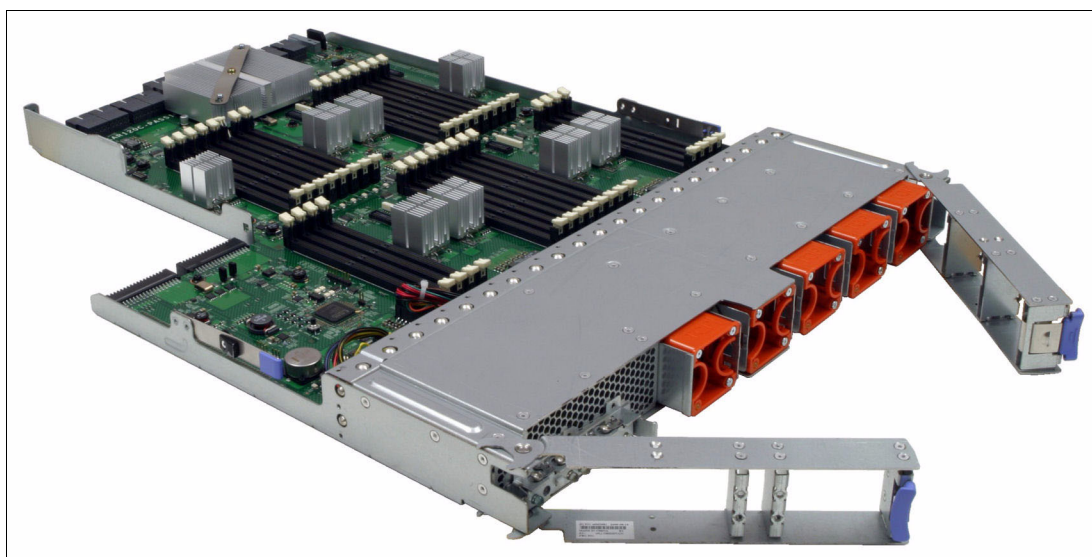


Figure 2-11 IBM MAX5 for the x3850 X5 and x3690 X5

MAX5 connects to these systems through QPI links and provides the EXA scalability interfaces. The eX5 chip set, described in 2.1, “eX5 chip set” on page 10, is contained in the MAX5 units.

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

Table 2-2 Memory capacity when MAX5 is used

	Memory capacity without MAX5	Memory capacity with MAX5
x3850 X5 two-node	4 TB	6TB
x3850 X5 single-node	2TB	3TB
x3690 X5	1TB	2TB
HX5 two-node	512 GB	Not available
HX5 single-node	256 GB	640 GB

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

- ▶ IBM System x3850 X5: “Memory DIMMs for MAX5” on page 72
- ▶ IBM System x3690 X5: 4.8.4, “MAX5 memory population order” on page 139
- ▶ IBM BladeCenter HX5: “MAX5 memory population order” on page 206

2.5 Scalability

As shown in Figure 2-12 on page 29, eX5 servers allow the following types of scaling:

- ▶ Memory scaling: A MAX5 unit can attach to an eX5 server through QPI link cables. This method provides the server with more 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.
- EXA scaling: Two servers, each with a MAX5 unit attached, can connect to form a single system image. The connections are formed by using EXA link cables, which are attached to the EXA link ports on the MAX5 units. This capability is unique to the x3850 X5s.

See Figure 2-12 for the types of scaling with eX5 systems.

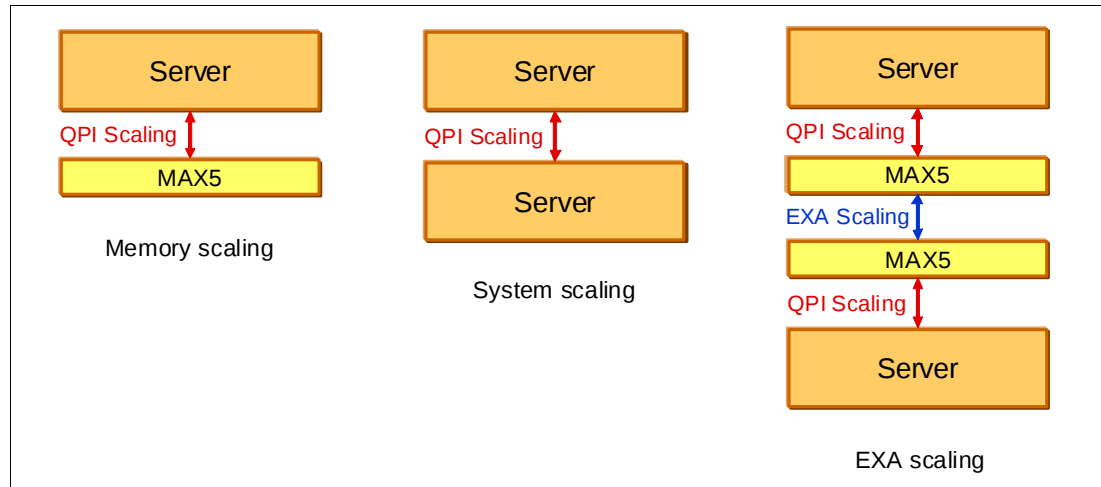


Figure 2-12 Types of scaling with eX5 systems

System scaling is possible for up to two nodes on HX5 and x3850 X5, and EXA scaling is possible on the x3850 X5. The scaling choices that you have available to you depend on the server and processors installed.

The scalability capabilities for eX5 systems with Intel Xeon processor E7 family installed are summarized in Table 2-3.

Table 2-3 Comparing the scalability features of the Intel Xeon E7 processors

	E7-2800	E7-4800	E7-8800
x3690 X5	Yes	Yes	Yes
x3690 X5 with MAX5	Yes ^a	Yes	Yes
HX5	Yes	Yes	Yes
HX5 with MAX5	Yes ^a	Yes	Yes
HX5 two-node	Not supported	Yes	Yes
x3850 X5	Not supported	Yes	Yes
x3850 X5 with MAX5	Not supported	Yes	Yes
x3850 X5 two-node without MAX5	Not supported	Not supported	Yes
x3850 X5 two-node with MAX5 (EXA scaling)	Not supported	Yes	Yes

a. E7-2803 and E7-2820 processors do not support MAX5

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

- ▶ IBM System x3850 X5: 3.6, “Scalability” on page 63
- ▶ IBM System x3690 X5: 4.6, “Scalability” on page 129
- ▶ BladeCenter HX5: 5.9, “Scalability” on page 196

2.6 Partitioning

You can operate the eX5 scaled system as two independent systems or as a single system, without physically accessing the systems. 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. And, you partition through the integrated management modules (IMMs) on the E7 processor models of x3850 X5 and x3950 X5.

Figure 2-13 depicts an HX5 system that is scaled to two nodes (left) and an HX5 system that is partitioned into two independent servers (right).

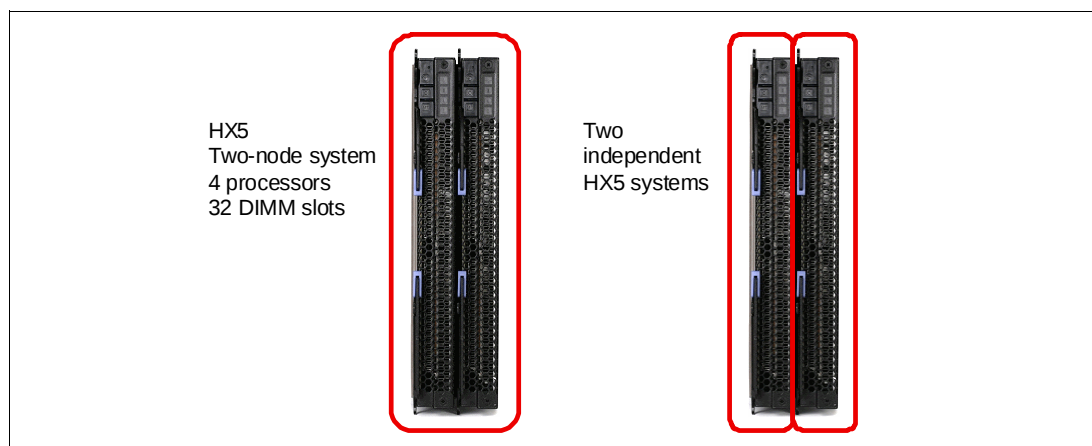


Figure 2-13 HX5 scaling and partitioning

Table 2-4 lists which configurations support partitioning.

Table 2-4 Support for partitioning

Configuration	Support for partitioning
x3690	Not supported
x3690 X5 with MAX5	Not supported
HX5 single-node	Not supported
HX5 with MAX5	Not supported
HX5 two-node	Yes
x3850 X5	Not supported
x3850 X5 with MAX5	Not supported
x3850 X5 two-node without MAX5	Not supported
x3850 X5 two-node with MAX5	Yes

Figure 2-14 shows a scalable complex configuration option for stand-alone mode through the AMM of the BladeCenter chassis.

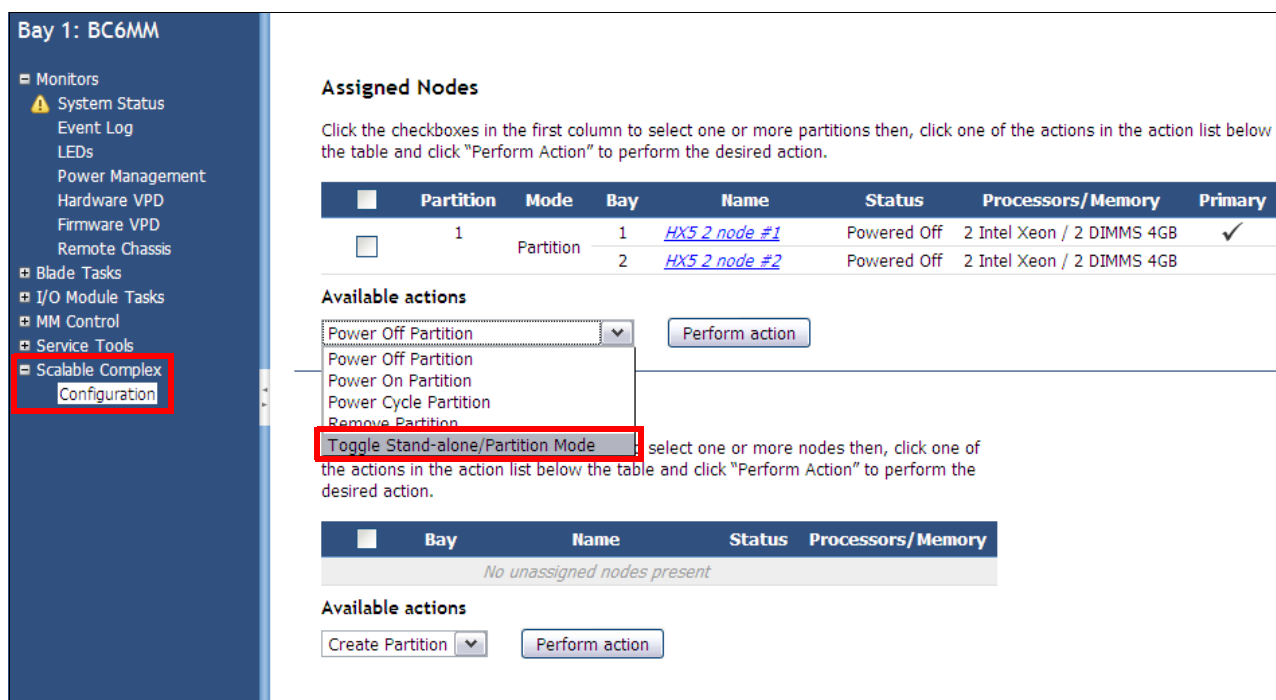


Figure 2-14 Option for putting a partition into stand-alone mode

Figure 2-15 shows an HX5 partition in stand-alone mode.

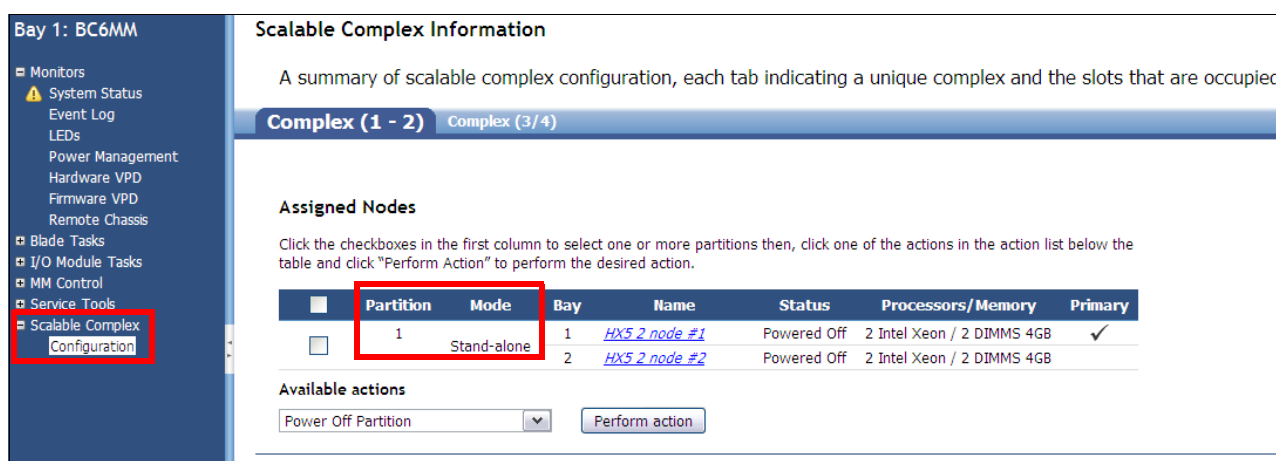


Figure 2-15 HX5 partition in stand-alone mode

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

Support for FlexNode partitioning is included with all scalable systems. Before a two-node solution can be used, you must create a partition. When the servers are scaled, they still act as single nodes until a partition is made.

2.7 Unified Extensible Firmware Interface system settings

The Unified Extensible Firmware Interface (UEFI) is a pre-boot environment that provides an interface between server firmware and the operating system. UEFI 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 and power consumption.

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

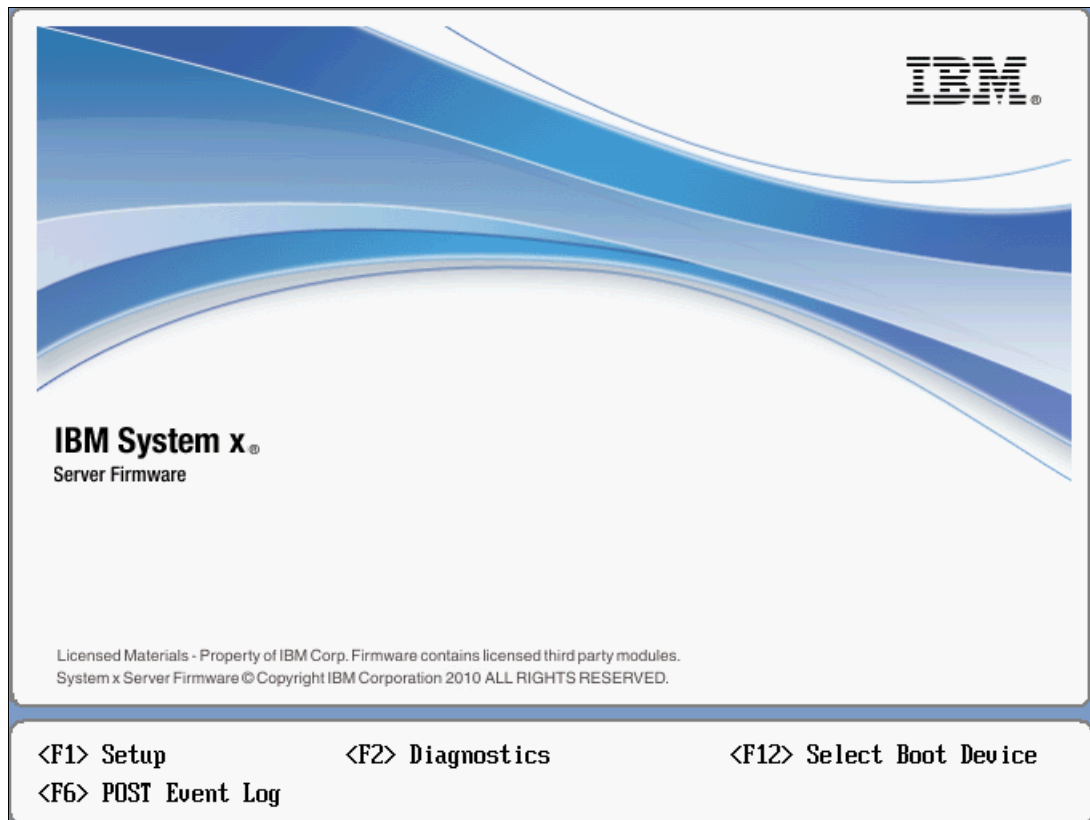


Figure 2-16 UEFI panel on system start

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

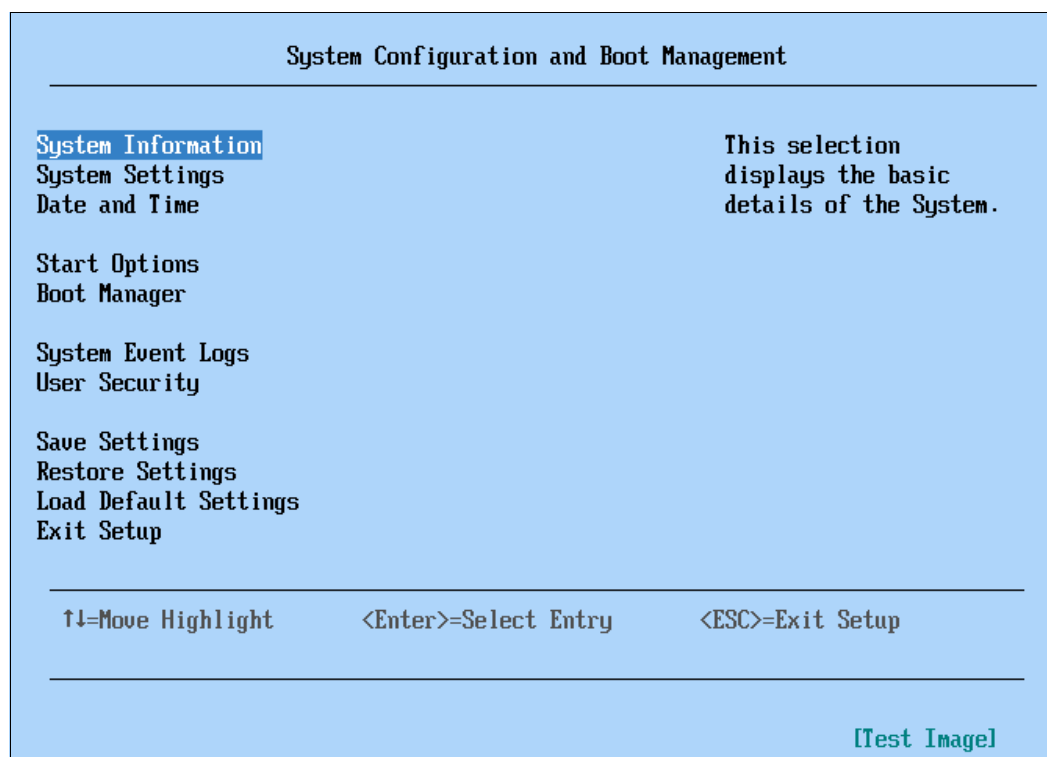


Figure 2-17 UEFI System Configuration and Boot Management page

To access the system settings options that are described here, choose **System Settings**. The page that is pictured in Figure 2-18 is displayed.

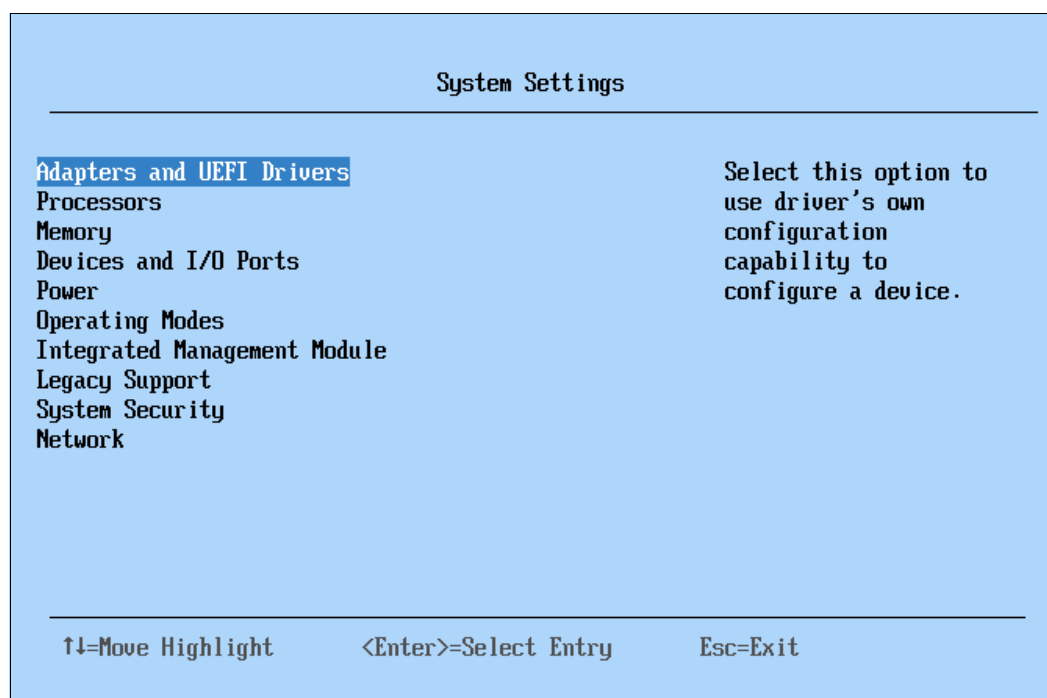


Figure 2-18 UEFI System Settings page

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 wanted operation conditions. The conditions for which these setups can be configured are now described.

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 then see the four operating modes from which to choose, as shown in Figure 2-19. When you choose a mode, the affected settings change to predetermined values, as shown.

Operating Modes		
Choose Operating Mode	<Custom Mode>	Efficiency: Highest performance per watt.
Memory Speed	<Max Performance>	
CKE Low Power	<Disable>	Acoustic: Lower acoustic level and lowest power.
Proc Performance States	<	
C1 Enhance Mode	< Acoustic Mode	
CPU C-States	< Efficiency Mode	
QPI Link Frequency	< Performance Mode	Performance: Highest performance.
Turbo Mode	< Custom Mode	
Turbo Boost Power Optimization	<	Custom: Customized settings.
Quiet Boot	<Enable>	

Figure 2-19 Operating modes in UEFI

These different modes are described.

Acoustic Mode

Figure 2-20 shows the Acoustic Mode predetermined values. They emphasize power-saving server operation to generate 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.

Operating Modes		
Choose Operating Mode	<Acoustic Mode>	Efficiency: Highest performance per watt.
Memory Speed	Minimal Power	Acoustic: Lower acoustic level and lowest power.
CKE Low Power	Enabled	
Proc Performance States	Enabled	
C1 Enhanced Mode	Enabled	
CPU C-States	Enabled	
QPI Link Frequency	Minimal Power	Performance: Highest performance.
Turbo Mode	Disabled	
Quiet Boot	<Enable>	Custom: Customized settings.

Figure 2-20 Acoustic Mode predetermined values

Efficiency Mode

Figure 2-21 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.

Operating Modes		
Choose Operating Mode	<Efficiency Mode>	Efficiency: Highest performance per watt.
Memory Speed	Power Efficiency	Acoustic: Lower acoustic level and lowest power.
CKE Low Power	Enabled	
Proc Performance States	Enabled	
C1 Enhanced Mode	Enabled	
CPU C-States	Enabled	
QPI Link Frequency	Power Efficiency	Performance: Highest performance.
Turbo Mode	Enabled	
Turbo Boost Power Optimization	Power Optimized	Custom: Customized settings.
Quiet Boot	<Enable>	

Figure 2-21 Efficiency Mode predetermined values

Performance Mode

Figure 2-22 shows the Performance Mode predetermined values. The server is set to 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.

Operating Modes		
Choose Operating Mode	<Performance Mode>	Efficiency: Highest performance per watt.
Memory Speed	Max Performance	Acoustic: Lower acoustic level and lowest power.
CKE Low Power	Disabled	
Proc Performance States	Enabled	
C1 Enhanced Mode	Disabled	
CPU C-States	Disabled	Performance: Highest performance.
QPI Link Frequency	Max Performance	
Turbo Mode	Enabled	
Turbo Boost Power Optimization	Traditional	
Quiet Boot	<Enable>	Custom: Customized settings.

Figure 2-22 Performance Mode predetermined values

Custom Mode

The default value that is set in new eX5 systems is Custom Mode, as shown in Figure 2-23. This 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.

Operating Modes		
Choose Operating Mode	<Custom Mode>	Efficiency: Highest performance per watt.
Memory Speed	<Max Performance>	Acoustic: Lower acoustic level and lowest power.
CKE Low Power	<Disable>	
Proc Performance States	<Enable>	
C1 Enhance Mode	<Enable>	
CPU C-States	<Disable>	Performance: Highest performance.
QPI Link Frequency	<Max Performance>	
Turbo Mode	<Enable>	
Turbo Boost Power Optimization	<Power Optimized>	
Quiet Boot	<Enable>	Custom: Customized 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 by 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
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. This approach enables a cross-platform management solution and simplifies 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 web page:
<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 shows the available options in the UEFI system power settings.

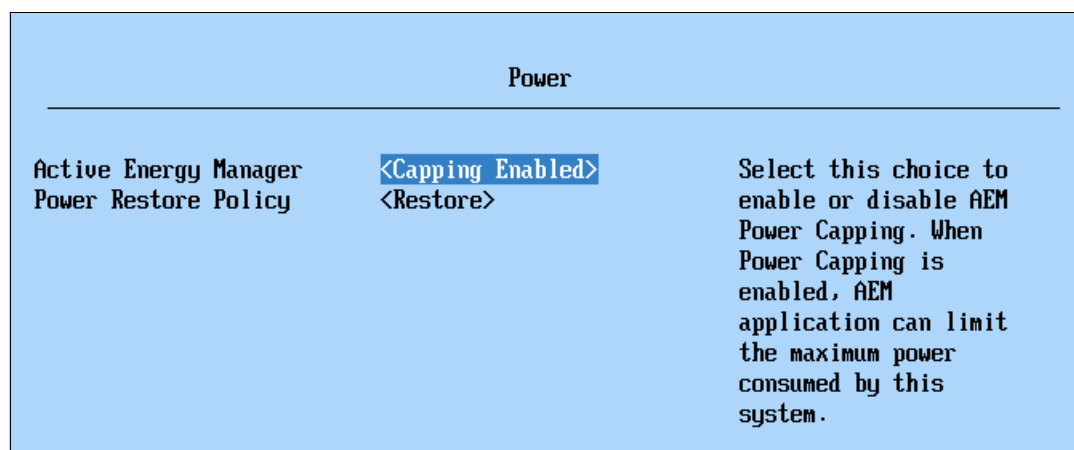


Figure 2-24 UEFI Power settings page

2.8 IBM eXFlash

IBM eXFlash is the name that is given to the 1.8-inch solid-state drives (SSDs), the backplanes, SSD hot-swap carriers, and indicator lights that are available for System x servers.

Each eXFlash 1.8-inch drive unit can replace four 2.5-inch serial-attached SCSI (SAS) hard disks or 2.5-inch form factor SSDs. You can install 1.8-inch eXFlash units according to the following specifications:

- ▶ The x3850 X5 can have either of the following configurations:
 - Up to four SAS or SATA drives, plus the eight 1.8-inch SSDs in one eXFlash unit
 - Sixteen 1.8-inch SSDs in two eXFlash units
- ▶ The x3950 X5 database-optimized models have two eXFlash units standard with space for 16 SSDs.
- ▶ The x3690 X5 can have up to 24 1.8-inch SSDs in three eXFlash units.

Spinning disks, although an excellent choice for cost per capacity, are not always the best choice when considering the cost of input/output operations per second (IOPS) and other factors.

In a production environment where the capacity requirements can be met by IBM eXFlash, the total cost per IOPS can be lower than any solution that requires 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, requiring, for example, a new rack or racks, extra power lines, or more 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 extra storage infrastructure that is 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 provides two grades of SSDs: Enterprise SSDs and Enterprise Value SSDs. The SSDs have similar read and write IOPS performance. However, the key difference between them is their endurance, which is how long they can do write operations, because SSDs have a finite number of program and erase cycles. Enterprise Value SSDs have a better cost/IOPS ratio but lower endurance when compared to Enterprise SSDs.

IBM Enterprise SSDs are 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, Enterprise SSDs 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 IBM ServerProven® for each machine are supported for use with SSDs.

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 wanted.

In a system that has two eXFlash backplanes that are installed, two controllers are required to connect to 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%. Each RAID controller controls only its own disks. With four M5015 controllers, each controller controls four disks. The effect of RAID-5 is that one disk per array is used for parity.

2.8.1 SSD and RAID controllers

You can use both RAID and non-RAID controllers. The IBM 6 Gb Performance Optimized HBA is optimized for read-intensive environments, and you can achieve maximum performance with only a single 6 Gb SSD HBA. A better choice for environments with a mixture of read and write activity is the ServeRAID M5014 or M5015 with the ServeRAID M5000 Performance Accelerator Key or the ServeRAID M5016.

In addition to using less power than rotating magnetic media, 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.

IOPS: I/O operations per second (IOPS) is used predominantly as a measure for database performance. Workloads that are 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.

More 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 large numbers of disks. When sizing for performance, it is common to greatly exceed the required capacity to reach the wanted number of IOPS.

Figure 2-25 on page 40 shows an eXFlash unit, with the status lights assembly on the left side.

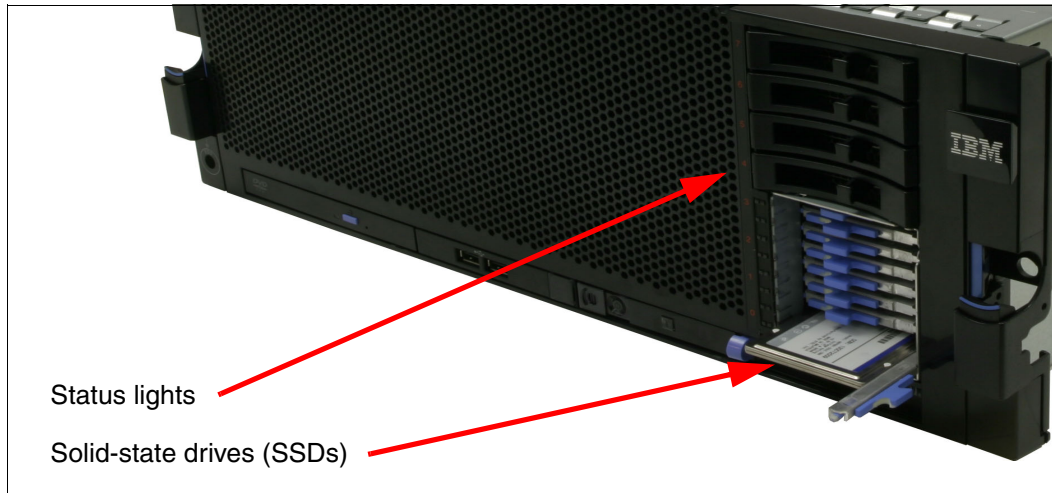


Figure 2-25 x3850 X5 with one eXFlash

Hot-swap capabilities: With the introduction of the 200 GB SSDs, the drives now support hot-swap capabilities. Therefore, the eXFlash trays have orange handles and not blue handles as shown in the figure.

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 88
- ▶ IBM System x3690 X5: 4.9.2, “IBM eXFlash and SSD 1.8-inch disk support” on page 151.

2.8.2 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. There is no 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.

IOPS: IOPS is a unit that is used predominantly as a measure for database performance. Workloads that are 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.

More 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 large numbers of disks. When you size for performance, it is common to greatly exceed the required capacity to reach the wanted IOPS.

If a typical disk drive can do 300 IOPS, and if the disk drive costs \$300, then the cost is \$1.00 per IOP. If a typical SSD can do 30,000 IOPS, and if it costs \$1000, then the cost per IOP is \$0.30. Configuring more disk drives to achieve the wanted number of IOPS can increase total system costs by requiring more disk controllers, disk enclosures, rack space, and power and cooling.

For more information about the devices that are mentioned here, see the relevant IBM Redbooks Product Guides:

- ▶ IBM SATA 1.8-inch and 2.5-inch MLC Enterprise SSDs for IBM System x
<http://www.redbooks.ibm.com/abstracts/tips0908.html>
- ▶ IBM SATA 1.8-inch and 2.5-inch MLC Enterprise Value SSDs
<http://www.redbooks.ibm.com/abstracts/tips0879.html>
- ▶ IBM 6 Gb Performance Optimized HBA
<http://www.redbooks.ibm.com/abstracts/tips0744.html>
- ▶ ServeRAID B5015 SSD Controller
<http://www.redbooks.ibm.com/abstracts/tips0763.html>
- ▶ ServeRAID M5015 and M5014 SAS / SATA Controllers
<http://www.redbooks.ibm.com/abstracts/tips0738.html>
- ▶ ServeRAID M5000 Series Performance Accelerator Key for IBM System x
<http://www.redbooks.ibm.com/abstracts/tips0799.html>
- ▶ ServeRAID M5016 SAS/SATA Controller
<http://www.redbooks.ibm.com/abstracts/tips0847.html>

For more information about storage for each of the eX5 systems, see the following sections:

- ▶ IBM System x3850 X5: 3.9, “Storage” on page 85
- ▶ IBM System x3690 X5: 4.9, “Storage” on page 146
- ▶ IBM BladeCenter HX5: 5.12, “Storage” on page 212

2.9 Integrated virtualization

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

2.9.1 VMware ESXi and vSphere

VMware ESXi is an embedded version of VMware ESX. The footprint of vSphere is relatively small because it does not provide the Service Console. Instead, it uses management tools, such as Virtual Center (vCenter), Remote Command-Line Interface (CLI), and Common Information Model (CIM) hardware monitoring. VMware ESXi includes full VMware File System (VMFS) support across Fibre Channel and iSCSI storage area networks (SANs), and network-attached storage (NAS). It supports eight-way virtual symmetrical multiprocessor systems (vSMPs).

Embedded virtualization keys are offered for the x3850 X5, x3690 X5, and HX5, as listed in Table 2-6.

Table 2-6 VMware vSphere 4.x memory key

Part number	Feature code	Description
41Y8296	A1NP	IBM USB Memory Key for VMware vSphere 4.1 Update 1
41Y8300	A2VC	IBM USB Memory Key for VMware vSphere 5.0
41Y8307	A383	IBM USB Memory Key for VMware vSphere 5.0 Update 1

Part number	Feature code	Description
41Y8311	A2R3	IBM USB Memory Key for VMware ESXi 5.1
41Y8298	A2G0	IBM Blank USB Memory Key for VMware vSphere downloads

For more information about USB keys, and to download the IBM customized version of VMware ESXi and VMware vSphere, visit the following web page:

<http://www.ibm.com/systems/x/os/vmware/esxi>

2.9.2 Red Hat RHEV-H (KVM)

The kernel-based virtual machine (KVM) that is supported by Red Hat Enterprise Linux (RHEL) 5.4 and later is available on the x3850 X5. The Red Hat Enterprise Virtualization Hypervisor (RHEV-H), or KVM, is standard with the purchase of RHEL 5.4 and later. All hardware components that were 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).

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 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, Windows 2012 with Hyper-V

Windows Server 2008 R2 and Windows Server 2012 with Hyper-V are also supported on the eX5 servers.

The following features are included:

- ▶ Cluster Shared Volumes
- ▶ Live migration
- ▶ Support for up to 64 logical cores
- ▶ Virtual machines snapshots



IBM System x3850 X5 and x3950 X5

The four-socket IBM System x3850 X5 and the IBM System x3950 X5 are introduced. The x3950 X5 models are optimized for specific workloads, such as virtualization and database workloads.

The MAX5 memory expansion unit is a 1U device that you connect to the x3850 X5 or x3950 X5. The MAX5 provides the server with an extra 32 DIMM sockets, ideal for applications that can take advantage of large amounts of memory.

The following topics are covered:

- ▶ 3.1, “Product features” on page 44
- ▶ 3.2, “Target workloads” on page 52
- ▶ 3.3, “Models” on page 53
- ▶ 3.4, “System architecture” on page 57
- ▶ 3.5, “MAX5” on page 60
- ▶ 3.6, “Scalability” on page 63
- ▶ 3.7, “Processor options” on page 68
- ▶ 3.8, “Memory” on page 70
- ▶ 3.9, “Storage” on page 85
- ▶ 3.10, “Optical drives” on page 99
- ▶ 3.11, “PCIe slots” on page 99
- ▶ 3.12, “I/O cards” on page 101
- ▶ 3.13, “Standard onboard features” on page 107
- ▶ 3.14, “Power supplies and fans of the x3850 X5 and MAX5” on page 110
- ▶ 3.15, “Integrated virtualization” on page 112
- ▶ 3.16, “Operating system support” on page 112
- ▶ 3.17, “Rack considerations” on page 113

3.1 Product features

The IBM System x3850X5 and x3950 X5 servers address the following requirements that many IBM enterprise clients need:

- ▶ Increased performance on a smaller IT budget
- ▶ The ability to increase database and virtualization performance without having to add more CPUs, especially valuable 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, although software licensing costs remain constant
- ▶ The flexibility to achieve the wanted memory capacity with larger capacity DIMMs
- ▶ The ability to pay for the system that clients need today, with the capability to grow both memory capacity and processing power when necessary in the future

The basic 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 model that is designed for specific workloads. The announced x3950 X5 models are optimized for database or virtualization applications.

x3850 X5 term is used for common features: 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 7143, is the second generation of the x3850 X5. It is a 4U four-socket Intel Xeon E7-based (Westmere EX) platform with 64 DIMM sockets. It can be scaled up to eight processor sockets, depending on the model, and 192 DIMM sockets. This configuration can be done by connecting a MAX5 memory expansion drawer and a second x3850 X5 with another MAX5 memory expansion drawer.

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

See Table 3-2 on page 51 for a comparison of eX4 x3850 M2 and eX5 x3850 X5.

The x3850 X5 offers the following key features:

- ▶ Up to four Intel Xeon E7 series processors (6, 8, and 10 core)
- ▶ Scalability to eight sockets by connecting two x3850 X5 servers
- ▶ 64 DDR3 DIMM sockets
- ▶ Opportunity to install up to eight memory cards, each with eight DIMM slots
- ▶ Seven Peripheral Component Interconnect Express (PCIe) 2.0 slots (one slot contains the Emulex 10 Gb Ethernet dual-port adapter)
- ▶ Up to eight 2.5-inch hard disk drives (HDDs) or sixteen 1.8-inch solid-state drives (SSDs)
- ▶ Standard Redundant Array of Independent Disks-0 (RAID-0) and RAID-1, optional RAID-5 and 50, RAID-6 and 60, and encryption

- ▶ Two 1 Gb Ethernet ports
- ▶ One Emulex 10 Gb Ethernet dual-port adapter (standard on all models except 7145-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 inch)
- ▶ Depth: 712 mm (28.0 inch)
- ▶ Height: 173 mm (6.8 inch) or four 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 are 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 that is 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 you remove that component.

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

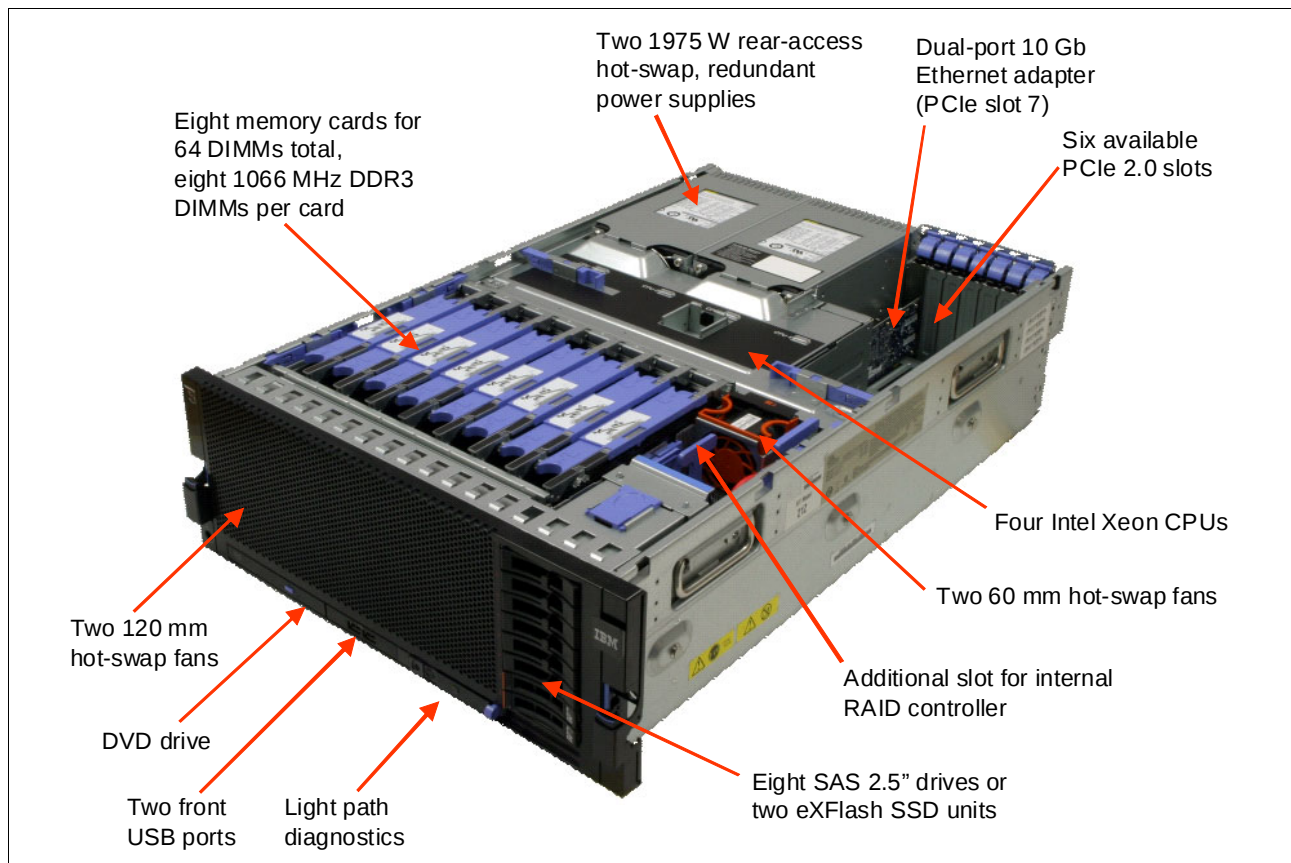


Figure 3-2 x3850 X5 internals

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

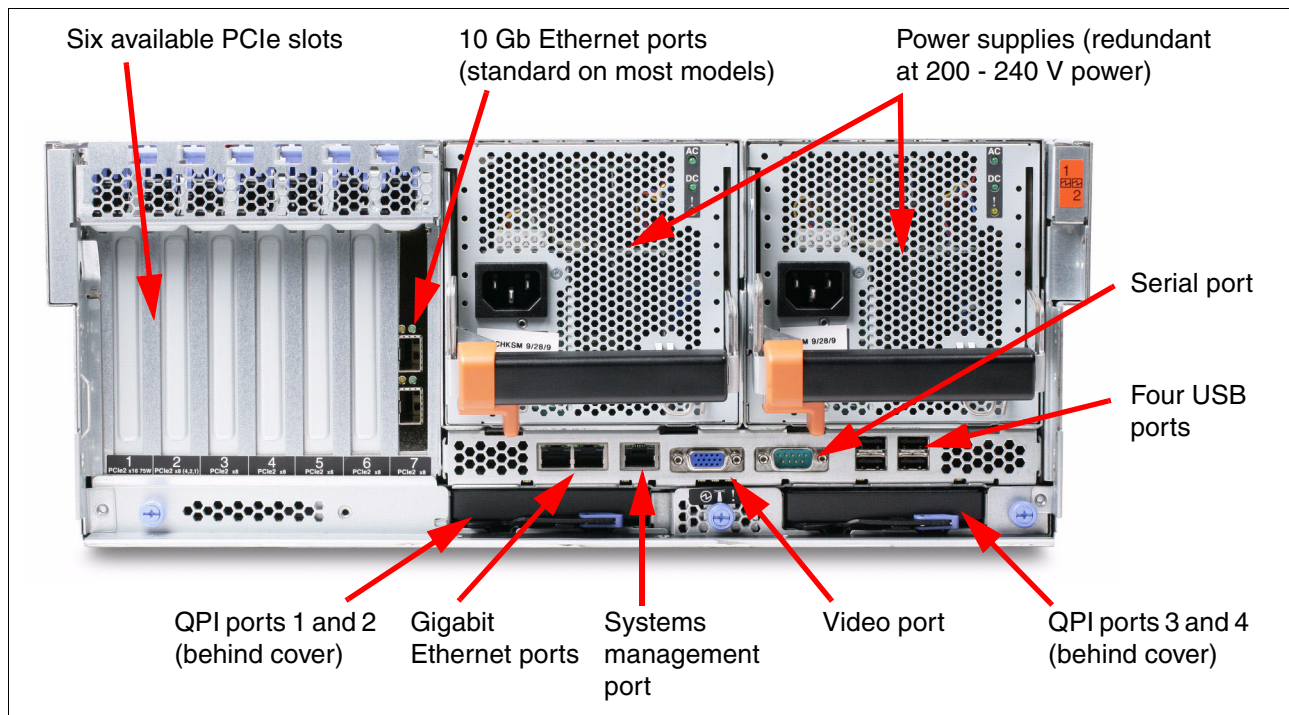


Figure 3-3 Rear of the x3850 X5

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 that are used to configure them. However, 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 eight-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 eight-socket solution. This configuration assumes 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 22.

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 53 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, two 675 watt power supplies, and five 40 mm hot-swap speed-controlled fans. The MAX5 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.

There is a second generation of the MAX5 called *MAX5 V2*. MAX5 V2 features newer versions of scalable memory buffers, which enable support for 1.35 V DIMMs, and 32 GB DIMMs.

Compatibility is summarized in Table 3-1. Certain combinations require minimum firmware levels as noted.

Table 3-1 MAX5 compatibility

MAX5 model	x3850 X5 with E7 processors (machine type 7143)
IBM MAX5, 59Y6265	Supported by minimum levels ^a
IBM MAX5 V2, 88Y6529	Supported

a. This combination requires these minimum firmware levels:
UEFI: G0E171T/A (signed)
IMM: YUOOC7E
pDSA: DSYT89O
FPGA: G0UD72B
ASU: 72L

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 buffer with eight memory slots (four DIMMs on each channel)
- ▶ 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)
- ▶ IBM EXA technology for configurations of two nodes with MAX5 units (EXA scaling); three connections operate at up to 10 GT/s
- ▶ Scalability:
 - Connects to an x3850 X5 server by using QPI cables
 - Connects to other MAX5 units, by using EXA link cables
 - Scales up to two nodes (two MAX5 units + two servers)
- ▶ Memory DIMMs:
 - Minimum: two DIMMs, 4 GB
 - Maximum: 32 DIMMs
 - MAX5: up to 512 GB of memory using 16 GB DIMMs
 - MAX5 V2: up to 1 TB of memory using 32 GB DIMMs
 - Type of DIMMs: PC3-10600, 1067 MHz, error correction code (ECC), DDR3 registered synchronous dynamic random access memory (SDRAM) DIMMs
 - DIMM sizes:
 - MAX5: Supports 2 GB, 4 GB, 8 GB, and 16 GB DIMMs
 - MAX5 V2: Supports 2 GB, 4 GB, 8 GB, 16 GB, and 32 GB DIMMs
 - Low voltage (1.35V) support for DIMMs with MAX5 V2

All DIMM sockets in the MAX5 are accessible regardless of the number of processors that are 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 (100 - 240 V ac auto-sensing)
 - Two power supplies standard and maximum (second power supply adds 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 inch)
 - Depth: 724 mm (28.5 inch)
 - Height: 44 mm (1.73 inch) (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 extra 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. With the second-generation x3850 X5 and MAX5 V2, three TB of RAM can be installed, using 32 GB DIMMs.

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

Figure 3-5 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 for use in configurations of a two-node x3850 X5 with two MAX5 units (called *EXA scaling*).

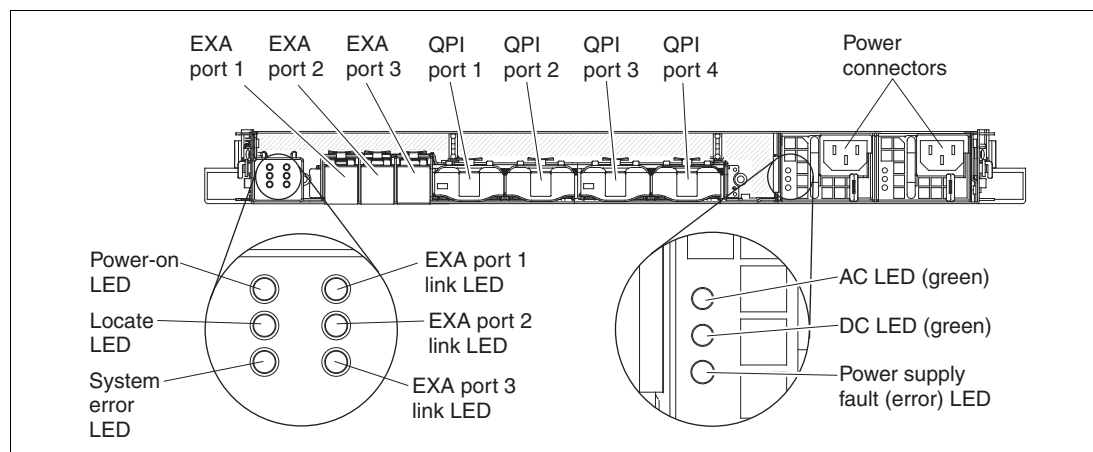


Figure 3-5 MAX5 connectors and LEDs

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.

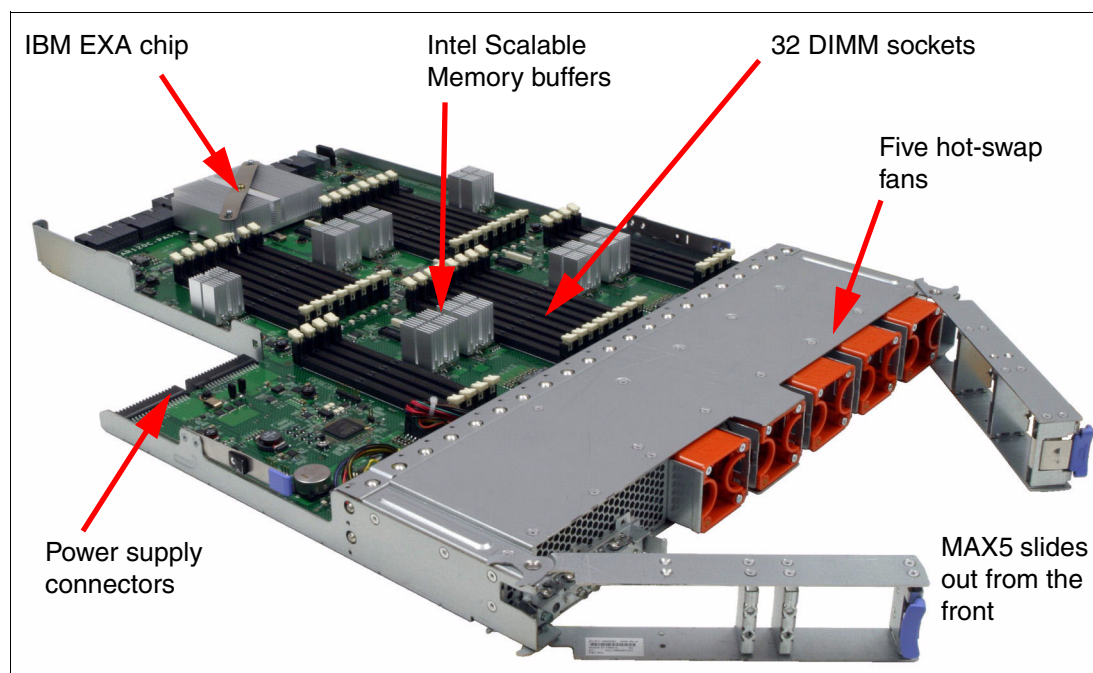


Figure 3-6 MAX5 memory expansion unit internal components

For an in-depth look at the MAX5 offering, see 3.5, “MAX5” on page 60.

3.1.4 Comparing the x3850 X5 to the x3850 M2

Table 3-2 shows a high-level comparison between the eX4-based x3850 M2 and the eX5-based x3850 X5.

Table 3-2 Comparison of the x3850 M5 to the x3850 X2

Subsystem	x3850 X5	x3850 M2
CPU card	▶ No voltage regulator modules (VRMs), Four voltage regulator downs (VRDs) ▶ Top access to CPUs and CPU card	▶ No voltage regulator downs (VRDs), Four voltage regulator modules (VRMs) ▶ Top access to CPU or 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 7510 "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 Callioc 2 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 ServRAID M1015 with RAID-0 and RAID-1 (most models) ▶ Optional ServeRAID M5015 with RAID-0, RAID-1, and RAID- 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 10 Gb 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 integrated management module (IMM) ▶ 16 MB VRAM	▶ ATI RN50 on Remote Supervisor Adapter (RSA2) ▶ 16 MB VRAM
Service processor	▶ Maxim VSC452 Integrated BMC (IMM) ▶ Remote presence feature standard	▶ RSA2 standard ▶ Optional remote presence feature
Disk drive support	▶ Eight 2.5-inch internal drive bays or 16 1.8-inch solid-state drive bays ▶ Support for SAS, SATA, and SSD	▶ Four 2.5-inch internal drive bays
USB, SuperIO design	▶ ICH10 chip set ▶ Six external USB ports, two internal ▶ No SuperIO ▶ No PS/2 keyboard or mouse connectors ▶ No diskette drive controller ▶ Optional optical drive	▶ ICH7 chip set ▶ Five external USB ports, one internal ▶ No SuperIO ▶ No PS/2 keyboard or 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

Subsystem	x3850 X5	x3850 M2
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 preinstalled with VMware ESXi. This feature allows clients to set up and run a virtualized environment simply and quickly.
- MAX5 expansion drawer: Most consolidated workloads benefit from increased memory capacity per processor.

As a general guideline, virtualization is a workload that is memory-intensive and I/O-intensive. Therefore, the x3850 X5 is an optimal platform for consolidated workloads.

VMware compatibility: If you use a MAX5 unit, you must use VMware ESXi 4.1 or later. VMware ESXi 4.0 does not support MAX5.

For more information, visit the following web page:

http://www.vmware.com/resources/compatibility/detail.php?device_cat=server&device_id=5317&release_id=144#notes

- Virtualization optimized models: Two virtualization workload-optimized models of the x3950 X5 are announced. See 3.3, “Models” on page 53 for more information.
- Processor support: The Intel E7 series processors support VT FlexMigration Assist and VMware Enhanced VMotion.

▶ Database

Database workloads require powerful CPUs, high memory throughput, and disk subsystems that are configured to deliver many input/output operations per second (IOPS). IBM predefined database models and SAP High-Performance Analytic Appliance (HANA) models use eight or ten core CPUs and use the power of eXFlash (high-IOPS SSDs) or high-performance PCIe-based SSD adapters. For more information about eXFlash, see 3.9.3, “IBM eXFlash and 1.8-inch SSD support” on page 88.

▶ Compute-intensive processors and core

The x3850 X5, with Intel E7 processors, scales up to 80 CPU cores per node, enabling large multi-processor workloads to be run within a single system. These servers can run workloads previously thought to be beyond the abilities of x86 processor-based systems.

► Memory intensive

A single node x3850 X5 supports up to 2 TB of RAM. A two node x3850 X5 supports up to 4 TB of RAM, and a two node x3850 X5 with MAX5 supports up to 6 TB RAM. These memory amounts enable processing of large workloads in memory, which can dramatically reduce the time that is required to run certain applications.

For the workload-specific model details, see 3.3.3, “Models” on page 53.

3.3 Models

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

For information about 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 web page:

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

3.3.1 x3850 X5 base models with Intel E7 processors

Table 3-3 lists the standard models of machine type 7143 (with Intel Xeon E7-4800 and E7-8800 series processors).

In this table, *std* is standard, *max* is maximum, and *C* is core (such as 4C four-core).

Table 3-3 Standard models of machine type 7143 (Intel Xeon E7-4800 and E7-8800 series processors)

Model	Intel Xeon processors (qty, model, cores, core speed, L3 cache, memory speed) (four max)	Scale to two nodes without MAX5 ^a	Scale to two nodes with MAX5 ^b	MAX5	Memory (standard memory cards) ^c	ServeRAID M1015 adapter	2.5 in. HDD disk bays (std / max)	Disk drives standard	10 Gb Ethernet standard ^d	Power supplies (std / max)
7143-B1x	2x E7-4807 6C 1.86 GHz 18 MB 800 MHz	No	Yes	Opt	2x 4 GB (1)	Opt	0 / 8	None	Opt	1 / 2
7143-B2x	2x E7-4820 8C 2.00 GHz 18 MB 978M Hz	No	Yes	Opt	4x 4 GB (2)	Yes	4 / 8	None	Yes	2 / 2
7143-B3x	2x E7-4830 8C 2.13 GHz 24 MB 1066 MHz	No	Yes	Opt	4x 4 GB (2)	Yes	4 / 8	None	Yes	2 / 2
7143-B5x	2x E7-4850 10C 2.00 GHz 24 MB 1066 MHz	No	Yes	Opt	4x 4 GB (2)	Yes	4 / 8	None	Yes	2 / 2
7143-B6x	2x E7-4860 10C 2.26 GHz 24 MB 1066 MHz	No	Yes	Opt	4x 4 GB (2)	Yes	4 / 8	None	Yes	2 / 2
7143-B7x	2x E7-4870 10C 2.40 GHz 30 MB 1066 MHz	No	Yes	Opt	4x 4 GB (2)	Yes	4 / 8	None	Yes	2 / 2
7143-C1x	2x E7-8850 8C 2.00 GHz 24 MB 1066 MHz	Yes	Yes	Opt	4x 4 GB (2)	Yes	4 / 8	None	Yes	2 / 2

Model	Intel Xeon processors (qty, model, cores, core speed, L3 cache, memory speed) (four max)	Scale to two nodes without MAX5 ^a	Scale to two nodes with MAX5 ^b	MAX5	Memory (standard memory cards) ^c	ServerRAID M1015 adapter	2.5 in. HDD disk bays (std / max)	Disk drives standard	10 Gb Ethernet standard ^d	Power supplies (std / max)
7143-C2x	2x E7-8860 10C 2.26GHz 24MB 1066	Yes	Yes	Opt	4x 4 GB (2)	Yes	4 / 8	None	Yes	2 / 2
7143-C3x	2x E7-8870 10C 2.40GHz 30MB 1066	Yes	Yes	Opt	4x 4 GB (2)	Yes	4 / 8	None	Yes	2 / 2

- a. The ability to form a two-node configuration without a MAX5 installed. Also known as *QPI scaling*. E7-4000 series processors do not support this type of scaling. See 3.6, “Scalability” on page 63.
- b. The ability to form a two-node configuration provided a MAX5 is also installed. Also known as *EXA scaling*. See 3.6, “Scalability” on page 63.
- c. The number in brackets is the number of memory cards standard in each model. Up to eight memory cards are supported. Each card holds up to eight DIMMs for a total of 64 DIMMs. The MAX5, when installed, adds 32 DIMM sockets for a total of 96 DIMMs.
- d. The Emulex 10 Gb Ethernet Adapter II is installed in PCIe slot 7. See 3.12.1, “Emulex 10 GbE Integrated Virtual Fabric Adapter II” on page 101.

3.3.2 Workload-optimized x3950 X5 models with Intel E7 processors

Table 3-4 lists the workload-optimized models with Intel E7 processors. (In the table, *std* is standard, and *max* is maximum).

Table 3-4 Workload-optimized models of machine type 7143 (Intel Xeon E7-4800 and E7-8800 series processors)

Model	Processor (qty, model, cores, core speed, L3 cache, memory speed) (four max)	Scale to two nodes without MAX5 ^a	Scale to two nodes with MAX5 ^b	MAX5	Std RAM / Mem cards (8 max) ^c	Standard RAID	Disk bays (std / max)	Std drives	GbE ^d	Multiburner drive
Database workload-optimized models										
7143-D3x	4x Xeon E7-4860 10C 2.26 GHz 24 MB 1066 MHz	No	Yes	Opt	32x 4 GB 8 cards	2x 6 Gb SSD HBA	16 / 16 1.8-inch SSD	16x 200 GB	2x 1 Gb 2x 10 Gb	Opt
7143-D4x	4x Xeon E7-4860 10C 2.26 GHz 24 MB 1066 MHz	No	Yes	Opt	32x 4 GB 8 cards	4x M5015 + perf keys	16 / 16 1.8-inch SSD	16x 200 GB	2x 1 Gb 2x 10 Gb	Opt
SAP HANA workload-optimized models										
7143-HAx	2x Xeon E7-8870 10C 2.40 GHz 30 MB 1066 MHz	Yes	No ^e	No ^e	16x 16 GB 4 cards	1x M5015 + battery	8 / 8 2.5-inch HDD	8x 900 GB 10K SAS 1x 1.2 TB PCIe	6x 1 Gb 4x 10 Gb	Yes
7143-HBx	4x Xeon E7-8870 10C 2.40 GHz 30 MB 1066 MHz	Yes	No ^e	No ^e	32x 16 GB 8 cards	1x M5015 + battery	8 / 8 2.5-inch HDD	8x 900 GB 10K SAS 1x 1.2 TB PCIe	6x 1 Gb + 4x 10 Gb	Yes

Model	Processor (qty, model, cores, core speed, L3 cache, memory speed) (four max)	Scale to two nodes without MAX5 ^a	Scale to two nodes with MAX5 ^b	MAX5	Std RAM / Mem cards (8 max) ^c	Standard RAID	Disk bays (std / max)	Std drives	GbE ^d	Multiburner drive
7143-HCx	4x Xeon E7-8870 10C 2.40 GHz 30 MB 1066 MHz	Yes ^f	No ^e	No ^e	32x 16 GB 8 cards	1x M5015 + battery	8 / 8 2.5-inch HDD	8x 900 GB 10K SAS 1x 1.2 TB PCIe	6x 1 Gb + 4x 10 Gb	Opt
Virtualization workload-optimized models										
7143-F1x (ESX)	4x Xeon E7-4860 10C 2.26 GHz 24 MB 1066 MHz	No	Yes	Std (V2)	Server: 64x 4 GB 8 cards MAX5: 32x 4 GB	1x M1015	4 / 8 2.5-inch HDD	Open	2x 1 Gb + 2x 10 Gb	Opt
7143-F2x (RH)	4x Xeon E7-4860 10C 2.26 GHz 24 MB 1066 MHz	No	Yes	Std (V2)	Server: 64x 4 GB 8 cards MAX5: 32x 4 GB	1x M1015	4 / 8 2.5-inch HDD	Open	2x 1 Gb + 2x 10 Gb	Opt
Virtualization workload-optimized models (China only)										
7143-B8x (ESX)	4x Xeon E7-4860 10C 2.26 GHz 24 MB 1066 MHz	No	Yes	Std (V1)	Server: 64x 4 GB 8 cards MAX5: 32x 4 GB	1x M1015	4 / 8 2.5-inch HDD	Open	2x 1 Gb + 2x 10 Gb	Opt
7143-B9x (ESX)	4x Xeon E7-4860 10C 2.26 GHz 24 MB 1066 MHz	No	Yes	Std (V2)	Server: 64x 4 GB 8 cards MAX5: 32x 4 GB	1x M1015	4 / 8 2.5-inch HDD	Open	2x 1 Gb + 2x 10 Gb	Opt

- a. The ability to form a two-node configuration without a MAX5 installed. Also known as *QPI scaling*. E7-4000 series processors do not support this type of scaling. See 3.6, “Scalability” on page 63.
- b. The ability to form a two-node configuration provided a MAX5 is also installed. Also known as *EXA scaling*. See 3.6, “Scalability” on page 63.
- c. Up to eight memory cards are supported in the server and each card holds up to eight DIMMs for a total of 64 DIMMs. The MAX5, when installed, adds 32 DIMM sockets for a total of 96 DIMMs. No memory cards are used in the MAX5.
- d. The H models include one Emulex 10GbE Integrated Virtual Fabric Adapter and one Intel Ethernet Quad Port Server Adapter I340-T4 for a total of two 10Gb ports and six 1Gb ports. F1x and F2x models include one Emulex 10GbE Integrated Virtual Fabric Adapter. D3x and D4x models include one Emulex 10GbE Integrated Virtual Fabric Adapter II.
- e. MAX5 is not currently certified for use with SAP HANA and is therefore not supported.
- f. Model HCx includes the QPI Scalability Kit (four cables), part number 46M0072. Use model HCx with a model HBx to form a two-node scaled complex.

The following list provides information about these models:

- Models 7143-D3x, D4x: These models are designed for database applications and use solid-state drives (SSDs) for the best I/O performance.

Backplane connections for 16x 1.8-inch SSDs are standard, as are 16x 200 GB high-performance SSDs. Model D3x includes two SSD host bus adapters and Model D4x includes four ServeRAID M5015 adapters with performance keys.

- Models 7143-HAx, HBx, HCx: The H models are optimized to run the SAP High-Performance Analytic Appliance (HANA) solution. HANA is an integrated, ready-to-run, hardware-software offering, featuring the new SAP In-Memory Computing Engine.

Models HAx and HBx include preinstalled software comprising SUSE Linux Enterprise Server (SLES) for SAP, IBM General Parallel File System (GPFS™), and the SAP HANA software stack. HCx is a model that is designed to be connected to a model HBx system to form an eight-processor system. HCx includes the four QPI cables necessary to join two systems together to form a two-node complex. HCx also includes the additional licenses to cover the extra four sockets but does not include any preinstalled software because it is designed as an add-on to the HBx offering.

All H models include either 256 GB or 512 GB of RAM, SAS disk drives, and a high-IOPS 1.2 TB solid-state storage PCIe adapter.

- Model 7143-F1x, B8x, B9x: These models are designed for virtualization applications and include VMware ESXi 4.1 Update 1 on an integrated bootable USB memory key.

The models come standard with the MAX5 memory expansion unit and 384 GB of memory that is implemented by using cost-effective 4 GB memory DIMMs (256 GB in the server and 128 GB in the MAX5).

F1x is available worldwide and includes a MAX5 V2, 88Y6529. Models B8x and B9x are for China only. B8x includes a MAX5 59Y6265. B9x includes a MAX5 V2 88Y6529.

- Model 7143-F2x: This model is designed for Open Virtualization and includes Red Hat Enterprise Linux with the Red Hat Enterprise Virtualization Hypervisor (kernel-based virtual machine (KVM)). The software is not preinstalled.

The model comes standard with the MAX5 memory expansion unit and 384 GB of memory that is implemented by using cost-effective 4 GB memory DIMMs (256 GB in the server and 128 GB in the MAX5).

3.4 System architecture

The system board architecture and the use of the QPI wrap card is explained.

3.4.1 System board

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

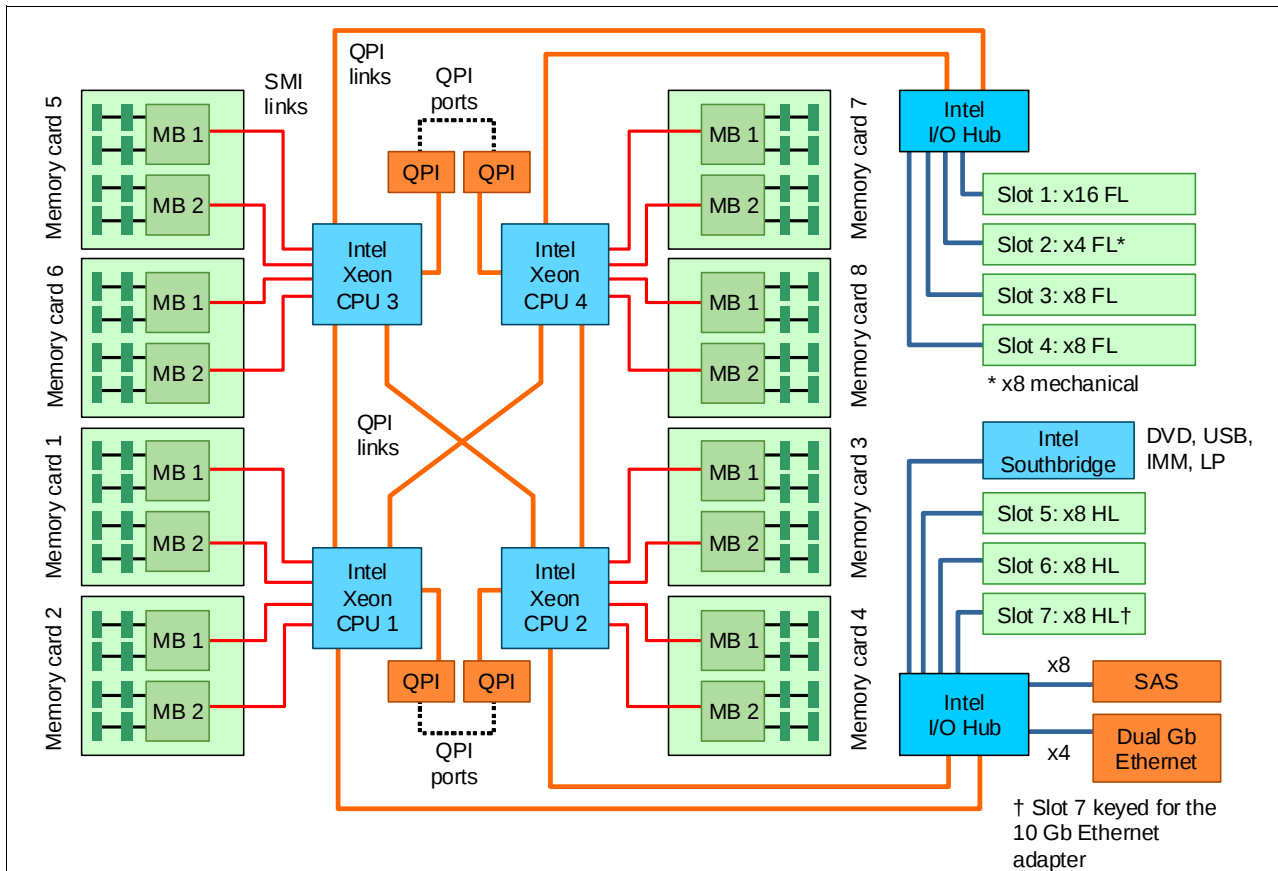


Figure 3-7 Block diagram for single-node x3850 X5

The following abbreviations are used in this figure:

- ▶ CPU: central processing unit
- ▶ FL: full length
- ▶ HL: half length
- ▶ IMM: integrated management module
- ▶ LP: light path diagnostics
- ▶ MB: memory bus
- ▶ QPI: QuickPath Interconnect

The dotted lines indicate where the QPI wrap cards are installed in a four-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 two-processor configurations and are removed when a MAX5 is connected.

Figure 3-12 on page 62 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 two-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.

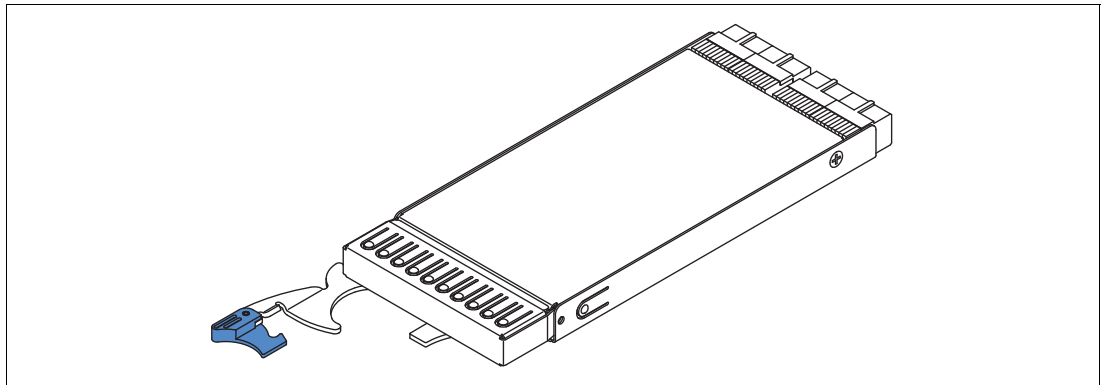


Figure 3-8 QPI Wrap Card

For single-node systems with 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. See Figure 3-9.

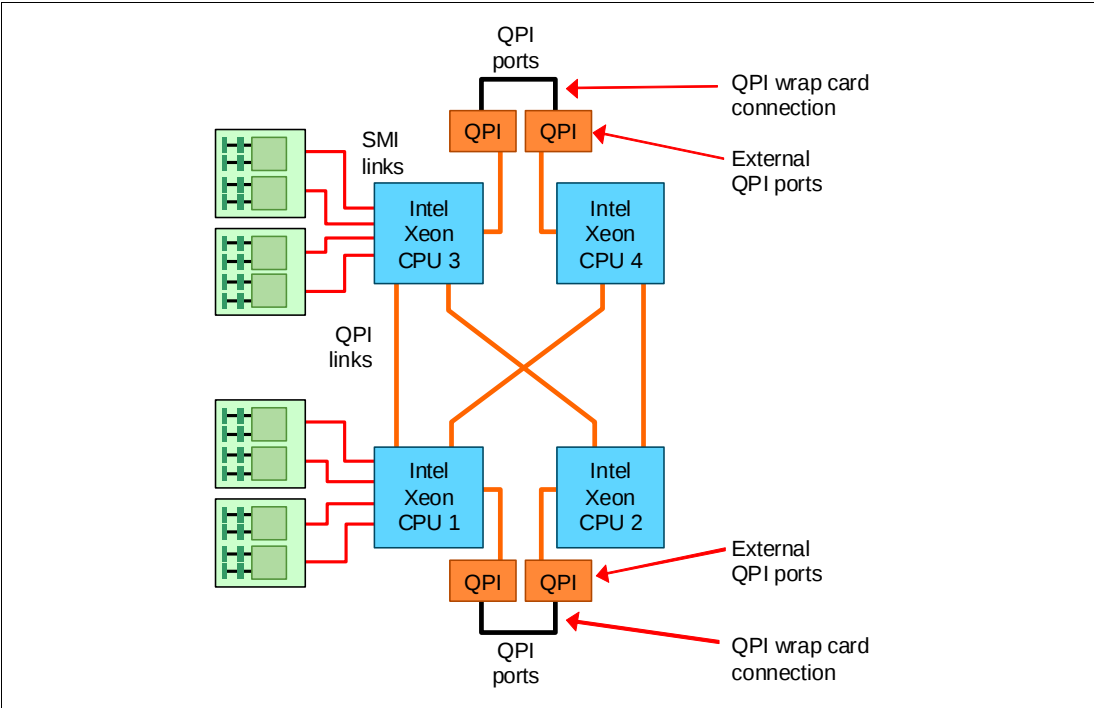


Figure 3-9 Location of QPI wrap cards

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 two-node configuration nor 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.

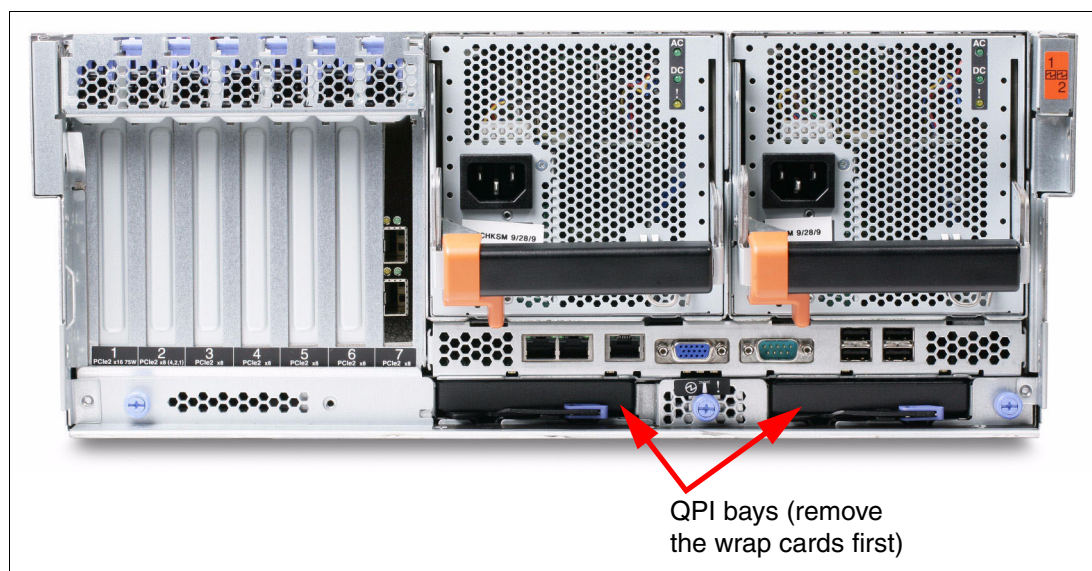


Figure 3-10 Rear of the x3850 X5 showing QPI connector wrap cards installed

3.5 MAX5

As introduced in 3.1.3, “IBM MAX5 memory expansion unit” on page 47, 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 53.

There are two MAX5 options available.

- ▶ IBM MAX5 for System x, part number 59Y6265 (also known as MAX5 V1)
- ▶ IBM MAX5 V2 for System x, part number 88Y6529

Both x3850 X5 machine types (7143 and 7145) support both MAX5 options, provided the firmware is at least UEFI level G0E171T/A. When used with the x3850 X5 machine type 7143 (Intel Xeon E7-4800 and E7-8800 series processors), MAX5 V2 supports low-voltage (operating at 1.35V DIMMs).

You can also order the MAX5 separately, as listed in Table 3-6. When you order a MAX5, remember to order the appropriate cable kits as well.

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

Part number	Feature code	Description
59Y6265	4199	IBM MAX5 for System x
88Y6529	A19H	IBM MAX5 V2 for System x

Part number	Feature code	Description
60Y0332	4782	IBM High Efficiency 675W Power Supply (MAX5 V1 only, 59Y6265)
59Y6267	4192	IBM MAX5 to x3850 X5 cable kit
59Y6271	4198	IBM eX5 MAX5 two-node EXA scalability kit

Compatibility is summarized in Table 3-7. Certain combinations require minimum firmware levels as noted.

Table 3-7 MAX5 compatibility

	x3850 X5 with E7 processors (machine type 7143)
IBM MAX5 V1, 59Y6265	Supported at minimum levels ^a
IBM MAX5 V2, 88Y6529	Supported

a. This combination requires the following minimum firmware levels:

UEFI: G0E171T/A (signed)
IMM: YUOOC7E
pDSA: DSYT89O
FPGA: G0UD72B
ASU: 72L

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

The eX5 chip connects directly through QPI links to all of the CPUs in the x3850 X5. And, it maintains a copy of the last-level cache of each CPU. Therefore, when a CPU requests content that is stored in the cache of another CPU, the MAX5 not only has that same data stored in its own cache. But, the MAX5 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 2.2.6, “QuickPath Interconnect” on page 13.

The MAX5 also has EXA scalability ports that are used in an EXA-scaled configuration (that is, a two-node and MAX5 configuration).

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
- ▶ Reduces snoop latencies

Figure 3-11 shows a diagram of the MAX5.

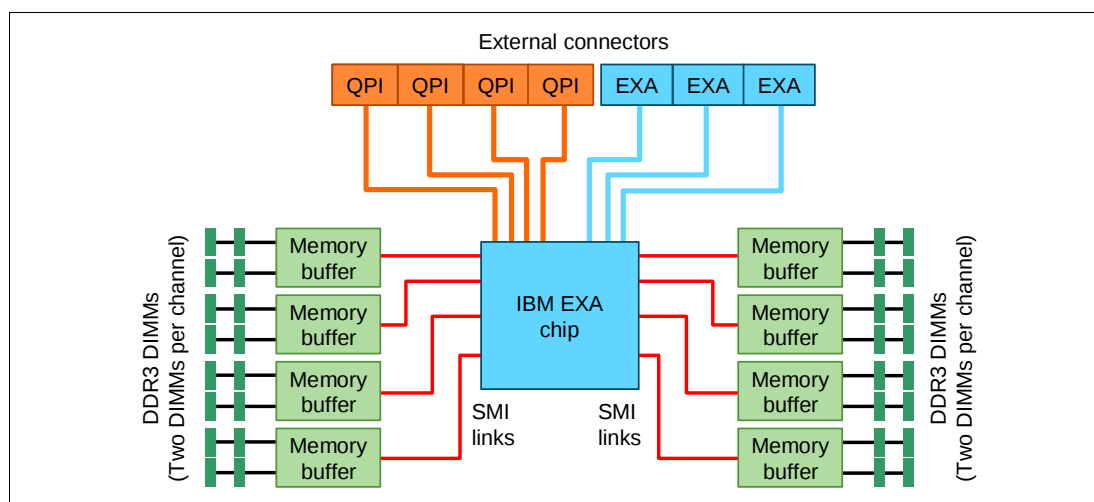


Figure 3-11 MAX5 block diagram

The MAX5 is connected to the x3850 X5 by using four cables that connect 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.

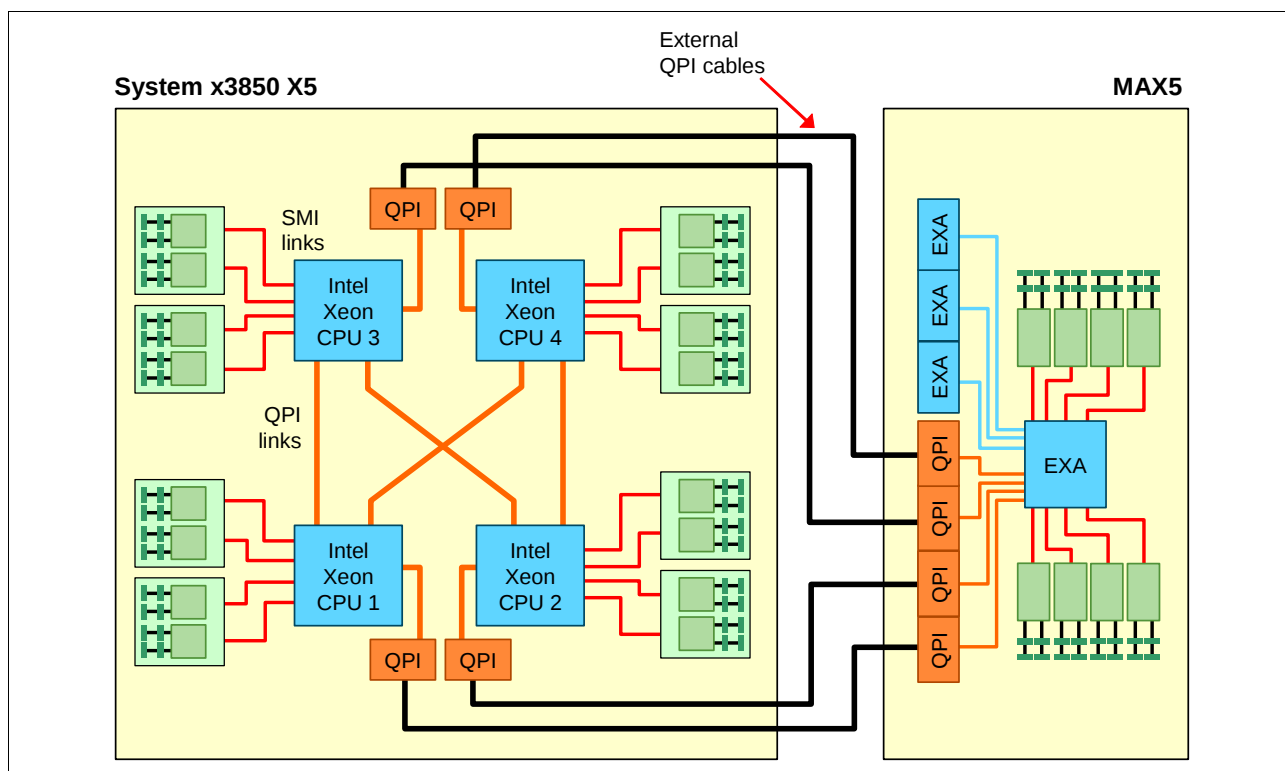


Figure 3-12 The x3850 X5: Connectivity of the system unit with the MAX5

We describe the connectivity of the MAX5 to the x3850 X5 in 3.6, “Scalability” on page 63.

Tip: As shown in Figure 3-12 on page 62, you maximize performance when you have four processors that are installed because you then have four active QPI links to the MAX5. However, configurations of two processors or three processors are still supported. If only two processors are required, consider the use of the x3690 X5.

For memory configuration information, see 3.8.6, “Memory mirroring” on page 82.

MAX5 V1 includes one power supply. The second power supply is optional (part 60Y0332) as listed in Table 3-6 on page 60 and provides redundancy. MAX5 V2 includes two power supplies so no additional power supplies are needed or available. MAX5 power supplies are hot-pluggable 675 W units. With two installed, the power subsystem is designed for N+N (fully redundant) operation and hot-swap replacement.

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 that is based on altitude and ambient temperature. In addition, a fan that is located inside each power supply cools the power modules.

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.

3.6 Scalability

The x3850 X5 can be expanded 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 that are connected to form a single image eight-socket server. This configuration is sometimes referred to as a *two-node server*.
- ▶ Two x3850 X5 servers that are connected together to form a single image eight-socket server with each server having a MAX5 attached. This configuration is sometimes referred to as a *two-node memory-expanded server*.

3.6.1 Memory scalability with MAX5

The MAX5 memory expansion unit allows the x3850 X5 to scale to an extra 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 3-8. Figure 3-13 also shows the connectivity.

Tip: As shown in Figure 3-12 on page 62, you maximize performance when you have four processors that are installed because you have four active QPI links to the MAX5. However, configurations of two and three processors are still supported.

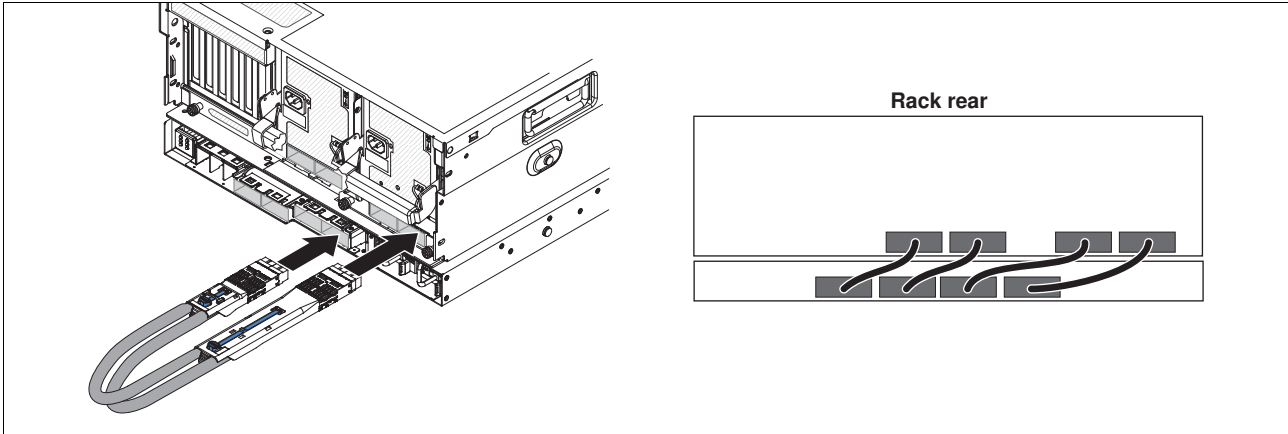


Figure 3-13 Connecting the MAX5 to a single-node x3850 X5

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 3-8.

Table 3-8 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 four cables)

To maximize performance, have all four processors and all memory cards installed and populated on the server. However, the x3850 X5 does support a single-node + MAX5 configuration with only two or three processors installed.

3.6.2 Two-node scalability

The two-node configuration also uses native Intel QPI scaling to create an eight-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 57. Figure 3-14 shows the cable routing.

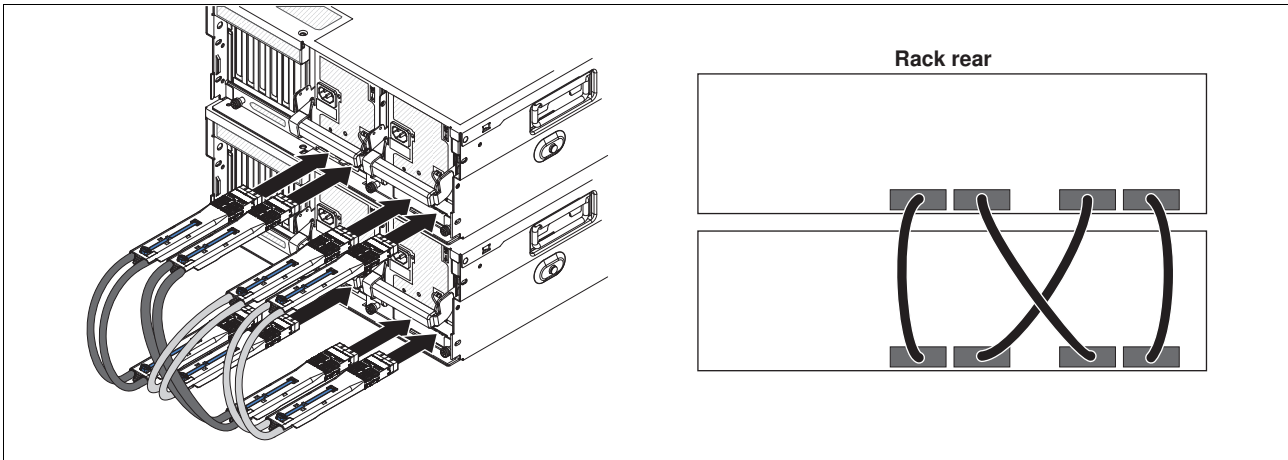


Figure 3-14 Cabling diagram for two-node x3850 X5

Connecting the two x3850 X5 servers to form a two-node system requires one IBM x3850 X5 and x3950 X5 QPI scalability kit, which consists of four QPI cables. See Table 3-9.

Table 3-9 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 four 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. These ports are located inside the server, close to the CPUs. Completing the QPI scaling of two x3850 X5 servers into a two-node complex does not require any other option.

In two-node configurations, both nodes must have all four processors and all memory cards installed and populated. All processors must be identical.

The following conditions are required to form a two-node configuration:

- ▶ Both servers must have processors in all four processor sockets.
- ▶ The processor specification must match among all processors in both servers.
- ▶ E7-8800 series processors are required.
- ▶ E7-4800 series processors cannot be used for two-node configurations.

For two-node configurations, DIMMs must be installed so that hemisphere mode can be enabled on the processors of both nodes. See Table 3-18 on page 76 to see the memory configurations that enable hemisphere mode.

Figure 3-15 shows the QPI links that are used to connect two x3850 X5 servers to each other.

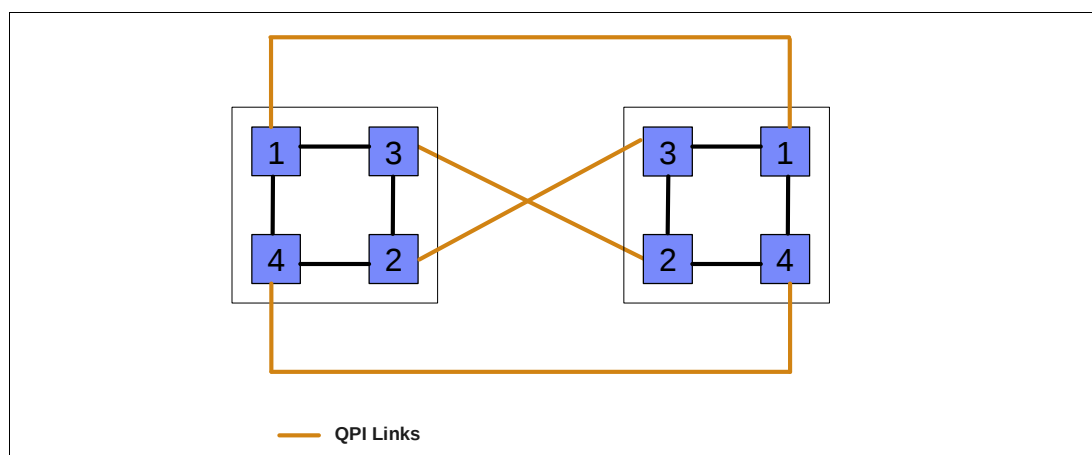


Figure 3-15 QPI links for a two-node x3850 X5

3.6.3 Two-node and MAX5 scalability

This configuration is only supported by the x3850 X5, machine type 7143. Machine type 7145 does not support this configuration.

With this configuration, each MAX5 is connected to an x3850 X5 by using QPI cables to form a 5U system. These two 5U systems are then cabled together with EXA cables to each MAX5 unit to form a single 10U system image:

- ▶ Eight processors for a total of up to 80 cores and 160 threads
- ▶ 192 DIMM sockets that support up 6 TB of RAM

In this configuration, install eight processors to maximize performance, but installing four processors (two in each node) is also supported. In addition, only Intel Xeon E7 processors are supported, that is, only x3850 X5 machine type 7143 supports this configuration. All processors must be identical.

Figure 3-16 shows the cable routing for a two-node x3850 X5 with MAX5.

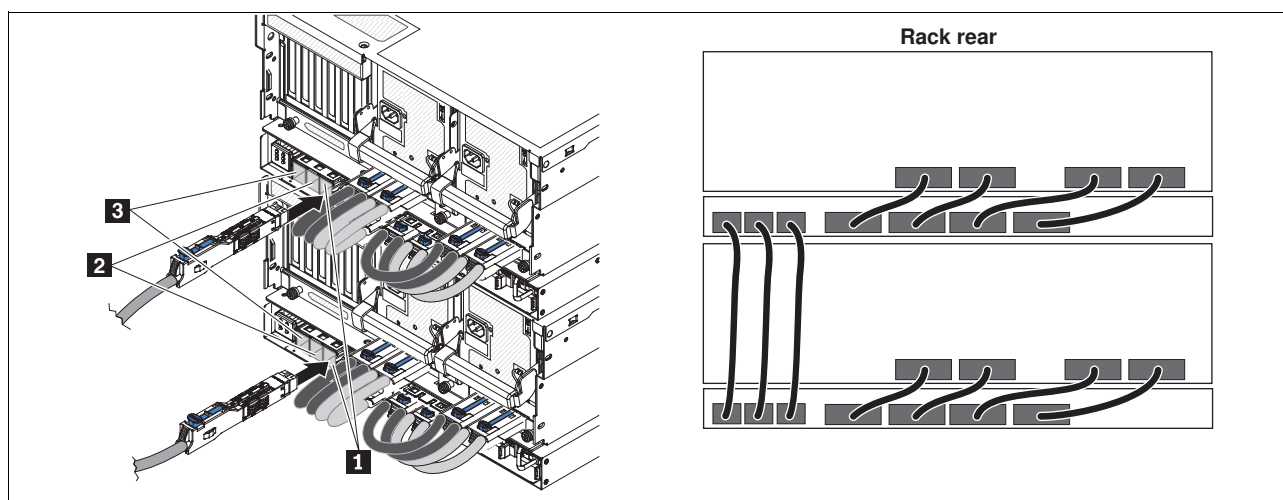


Figure 3-16 Cabling diagram for two-node x3850 X5 with MAX5

Connecting two x3850 X5 servers and two MAX5 units to form a two-node system requires three cable kits, as listed in Table 3-10.

Table 3-10 Cabling that is required for a two-node system with MAX5 configuration

Part number	Feature code	Quantity	Description
59Y6267	4192	2	IBM MAX5 to x3850 X5 cable kit (quantity four cables)
59Y6271	4198	1	IBM eX5 MAX5 two-node EXA scalability kit (three cables)

The MAX5 to x3850 X5 cable kit, 59Y6267, contains the four cables that are needed to connect an x3850 X5 to a MAX5 unit. The EXA scalability kit, 59Y6271, contains the three cables that are needed to connect the two MAX5 units together.

Figure 3-17 shows the block diagram of the configuration.

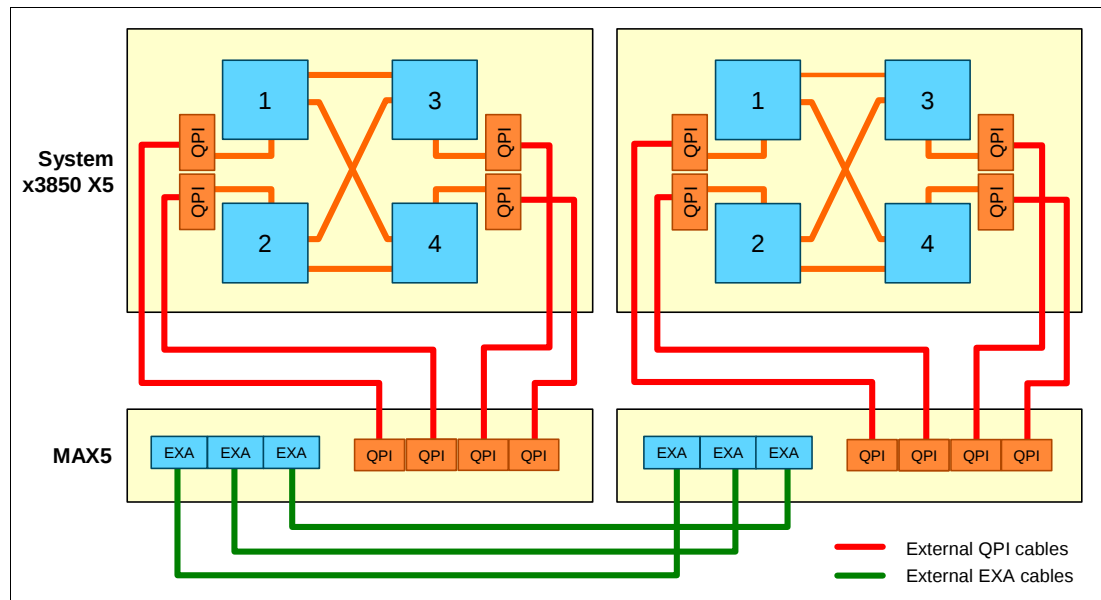


Figure 3-17 Block diagram for the two-node plus MAX5 configuration

3.6.4 FlexNode partitioning

The x3850 X5 supports a configuration of two systems (nodes) that are physically cabled together and acting as a single image eight-processor system. In addition, you can use FlexNode partitioning to reconfigure that two-node system back to being two independent four-processor servers without having to remove cables. This reconfiguration process can be automated, allowing for flexibility when the workloads require it.

The following requirements pertain to FlexNode scaling:

- Configuration of two-node x3850 X5 with MAX5 units
- Machine type 7143 only (7145 is not supported)
- MAX5 V2 only (MAX5, 59Y6265, is not supported)

For more information about partitioning, see 2.6, “Partitioning” on page 30.

3.6.5 MAX5 and Xcel4v Dynamic Server Cache

Another performance feature of MAX5 expansion is the Xcel4v L4 cache. When using a single node (chassis) with a MAX5 expansion unit, the eX5 chipset responds to processor cache snoop requests to accelerate cache updates. When two nodes are used, 256 MB of virtual cache per node (taken from main memory) is used to augment the caches of the processor. In a two-chassis configuration, this amounts to 512 MB of L4 cache.

This feature is a unique IBM enhancement that is not offered by other x86 server architectures that use either Intel or AMD processors.

IBM X3, eX4, and eX5 servers achieve well over 100 #1 results on industry-standard benchmarks, such as TPC-C, TPC-E, TPC-H, SAP SD, vConsolidate, VMmark, and others. See 1.3, “Energy efficiency” on page 6 for details about several popular benchmarks.

3.7 Processor options

The x3850 X5 is supported by two, three, or four processors. The tables in the following subsections show the option part numbers for the supported processors. In a two-node system, you must have eight processors, which must all be identical.

3.7.1 Intel Xeon E7 processor options

Table 3-11 lists the available Intel Xeon E7-4800 and E7-8800 series processor options that are supported on the x3850 X5, machine type 7143 (Xeon E7-4800 and E7-8800 series processors).

Table 3-11 Processor options for machine type 7143 (Xeon E7-4800 and E7-8800 series processors)

Part number	Feature code	Intel Xeon processor description (model, cores, processor speed, L3 cache, memory speed, TDP power rating)	Can scale to two nodes without MAX5	Can scale to two nodes with MAX5 ^a
69Y1889	A14Z	Intel Xeon Processor E7-4807 6C 1.86GHz 18MB 95W	No	Yes
69Y1890	A150	Intel Xeon Processor E7-4820 8C 2.00GHz 18MB 105W	No	Yes
69Y1891	A151	Intel Xeon Processor E7-4830 8C 2.13GHz 24MB 105W	No	Yes
88Y5358	A153	Intel Xeon Processor E7-4850 10C 2.00GHz 24MB 130W	No	Yes
69Y1892	A152	Intel Xeon Processor E7-4860 10C 2.26GHz 24MB 130W	No	Yes
69Y1893	A14T	Intel Xeon Processor E7-4870 10C 2.40GHz 30MB 130W	No	Yes
69Y1896	A14V	Intel Xeon Processor E7-8830 8C 2.13GHz 24MB 105W	Yes	Yes
69Y1894	A14U	Intel Xeon Processor E7-8837 8C 2.67GHz 24MB 130W	Yes	Yes
88Y5357	A154	Intel Xeon Processor E7-8850 10C 2.00GHz 24MB 130W	Yes	Yes
69Y1898	A14X	Intel Xeon Processor E7-8860 10C 2.26GHz 24MB 130W	Yes	Yes
69Y1897	A14W	Intel Xeon Processor E7-8867L 10C 2.13GHz 30MB 105W	Yes	Yes
69Y1899	A14Y	Intel Xeon Processor E7-8870 10C 2.40GHz 30MB 130W	Yes	Yes

a. Requires IBM MAX5 V2, 88Y6529. The initial model, IBM MAX5, 59Y6265, is not supported.

Except for the Intel Xeon E7-4807, all processors that are listed in Table 3-11 on page 68 support Intel Turbo Boost Technology. When a processor operates below its thermal and electrical limits, Turbo Boost can dynamically increase the clock frequency of the processor by 133 MHz on short and regular intervals until an upper limit is reached. See 2.2.5, “Turbo Boost Technology” on page 12 for more information.

All of the E7 processors that are shown in Table 3-11 on page 68 support Intel Hyper-Threading Technology, which is an Intel technology that can improve the parallelization of workloads. When Hyper-Threading is enabled in the Unified Extensible Firmware Interface (UEFI), the operating system treats each processor core as two independently addressable processing units. For more information, see 2.2.4, “Hyper-Threading Technology” on page 12.

The x3850 X5 includes at least two CPUs as standard. Two CPUs are required to access all seven of the PCIe slots (as shown in Figure 3-7 on page 57).

- ▶ 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 70.

3.7.2 Population guidelines

Use these population guidelines when you install processors:

- ▶ Each CPU requires a minimum of two DIMMs to operate.
- ▶ All processors must be identical.
- ▶ Configurations of two, three, or four processors are supported.
- ▶ The number of installed processors dictates how many 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 57.
- ▶ When you install 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 58.
- ▶ When you use a MAX5 memory expansion unit, as shown in Figure 3-12 on page 62, you maximize performance when you have four installed processors. This scenario is true because there are four active QPI links to the MAX5. However, two or three processor configurations are supported.
- ▶ Consider the E7-8837 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. This configuration is preferred rather than using two processors with more cores or a higher core frequency. Having four processors enables all memory channels and maximizes memory bandwidth. We describe this topic in 3.8, “Memory” on page 70.

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.

The following topics are covered:

- ▶ 3.8.1, “Memory cards” on page 70
- ▶ 3.8.2, “Memory DIMMs for the x3850 X5” on page 71
- ▶ 3.8.3, “Memory DIMMs for MAX5” on page 72
- ▶ 3.8.4, “DIMM population sequence” on page 74
- ▶ 3.8.5, “Maximizing memory performance” on page 79
- ▶ 3.8.6, “Memory mirroring” on page 82
- ▶ 3.8.7, “Memory sparing” on page 84
- ▶ 3.8.8, “Effect on performance by using mirroring or sparing” on page 84

3.8.1 Memory cards

The x3850 X5, like its predecessor, the x3850 M2, uses memory cards to which the memory DIMMs are attached, as shown in Figure 3-18.

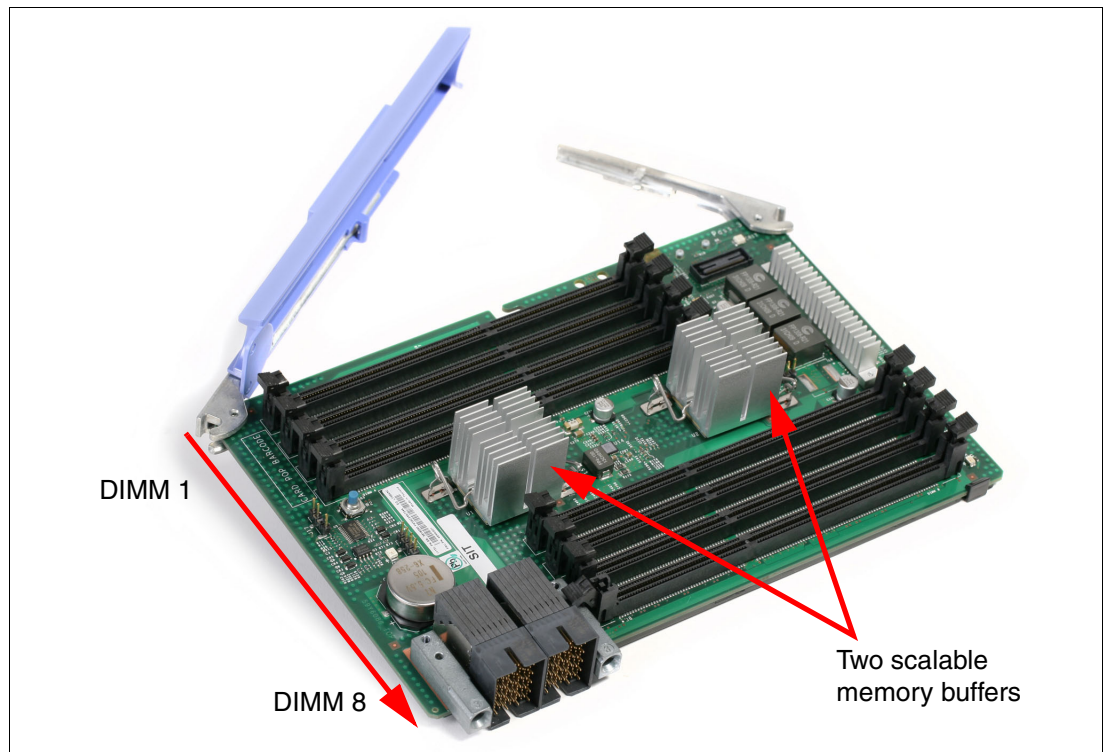


Figure 3-18 x3850 x5 memory card

Standard models contain two or more memory cards. You can configure more cards, as listed in Table 3-12.

Table 3-12 IBM System x3850 X5 and x3950 X5 memory card

Part number	Feature code	Description
69Y1888	A14D	IBM x3850 X5 and x3950 X5 Memory Expansion Card for use in systems with E7 processors (machine type 7143)

The memory cards are installed in the server, as shown in Figure 3-19. Each processor is electrically connected to two memory cards as shown (for example, processor 1 is connected to memory cards 1 and 2).

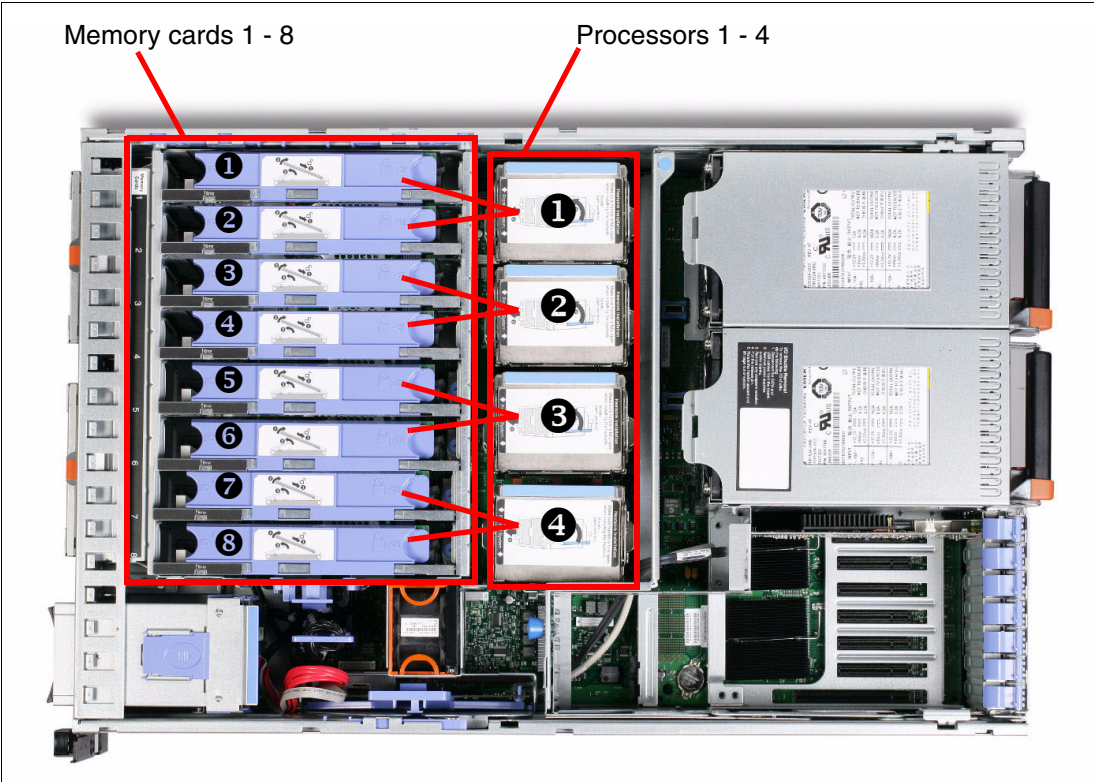


Figure 3-19 Memory card and processor enumeration

3.8.2 Memory DIMMs for the x3850 X5

The memory DIMMs that are available for the x3850 X5 are now described.

When DIMMs with x4 dynamic random access memory (DRAM) modules are used, double device data correction (DDDC) is automatically enabled. For more information about DDDS, see “Redundant bit steering and double device data correction” on page 25.

Several DIMMs listed are low-voltage DIMMs (with “PC3L” in the description). When all DIMMs populated are low voltage, the memory runs at 1.35 V. Otherwise, memory runs at standard voltage, 1.5 V.

Table 3-13 shows the available DIMMs that are supported in the x3850 X5 server with Intel Xeon E7 processors - machine type 7143.

Table 3-13 x3850 X5 (E7 processors with machine type 7143) supported DIMMs

Part number	x3850 X5 feature code	Description
44T1592 ^a	1712	2GB (1x2GB, 1Rx8, 1.5V) PC3-10600 CL9 ECC DDR3 1333MHz LP RDIMM
44T1599 ^a	1713	4GB (1x4GB, Dual Rankx8) PC3-10600 CL9 ECC DDR3-1333MHz LP RDIMM
49Y1407	8942	4GB (1x4GB, 2Rx8, 1.35V) PC3L-10600 CL9 ECC DDR3 1333MHz LP RDIMM
46C7482 ^a	1706	8GB (1x8GB, Quad Rankx8) PC3-8500 CL7 ECC DDR3 1066MHz LP RDIMM
49Y1399	A14E	8GB (1x8GB, 4Rx8, 1.35V) PC3L-8500 CL7 ECC DDR3 1066MHz LP RDIMM
46C7483 ^a	1707	16GB (1x16GB, Quad Rankx4) PC3-8500 CL7 ECC DDR3 1066MHz LP RDIMM
49Y1400	8939	16GB (1x16GB, 4Rx4, 1.35V) PC3L-8500 CL7 ECC DDR3 1066MHz LP RDIMM
49Y1564	A1QT	16GB (4Gb, 2Rx4, 1.35V)PC3L-10600 DDR3-1333 LP RDIMM
90Y3101	A1CP	32GB (1x32GB, 4Rx4, 1.35V) PC3L-8500 CL7 ECC DDR3 1066MHz LP RDIMM

a. This part has been withdrawn from marketing.

The speed at which the memory that is connected to the Xeon E7 and 7500 processors runs depends on the capabilities of the processor. Scalable memory interconnect (SMI) links run from the memory controller that is integrated in the processor to the scalable 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 17.

3.8.3 Memory DIMMs for MAX5

The MAX5 memory expansion unit has 32 DIMM sockets and is designed to augment the installed memory in the attached x3850 X5 server. The following tables show the available memory options that are supported in the MAX5 memory expansion unit, both MAX5, part number 59Y6265 and MAX5 V2, part number 88Y6529.

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 x8 or x4 (but not both at the same time).

Table 3-14 lists the DIMM options that are also supported in the MAX5 V2. When used in the MAX5 V2, the DIMMs have separate feature codes.

Table 3-14 DIMMs supported in MAX5 V2, 88Y6529

Part number	MAX5 V2 feature code	Description
44T1592	2429	2 GB MAX5 1x2 GB 1Rx8 1.5V PC3-10600 CL9 ECC DDR3 1333 MHz LP RDIMM
9Y1407	A1MH	4 GB MAX5 (1x4 GB, 2 Gb, 2Rx8, 1.35 V) PC3L-10600R-999 LP ECC RDIMM
44T1599	2431	4 GB MAX5 1x4 GB DualRankx8 PC310600 CL9 ECC DDR3 1333 MHz LP RDIMM
46C7482	2432	8 GB MAX5 1x8 GB QuadRankx8 PC3-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
49Y1399	A1N7	8 GB MAX5 1x8 GB, 4Rx8, 1.35 V PC3L-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
49Y1564	A3E1	16 GB MAX5 1x16GB 2Rx4 1.35V PC3L-10600 CL9 ECC DDR3 1333MHz LP RDIMM
46C7483	2433	16 GB MAX5 1x16 GB QuadRankx4 PC3-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
49Y1400	A1N8	16 GB MAX5 1x16 GB 4Rx4 1.35 V PC3L-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
90Y3206	A1R2	32 GB MAX5 (4 GB, 4Rx4, 1.35 V) PC3L-8500 DDR3-1066 MHz LP RDIMM

When DIMMs with x4 DRAM modules are used, DDDC is automatically enabled. For more information about DDDS, see “Redundant bit steering and double device data correction” on page 25.

Certain DIMMs listed in Table 3-14 are low-voltage DIMMs (with “PC3L” in the description). When all DIMMs populated are low voltage, the memory runs at 1.35 V. Otherwise, memory runs at 1.5 V. MAX5 V2 (88Y6529) supports low-voltage DIMM operation.

Although 1333 MHz memory DIMMs are supported in the MAX5 V2, the memory DIMMs run at a maximum speed of 1066 MHz. Actual memory speed depends on the processors that are installed in the attached server.

MAX5 V2 memory options: The 16 GB memory options and the 32 GB memory option, 90Y3206, are supported in the MAX5 V2 only when it is the only type of memory (x4 DRAM) that is used in the MAX5 V2. No other memory options can be used in the MAX5 V2 if either of these options are installed in the MAX5 V2.

Table 3-15 indicates the DIMM options that are also supported in the MAX5, 59Y6265. When used in the MAX5, the DIMMs have separate feature codes. When low-voltage DIMMs are used in the MAX5, they run at 1.5 V.

Table 3-15 DIMMs supported in MAX5, 59Y6265

Part number	MAX5 feature code	Description
44T1592	2429	2 GB MAX5 1x2 GB 1Rx8 1.5 V PC3-10600 CL9 ECC DDR3 1333 MHz LP RDIMM
44T1599	2431	4 GB MAX5 1x4 GB DualRankx8 PC310600 CL9 ECC DDR3 1333 MHz LP RDIMM
9Y1407	A1MH	4 GB MAX5 (1x4 GB, 2 Gb, 2Rx8, 1.35 V) PC3L-10600R-999 LP ECC RDIMM
46C7482	2432	8 GB MAX5 1x8 GB QuadRankx8 PC3-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
49Y1399	A1N7	8 GB MAX5 1x8 GB, 4Rx8, 1.35 V PC3L-8500 CL7 ECC DDR3 1066 MHz LP RDIMM

Part number	MAX5 feature code	Description
46C7483	2433	16 GB MAX5 1x16 GB QuadRankx4 PC3-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
49Y1400	A1N8	16 GB MAX5 1x16 GB 4Rx4 1.35 V PC3L-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
90Y3206	A1R2	32 GB MAX5 (4 GB, 4Rx4, 1.35 V) PC3L-8500 DDR3-1066 MHz LP RDIMM

MAX5 memory options: The 16 GB memory options and the 32 GB memory option, 90Y3206, are supported in the MAX5 only when it is the only type of memory (x4 DRAM) that is used in the MAX5. No other memory options can be used in the MAX5 if either of these options are installed in the MAX5.

Redundant Bit Steering: In the x3850 X5 with Intel Xeon E7 processors, DDDC, the Intel implementation of Redundant Bit Steering (RBS), is supported. See “Redundant bit steering and double device data correction” on page 25 for details.

The MAX5 memory expansion unit supports RBS, but only with x4 memory and not x8 memory. As shown in Table 3-14 on page 73 and Table 3-15, the 16 GB DIMMs, and the 32 GB DIMM, part number 90Y3206, use x4 DRAM technology. RBS is automatically enabled in the MAX5, if all installed DIMMs are x4 DIMMs.

3.8.4 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 81 for performance effects when this guideline is not followed.

Tip: The following tables 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*, at this web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5085479>

x3850 X5 single-node and two-node configurations

The information in Table 3-16 on page 75 is the same if you use a single-node configuration or if you use a two-node configuration. In a two-node configuration, you install in the same order twice, once for each server.

Table 3-16 shows the NUMA-compliant memory installation sequence for two processors.

Table 3-16 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 22.

Table 3-17 shows the NUMA-compliant memory installation sequence for three processors.

Table 3-17 NUMA-compliant DIMM installation (three processors) for 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 22.

Three-processor systems: For a three-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 3-18 shows the NUMA-compliant memory installation sequence for four processors.

Table 3-18 NUMA-compliant DIMM installation (four processors) for 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	
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 22.

MAX5 configurations

The memory that is 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 17, the memory speed is derived from the QPI link speed of the installed processors. This process in turn dictates the maximum SMI link speed and dictates the memory speed.

The tables in 3.7, “Processor options” on page 68 summarize the memory speeds of all the available processors.

One important consideration when installing memory in MAX5 configurations is that the server must be fully populated before you add DIMMs to the MAX5. As we described in 2.3.2, “Memory dual inline memory module placement” on page 18, 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 3-20 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.

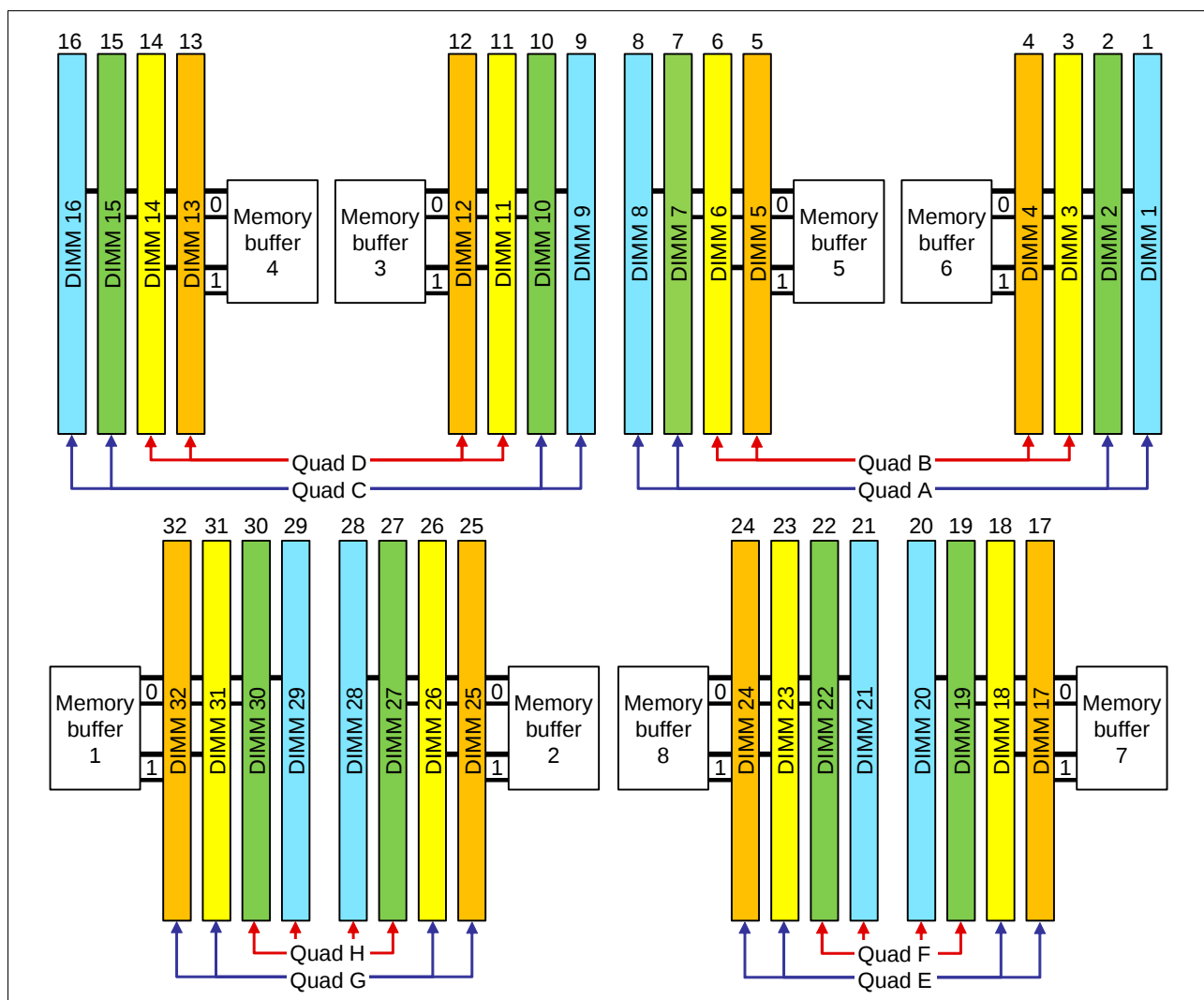


Figure 3-20 DIMM numbering on MAX5

Table 3-19 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 3-20.

Table 3-19 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

DIMM pair	DIMM slots
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

Hemisphere mode

Hemisphere mode is a memory performance feature of the processors that is used in the x3850 X5. In a single node system where memory is installed in quads, hemisphere mode is automatically engaged and typically results in a performance boost.

It is recommended for performance reasons that you configure memory such that hemisphere mode is enabled. The tables in 3.8.4, “DIMM population sequence” on page 74 indicate which configurations support hemisphere mode.

Hemisphere mode can best be thought of as a four-way interleave (as opposed to the standard two-way interleave in this system). The requirement is to install memory options in sets of four, that is, two matched pairs at a time.

When using QPI scaling to join two x3850 X5 nodes together with MAX5 units, memory *must* be configured in hemisphere mode. One matched pair is installed in the primary node and the other is installed in the secondary node. The installation uses the same memory cards and positions on those memory cards across both nodes.

MAX5 memory as seen by the operating system

MAX5 can run in 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.

Important: MAX5 requires VMware vSphere 4.1 or later.

3.8.5 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 57. 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

Installation and configuration of memory DIMMS

The x3850 X5 supports various 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 3-21 on page 80 shows eight possible memory configurations for the two memory cards and 16 DIMMs connected to one processor socket (one processor and two memory cards are shown). Each configuration has a relative performance score. Note the key information from the chart within the figure:

- ▶ 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 than one DIMM per channel (compare configuration 1 with 2, and compare configuration 5 with 6).

Hemisphere mode: Configurations 1 and 2 are the only configurations in which hemisphere mode is enabled. These configurations that are compared to the rest show the importance of hemisphere mode to memory performance. For more information about hemisphere mode, see section 2.3.5, “Hemisphere mode” on page 22.

Figure 3-21 shows relative memory performance based on DIMM placement.

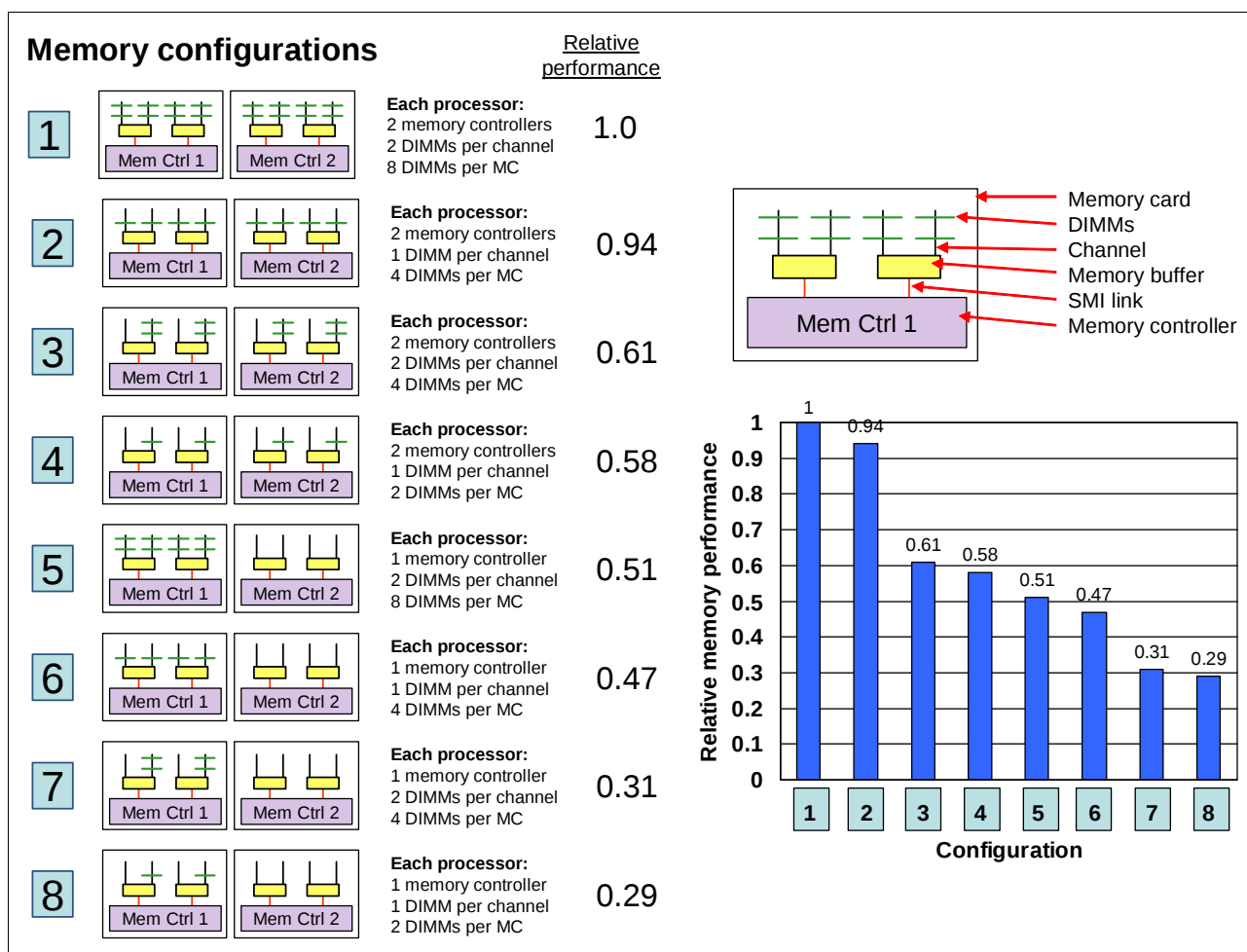


Figure 3-21 Relative memory performance that is based on DIMM placement

Use the following general memory population rules:

- ▶ DIMMs must be installed in matching pairs.
- ▶ Each memory card requires at least two DIMMs.
- ▶ Have identical amounts of RAM for each processor and memory card.
- ▶ 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 the next largest DIMMs, and finish with the smallest capacity DIMMs that you have.

Because 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, there is 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 homogeneous memory population provides.
- ▶ Cost-saving requirements might dictate using a lower memory capacity for several of the DIMMs of the platform.
- ▶ Certain configurations might attempt to use the DIMMs that came with the base platform, along with optional DIMMs of a separate type.

Figure 3-22 on page 82 illustrates the relative performance of three mixed memory configurations as compared to a baseline of a fully populated memory configuration. Although these configurations use 4 GB (4R x8) and 2 GB (2R x8) DIMMs as specified, similar trends to this data are expected when you use 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:

- ▶ 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. However, 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 more 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.

The figure displays four system architectures (A, B, C, D) illustrating the impact of memory configuration on relative performance. Each diagram shows a Processor connected to two Memory controllers, which are further connected to Buffers and DIMMs.

- A**: Relative performance: 100. This configuration uses 8 GB of 2 GB DIMMs (4 banks × 2 DIMMs/bank).
- B**: Relative performance: 97. This configuration uses 8 GB of 4 GB DIMMs (4 banks × 2 DIMMs/bank).
- C**: Relative performance: 92. This configuration uses 8 GB of 4 GB DIMMs (4 banks × 2 DIMMs/bank), with some DIMMs being empty sockets.
- D**: Relative performance: 82. This configuration uses 8 GB of 4 GB DIMMs (4 banks × 2 DIMMs/bank), with some DIMMs being empty sockets.

Legend:

- DIMM 2 GB DIMM
- DIMM 4 GB DIMM
- Empty DIMM socket

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.

Memory mirroring is supported by 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.

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 24.

x3850 X5 memory mirroring population order

Table 3-20 shows DIMM placements for each solution.

Table 3-20 x3850 X5 memory mirroring four-processor two-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
4	x				x											
8	x				x				x							
12	x				x				x				x			
16	x				x				x				x			
20	x	x			x	x			x				x			
24	x	x			x	x			x				x			
28	x	x			x	x			x	x			x			
32	x	x			x	x			x	x			x	x		
36	x	x	x		x	x	x		x	x			x	x		
40	x	x	x		x	x	x		x	x			x	x		
44	x	x	x		x	x	x		x	x	x		x	x		
48	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	
56	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				
64	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x	x

Table 3-21 shows the memory mirroring card pairs.

Table 3-21 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 3-22 shows the installation guide for MAX5 memory mirroring.

Table 3-22 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	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	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	

3.8.7 Memory sparing

Sparing provides a degree of redundancy in the memory subsystem, but not to the extent of mirroring. For more information about memory sparing, see “Memory sparing” on page 24. 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 to all memory cards in the system.

► Rank sparing

Two ranks per memory card are configured as spares. The ranks must 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.

3.8.8 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 four processors and populated with sixty-four 4 GB quad-rank DIMMs.

Figure 3-23 shows the peak system-level memory throughput for various memory modes that are measured by 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.

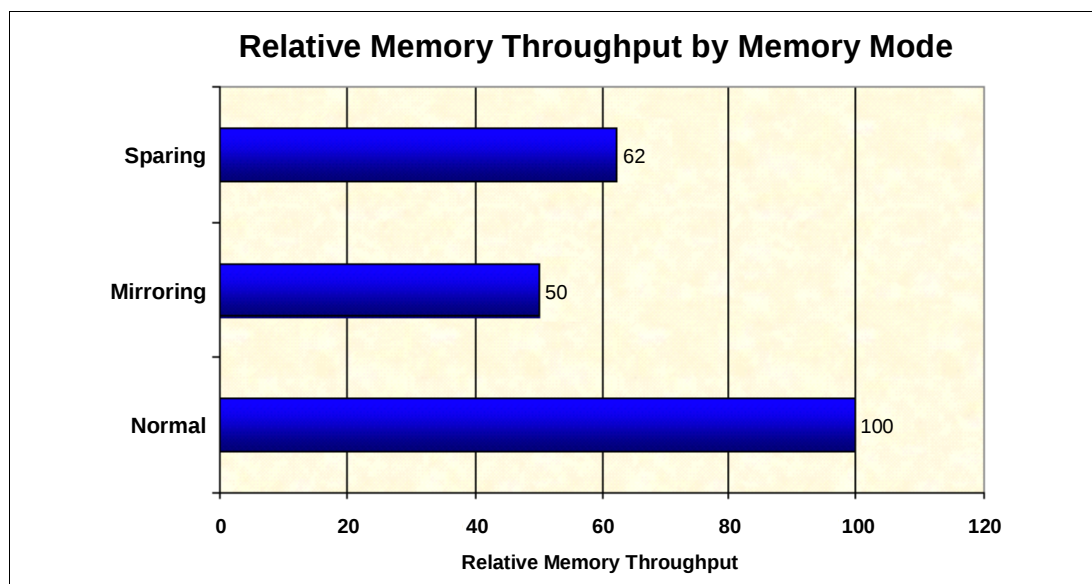


Figure 3-23 Relative memory throughput by memory mode

3.9 Storage

We now 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. The following topics are covered:

- ▶ 3.9.1, “Internal disks” on page 85
- ▶ 3.9.2, “SAS and SSD 2.5-inch disk support” on page 86
- ▶ 3.9.3, “IBM eXFlash and 1.8-inch SSD support” on page 88
- ▶ 3.9.4, “SAS and SSD controllers” on page 92
- ▶ 3.9.6, “External direct-attach storage connectivity” on page 97

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 3-24 shows the internal bays with eight 2.5-inch SAS drives.



Figure 3-24 Front view of the x3850 X5 showing eight 2.5-inch SAS drives

3.9.2 SAS and SSD 2.5-inch disk support

Backplane, controller, and drive options for 2.5-inch disk drives and SSDs are now described.

SAS disks and SSDs that are 2.5-inch use the same backplane options. Most standard models of the x3850 X5 include one SAS backplane that supports four 2.5 inches drives as listed in 3.3, “Models” on page 53. You can add a second identical backplane to increase the supported number of SAS disks to eight (using part number 59Y6135) as shown in “x3850 X5 backplane options” on page 86. The standard backplane is always included in the lower of the two backplane bays.

Table 3-23 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 in. drives

The SAS backplane uses a short SAS cable (included with the part number 59Y6135). The backplane 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 3-24 lists the supported RAID controllers. For more information about each RAID controller, see 3.9.4, “SAS and SSD controllers” on page 92.

Table 3-24 RAID controllers that are compatible with SAS backplane and SAS disk drives

Part number	Feature code	Description
46M0831	0095	ServeRAID M1015 SAS / SATA Controller for System x
46M0916	3877	ServeRAID M5014 SAS / SATA Controller for System x ^a
46M0829	0093	ServeRAID M5015 SAS / SATA Controller for System x
90Y4304	A2NF	ServeRAID M5016 SAS/SATA Controller for System x

Part number	Feature code	Description
88Y5874	A39Q	ServeRAID M5016 Battery Tray ^b
46M0969	3889	ServeRAID B5015 SSD
46M0832	9749	ServeRAID M1000 Series Advance Feature Key
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 M5014.

b. The ServeRAID M5016 Battery Tray is used to house the M5016 power module remotely from the controller. The tray replaces the existing tray that is supplied with the server and supports up to two power modules. Only one ServeRAID M5016 Battery Tray can be installed in the x3850 X5 because the x3850 X5 supports a maximum of two ServeRAID M5016 adapters.

Table 3-25 lists the 2.5-inch drives that are supported in the x3850 X5. These drives are supported by the SAS hard disk backplane, 59Y6135

Table 3-25 Supported 2.5-inch drives

Part number	Feature	Description
Solid-state drive (SSD)		
00W1125	A3HR	IBM 100GB SATA 2.5" MLC HS Enterprise SSD
43W7718	A2FN	IBM 200GB SATA 2.5" MLC HS Enterprise SSD
49Y5839	A3AS	IBM 64GB SATA 2.5" MLC HS Enterprise Value SSD
90Y8648	A2U4	IBM 128GB SATA 2.5" MLC HS Enterprise Value SSD
90Y8643	A2U3	IBM 256GB SATA 2.5" MLC HS Enterprise Value SSD
49Y5844	A3AU	IBM 512GB SATA 2.5" MLC HS Enterprise Value SSD
49Y6129	A3EW	IBM 200GB SAS 2.5" MLC HS Enterprise SSD
49Y6134	A3EY	IBM 400GB SAS 2.5" MLC HS Enterprise SSD
49Y6139	A3F0	IBM 800GB SAS 2.5" MLC HS Enterprise SSD
2.5-inch 15K SAS hot-swap (HS) hard disk drive (HDD)		
90Y8926	A2XB	IBM 146GB 15K 6Gbps SAS 2.5" SFF G2HS HDD
42D0677	5536	IBM 146GB 15K 6Gbps SAS 2.5" SFF Slim-HS HDD
81Y9670	A283	IBM 300GB 15K 6Gbps SAS 2.5" SFF HS HDD
2.5-inch 15K SAS hot-swap self-encrypting drive (SED)		
44W2294	5412	IBM 146 GB 15K 6 Gbps SAS 2.5-inch SFF Slim-HS SED ^a
90Y8944	A2ZK	IBM 146 GB 15K 6Gbps SAS 2.5" SFF G2HS SED ^a
2.5-inch 10K SAS hot-swap HDD		
90Y8877	A2XC	IBM 300GB 10K 6Gbps SAS 2.5" SFF G2HS HDD

Part number	Feature	Description
42D0637	5599	IBM 300GB 10K 6Gbps SAS 2.5" SFF Slim-HS HDD
90Y8872	A2XD	IBM 600GB 10K 6Gbps SAS 2.5" SFF G2HS HDD
49Y2003	5433	IBM 600GB 10K 6Gbps SAS 2.5" SFF Slim-HS HDD
81Y9662	A3EG	IBM 900GB 10K 6Gbps SAS 2.5" SFF G2HS SED
81Y9650	A282	IBM 900GB 10K 6Gbps SAS 2.5" SFF HS HDD
2.5-inch 10K SAS hot-swap SED		
90Y8913	A2XF	IBM 300GB 10K 6Gbps SAS 2.5" SFF G2HS SED ^a
90Y8908	A3EF	IBM 600GB 10K 6Gbps SAS 2.5" SFF G2HS SED ^a
44W2264	5413	IBM 300GB 10K 6Gbps SAS 2.5" SFF Slim-HS SED ^a
2.5-inch NL SAS hot-swap HDD		
81Y9690	A1P3	IBM 1TB 7.2K 6Gbps NL SAS 2.5" SFF HS HDD
90Y8953	A2XE	IBM 500GB 7.2K 6Gbps NL SAS 2.5" SFF G2HS HDD
42D0707	5409	IBM 500GB 7200 6Gbps NL SAS 2.5" SFF Slim-HS HDD
2.5-inch NL Serial ATA (SATA) hot-swap HDDs		
81Y9730	A1AV	IBM 1TB 7.2K 6Gbps NL SATA 2.5" SFF HS HDD
81Y9722	A1NX	IBM 250GB 7.2K 6Gbps NL SATA 2.5" SFF HS HDD
81Y9726	A1NZ	IBM 500GB 7.2K 6Gbps NL SATA 2.5" SFF HS HDD
42D0752	5407	IBM 500GB 7200 NL SATA 2.5" SFF Slim-HS HDD

a. To use the self-encrypting capabilities of these drives, the ServeRAID M1000 Series Advanced Feature Key is required (for the ServeRAID M1015 adapter). Or, either of the ServeRAID M5000 keys are required (for the ServeRAID M5014 or M5015 adapters), as listed in Table 3-24 on page 86.

The 2.5-inch drives use 70% less space than 3.5-inch drives, use half the power, produce less noise, seek faster, and offer increased reliability. SEDs provide cost-effective advanced data security with Advanced Encryption Standard (AES) 128-disk encryption. To use the encryption capabilities, you must also use either a ServeRAID M5014 or M5015 RAID controller. Plus, you must also use either the ServeRAID M5000 Advance Feature Key or the Performance Accelerator Key, or a ServeRAID M5016 controller. SEDs can be used in place of non-SEDs, although the data is not encrypted. Further detail on data encryption on these drives is covered in 3.9.4, "SAS and SSD controllers" on page 92.

For more information about SEDs, see the IBM Redbooks Product Guide, *Self-Encrypting Drives for IBM System x*, TIPS0761, available at this web page:

<http://www.ibm.com/redbooks/abstracts/tips0761.html>

3.9.3 IBM eXFlash and 1.8-inch SSD support

Support for eXFlash and 1.8-inch SSDs is now described.

IBM eXFlash SSD offerings

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 3-4 on page 54. 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 3-26).

Table 3-26 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 and configuration cable, which are both shipped standard. Up to two SSD backplanes and 16 SSDs are supported in the x3850 X5.

Table 3-27 lists the 1.8-inch SSD options that are supported in the x3850 X5. These drives are supported by the eXFlash 8 disk backplane, part number 59Y6213.

Table 3-27 Supported 1.8-inch SSDs

Part number	Feature code	Description
00W1120	A3HQ	IBM 100GB SATA 1.8" MLC Enterprise SSD
49Y6119	A3AN	IBM 200GB SATA 1.8" MLC Enterprise SSD
49Y6124	A3AP	IBM 400GB SATA 1.8" MLC Enterprise SSD
49Y5834	A3AQ	IBM 64GB SATA 1.8" MLC Enterprise Value SSD
00W1222	A3TG	IBM 128GB SATA 1.8" MLC Enterprise Value SSD
00W1227	A3TH	IBM 256GB SATA 1.8" MLC Enterprise Value SSD
49Y5993	A3AR	IBM 512GB SATA 1.8" MLC Enterprise Value SSD
43W7726	5428	IBM 50GB SATA 1.8" MLC SSD ^a
43W7746	5420	IBM 200GB SATA 1.8" MLC SSD ^a

a. These SSDs are designated simple swap, which means they can be removed and replaced without tools. All the other SSDs listed are hot swap.

The failure rate of SSDs is low because, in part, the drives have no moving parts. These SSDs feature enterprise-grade multi-layer cell (eMLC) NAND flash chips. The SSDs also include discrete capacitors (to assure there is enough energy to fully commit writes to the cells in the event of a power disruption) and reliability features. Examples include data error checking and correction, I/O path error checking and correction, and thermal monitoring and reporting.

Controllers inside the SSDs use wear level algorithms and record and report cell usage counts. With the use of these technologies, the using RAID redundancy might not always be necessary. Therefore, in certain cases, RAID level 0 might be an acceptable solution.

Enterprise Value SSDs and Enterprise SSDs have similar read and write IOPS performance. However, the key difference between them is their endurance, that is, how long they can perform write operations because SSDs have a finite number of program and erase cycles. Enterprise Value SSDs have a better cost/IOPS ratio but lower endurance when compared to Enterprise SSDs.

For more information about eXFlash and SSD information, including a brief overview of the benefits of using eXFlash, see 2.8, “IBM eXFlash” on page 38.

Spanning an array between two chassis: Spanning an array on any disk type between two chassis (two-node configuration) is not possible with hardware RAID adapters. Spanning is not possible because the RAID controllers in each node operate separately. This limitation also applies to multiple RAID adapters within an x3850 X5. Software array spanning can be used in these cases.

Figure 3-25 shows an x3850 X5 with one of two eXFlash units installed.

Hot-swap capabilities: With the introduction of the SSDs listed in Table 3-27, the drives support hot-swap capabilities. Therefore, the eXFlash trays have orange handles and not blue handles as shown in Figure 3-25.



Figure 3-25 IBM eXFlash with eight SSDs

Table 3-28 lists the supported controllers.

Table 3-28 Controllers that are supported by the eXFlash SSD backplane option

Part number	Feature code	Description
46M0912	3876	IBM 6 Gb Performance Optimized HBA (No RAID support)
46M0916	3877	ServeRAID M5014 SAS / SATA Controller ^a
46M0829	0093	ServeRAID M5015 SAS / SATA Controller ^a
46M0969	3889	ServeRAID B5015 SSD ^a

Part number	Feature code	Description
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
90Y4304	A2NF	ServeRAID M5016 SAS/SATA Controller for System x
88Y5874	A39Q	ServeRAID M5016 Battery Tray ^{ab}

- a. When using SSD drives, disable the write-back cache to reduce latency, by using the controller settings or by adding the ServeRAID M5000 Series Performance Accelerator Key. The ServeRAID M5016 includes the Performance Accelerator Key functionality. See “ServeRAID M5000 Series Performance Accelerator Key” on page 92 for more information.
- b. The ServeRAID M5016 Battery Tray is used to house the M5016 power module remotely from the controller. The tray replaces the existing tray that is supplied with the server and supports up to two power modules. Only one ServeRAID M5016 Battery Tray can be installed in the x3850 X5 because the x3850 X5 supports a maximum of two ServeRAID M5016 adapters.

When using the ServeRAID M5014, M5015, or M5016 with SSDs only, do not use the write-back caching, for performance reasons. If using an M5014 controller in a mixed SSD and SAS environment cache, order the battery along with the Performance Accelerator Key. The M5015 comes standard with a cache battery, and is write-caching enabled. The M5016 has flash backed caching enabled by default. If the ServeRAID controller that is being used is already set up and you want to disable the write-back cache, use the MegaRAID web basic input/output system (BIOS) or MegaRAID Storage Manager. See Figure 3-26.

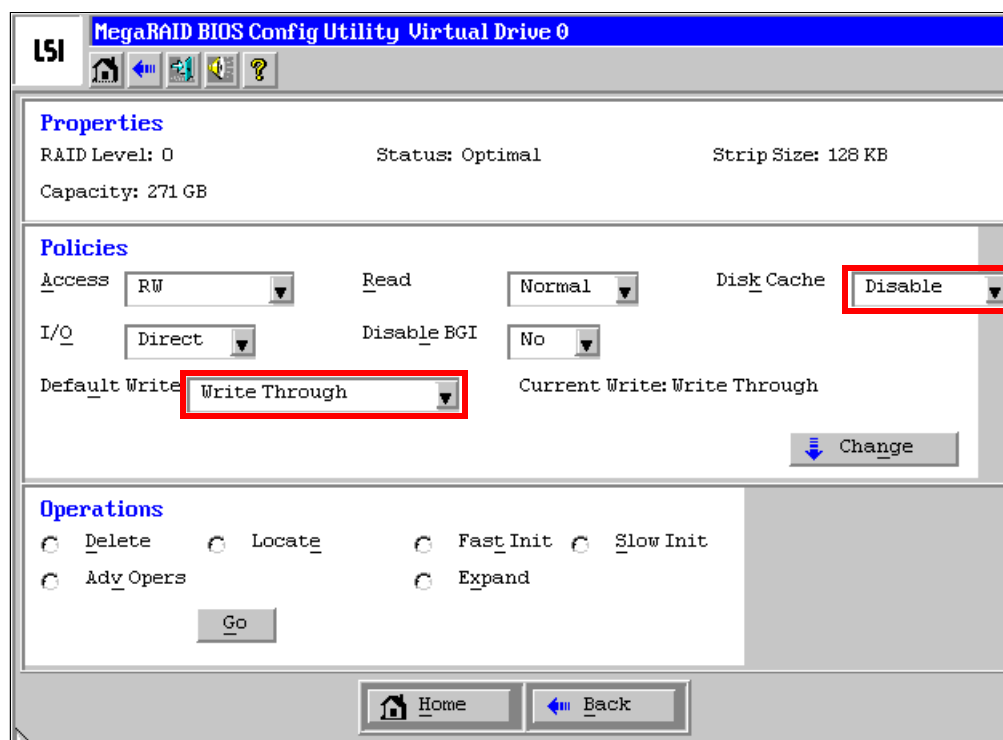


Figure 3-26 Disabling battery cache on the controller in MegaRAID web BIOS

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 the following benefits:

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

Keys enabled for ServeRAID M5016: The ServeRAID M5016 comes with the M500 Series Performance Accelerator Key and Advanced Feature Key functionality enabled.

We cover these controllers in detail in 3.9.4, “SAS and SSD controllers” on page 92.

3.9.4 SAS and SSD controllers

Table 3-29 lists the disk controllers that are supported in the x3850 X5 for internal storage connectivity.

Table 3-29 Disk controllers that are compatible with the x3850 X5

Part number	Feature code	Name	Supports 2.5 in. SAS backplane	Supports eFlash SSD backplane	Dedicated slot ^a	Write-cache protection	Cache	RAID support
44E8689	3577	ServeRAID BR10i	Yes	No	Yes	No	None	0, 1, 1E
46M0831	0095	ServeRAID M1015	Yes	No	Yes	No	None	0, 1, 10, 5, 50 ^b
46M0916	3877	ServeRAID M5014	Yes	Yes	Yes	Battery, Optional	256 MB	0, 1, 10, 5, 50, 6, 60 ^c
46M0829	0093	ServeRAID M5015	Yes	Yes	Yes	Yes ^d , Battery	512 MB	0, 1, 10, 5, 50, 6, 60 ^c
90Y4304	A2NF	ServeRAID M5016	Yes	Yes	Yes	Capacitor ^e	1 GB	0, 1, 10, 5, 50, 6, 60
46M0912	3876	IBM 6 Gb Performance Optimized 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 96.

b. M1015 support for RAID-5 and RAID-50 requires the M1000 Advanced Feature Key (46M0832, 9749).

c. M5014, M5015, and M5025 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 ServeRAID M5016 uses a capacitor to power the adapter long enough to back up the contents of the write cache to a flash module. This process eliminates the need for consumable batteries.

RAID levels 0 and 1 are standard on all models. All servers include the blue mounting bracket (see Figure 3-27 on page 97). The bracket 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 by the 2.5-inch SAS backplane can be used in this slot. See Table 3-29 on page 92 for a summary of these supported options.

ServeRAID M1015 Controller

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

- ▶ Eight internal 6 Gbps SAS and 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
- ▶ Logical unit number (LUN) sizes up to 64 TB
- ▶ Configurable stripe size up to 64 KB
- ▶ Compliant with Disk Data Format (DDF) configuration on disk (COD)
- ▶ Self-Monitoring, Analysis, and Reporting Technology (S.M.A.R.T.) support

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

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

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

ServeRAID M5014 and M5015 controllers

The ServeRAID M5014 and M5015 adapters 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 IBM PowerPC® processor with LSI SAS2108 6 Gbps 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 you use all SSD drives. If you use 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 RAID-60 support with the optional M5000 Advanced Feature Key)
- ▶ Connection of up to 32 SAS or SATA drives
- ▶ SAS and SATA drive support (although 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 DDF COD
- ▶ S.M.A.R.T. support
- ▶ Support for the optional M5000 Series Performance Accelerator Key, which is recommended when you use SSD drives in a mixed environment with SAS and SSD. The following features are enabled:
 - 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: The Performance Accelerator Key uses the same features as the Advanced Feature Key. However, the Performance Accelerator Key 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 web page:

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

ServeRAID M5016 controller

The ServeRAID M5016 adapter has the following specifications:

- ▶ Eight internal 6 Gbps SAS/SATA ports
- ▶ Two Mini-SAS internal connectors (SFF-8087)

- ▶ Six Gbps throughput per port
 - ▶ An 800 MHz dual-core PowerPC processor with LSI SAS2208 6 Gbps RoC controller
 - ▶ PCI Express x8 Gen 2 host interface
 - ▶ One GB of onboard data cache (DDR3 running at 1333 MHz)
 - ▶ CacheVault technology to protect data in cache in case of critical power or server failure
- CacheVault flash cache protection uses NAND flash memory that is powered by a supercapacitor to protect data that is stored in the controller cache. This module eliminates the need for a lithium-ion battery that is commonly used to protect DRAM cache memory on PCI RAID controllers.

To avoid the possibility of data loss or corruption during a power or server failure, CacheVault technology transfers the contents of the DRAM cache to NAND flash (CacheVault flash module (CVFM)). This process is done by using power from the CacheVault power module (CVPM). After the power is restored to the M5016 RAID controller, CacheVault technology transfers the contents of the NAND flash back to the DRAM, which will then be flushed to disk.

- ▶ Supports RAID levels 0, 1, 5, 6, 10, 50, and 60
- ▶ Connects to up to 128 SAS or SATA drives
- ▶ Intermix of SAS and SATA drives are supported, but the mixing of SAS and SATA drives in the same RAID array is not supported
- ▶ Supports up to 64 logical volumes
- ▶ Supports LUN sizes up to 64 TB
- ▶ Configurable stripe size up to 1 MB
- ▶ Compliant with DDF COD
- ▶ S.M.A.R.T. support
- ▶ 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 Product Guide *ServeRAID M5016 SAS/SATA Controller*, TIPS0847, which is available at the following web page:

<http://www.redbooks.ibm.com/abstracts/tips0847.html>

IBM 6 Gb Performance Optimized Host Bus Adapter

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

The 6 Gb Performance Optimized HBA has the following high-level specifications:

- ▶ PCI Express 2.0 host interface
- ▶ A 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 533 MHz
- ▶ UEFI support

For more information, see the IBM Redbooks Product Guide *IBM 6 Gb Performance Optimized HBA*, TIPS0744, which is available at the following web page:

<http://www.redbooks.ibm.com/abstracts/tips0744.html>

Important: Two variants of the 6 Gb host bus adapter exist. The SSD variant (part number 46M0912) has no external port. Do not confuse this variant 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 RAID-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
- ▶ 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 that are supported in a system
- ▶ Support for up to four arrays and logical volumes

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

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

Important: The ServeRAID B5015 SSD 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 3-29 on page 92, certain supported controllers can be installed in a single PCIe x8 dedicated slot on the side of the server, near the front.

Figure 3-27 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 adapter). The latch and edge clips allow the card to be removed and replaced with another supported card as required.

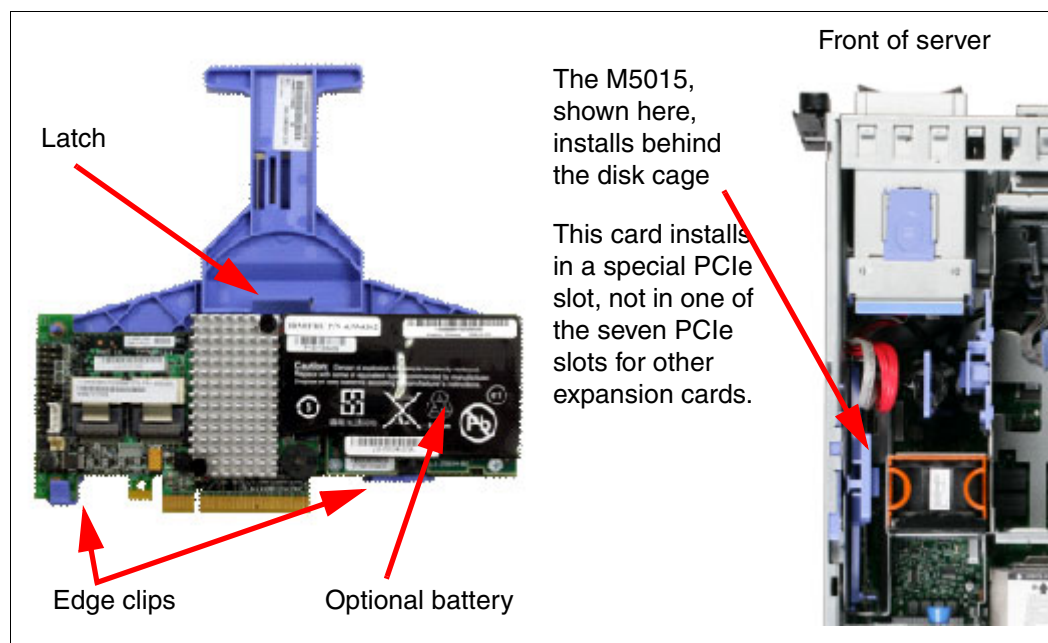


Figure 3-27 ServeRAID M5015 SAS and SATA Controller

3.9.6 External direct-attach storage connectivity

The ServeRAID M5025 offers two external SAS ports to connect to external storage. Table 3-30 lists the cards and support cables and feature keys.

Table 3-30 External ServeRAID card

Option	Feature code	Description
46M0830	0094	IBM 6 Gb ServeRAID M5025 External RAID
39R6531	3707	IBM 3 m SAS External Cable for ServeRAID M5025 to an EXP2512 (1747-HC1) or EXP2524 (1747-HC2)
39R6529	3708	IBM 1 m 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 that are implemented through two four-lane (x4) connectors
- Two mini-SAS external connectors (SFF-8088)
- Six Gbps throughput per SAS port

- ▶ A 800 MHz PowerPC processor with LSI SAS2108 6 Gbps RoC controller
- ▶ PCI Express 2.0 x8 host interface
- ▶ A 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 the optional M5000 Advanced Feature Key or the optional M5000 Performance Key)
- ▶ Connections:
 - Up to 240 SAS or SATA drives
 - Up to nine daisy-chained enclosures per port
- ▶ SAS and SATA drives 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 DDF 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
 - 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 that uses the optional M5000 Series Performance Accelerator Key, which enables the following features:
 - RAID levels 6 and 60
 - Performance optimizations for SSDs
 - 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 Product Guide *ServeRAID M5025 SAS/SATA Controller for IBM System x*, TIPS0739, available at this web page:

<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 System x3850/3950 X5 (7145) Configuration and Options Guide, which is available at this web page:

<http://www-947.ibm.com/support/entry/portal/docdisplay?lnodocid=scod-3zvq5w>

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 3-31 lists the supported part numbers.

Table 3-31 Optical drives

Part number	Feature code	Description
46M0901	4161	IBM UltraSlim Enhanced SATA DVD-ROM
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 10 Gb Ethernet Adapter that is standard in most models (see 3.3, “Models” on page 53). We describe the Emulex 10 Gb Ethernet Adapter in 3.12.1, “Emulex 10 GbE Integrated Virtual Fabric Adapter II” on page 101.

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 use one of the seven available PCIe slots. For further details about supported RAID cards, see 3.9.4, “SAS and SSD controllers” on page 92.

Table 3-32 lists the PCIe slots.

Table 3-32 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 10 Gb Ethernet Adapter)
Dedicated	PCI Express 2.0 x8	Dedicated RAID controller internal 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 5 Gbps per lane. PCI Express 1.1 is rated for 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). This 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 is extended in length to 106 pins, making it a nonstandard connector. It still accepts PCIe x8, x4, and x1 standard adapters. It is the only slot that is compatible with the extended edge connector on the Emulex 10 Gb 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. They 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 3-33 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 3-33 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 that is expressed as GB for gigabyte or Gb for gigabit. Ten bits of traffic correspond to one byte of data because of the 8:10 encoding scheme. A single PCIe 2.0 lane provides a unidirectional bandwidth of 500 MBps or 5 Gbps.

Two extra power connectors, one 2x4 and one 2x3, are provided on the system board for high-power adapters, such as graphics cards. If you are required to use an x16 PCIe card that is not shown as supported in ServerProven, initiate the ServerProven Opportunity Request For Evaluation (SPORE) process.

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

<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 qualified on the x3850 X5. You can obtain the support caveats for third-party options in 3.12.2, "Optional adapters" on page 104.

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

3.12 I/O cards

The I/O cards that are suitable for the x3850 X5 are now described.

3.12.1 Emulex 10 GbE Integrated Virtual Fabric Adapter II

As described in 3.3, “Models” on page 53, most models include the Emulex 10 GbE Integrated Virtual Fabric Adapter II. The card is installed in PCIe slot 7. Slot 7 is a nonstandard x8 slot that is slightly longer than normal, as shown in Figure 3-28.

The integrated 10 Gb adapter is a custom version of the equivalent adapter available as a System x option:

- The Emulex 10 GbE Integrated Virtual Fabric Adapter II (feature code A148, standard in most models) has the same features and functions as the Emulex 10 Gb Virtual Fabric Adapter II for IBM System x, part number 49Y7950.

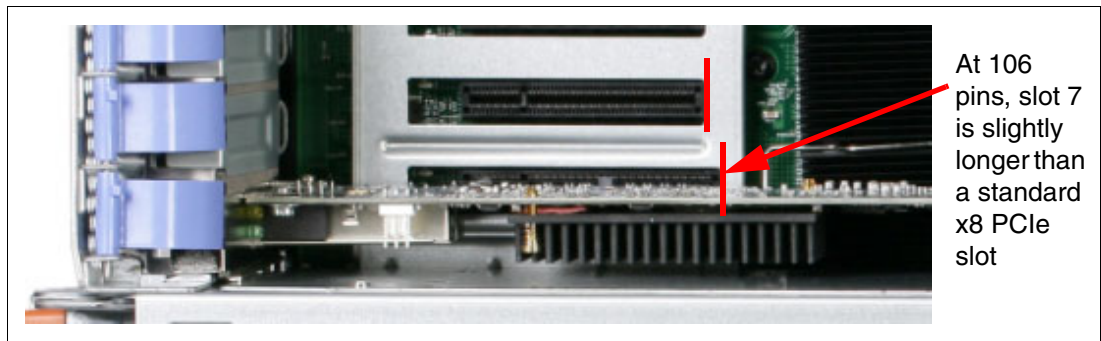


Figure 3-28 Top view of slot 6 and 7 showing that slot 7 is slightly longer than slot 6

The integrated 10 Gb Ethernet Adapter in the x3850 X5 is called an *extended edge connector* because it is customized to be longer than a usual 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, as shown in Figure 3-29.

Only the x3850 X5 and the x3690 X5 have slots that are compatible with the custom-built integrated 10 Gb Ethernet Adapter.

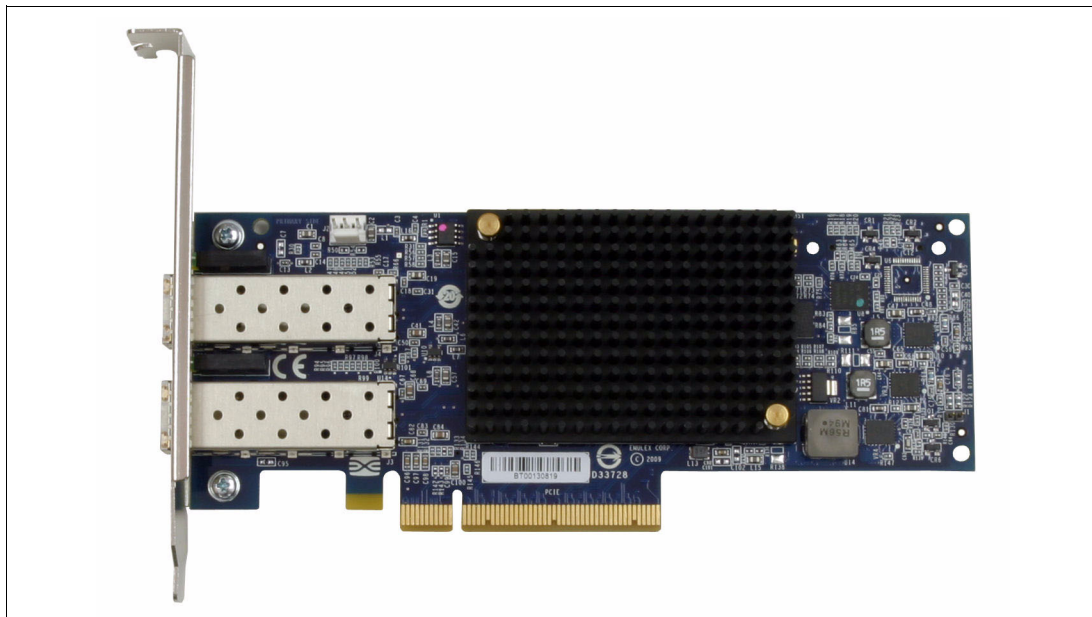


Figure 3-29 Emulex 10 GbE Integrated Virtual Fabric Adapter II

General details about this card can be found in *Emulex 10GbE Virtual Fabric Adapter II and III family for IBM System x*, which is available at the following web page:

<http://www.redbooks.ibm.com/abstracts/tips0844.html>

Fibre Channel over Ethernet and Internet Small Computer System Interface

upgrades: The Emulex 10 GbE Virtual Fabric Adapter II card supports the Internet Small Computer System Interface (iSCSI) hardware initiator or Fibre Channel over Ethernet (FCoE) feature upgrade. The part number for this upgrade is 49Y4265, feature code 2436.

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

- ▶ Dual-channel, 10 Gbps Ethernet controller
- ▶ Line rate 10 Gbps performance
- ▶ Two small form-factor pluggable+ (SFP+) empty cages to support either of the following items:
 - SFP+ SR link with SFP+ SR Module with LC connectors
 - SFP+ twinaxial copper link with SFP+ direct-attached copper module/cable

Transceivers: Servers that include the Emulex 10Gb Ethernet Adapter do not include transceivers. You must order transceivers separately if needed, as listed in Table 3-34.

- ▶ TCP/IP stateless offloads
- ▶ TCP chimney offload

- ▶ Based on Emulex OneConnect technology
 - Deployment of this adapter and other Emulex OneConnect-based adapters with OneCommand Manager
- ▶ iSCSI hardware initiator or FCoE support as a feature entitlement upgrade
- ▶ Hardware parity, cyclic redundancy check (CRC), ECC, and other advanced error checking
- ▶ PCI Express 2.0 x8 host interface
- ▶ Low-profile form-factor design
- ▶ IPv4/IPv6 TCP and User Datagram Protocol (UDP) checksum offload
- ▶ Virtual local area network (VLAN) insertion and extraction
- ▶ Support for jumbo frames up to 9000 bytes
- ▶ Preboot Execution Environment (PXE) 2.0 network boot support
- ▶ Interrupt coalescing
- ▶ Load balancing and failover support
- ▶ Interoperability with IBM Systems Networking 10 Gb Top of Rack (ToR) switch for FCoE functions
- ▶ Interoperability with Cisco Nexus 5000 and Brocade 10 Gb Ethernet switches for NIC/FCoE
- ▶ Support for two types of virtual NIC (vNIC) operating modes, and a physical NIC (pNIC) operating mode:

- IBM Virtual Fabric Mode

Also known as vNIC1 mode. Works with the IBM RackSwitch™ G8124E and G8264. In this mode, the Emulex adapter communicates with the IBM switch to obtain vNIC parameters (using DCBX). A special tag is added within each data packet and is later removed by the NIC or switch for each vNIC group to maintain separation of the virtual data paths.

Each physical port is divided into four virtual ports for a maximum of eight virtual NICs per adapter. Bandwidth for each vNIC can be configured by the IBM switch from 100 Mbps to 10 Gbps. The vNICs can also be configured to have 0 bandwidth if you allocate the available bandwidth to fewer than four vNICs per physical port. Bandwidth allocations can be dynamically changed through the IBM switch. Rebooting the server is not required for the change to take effect.

Storage protocols (FCoE and iSCSI) on vNICs are not supported.

- Switch Independent Mode

Also known as vNIC2 mode. Works with any 10 Gb Ethernet switch. Switch Independent Mode offers the same capabilities as IBM Virtual Fabric Mode in terms of the number of vNICs and the bandwidth each can be configured to have. Switch Independent Mode extends the existing client VLANs to the virtual NIC interfaces. The IEEE 802.1Q VLAN tag is essential to the separation of the vNIC groups by the NIC adapter or driver and the switch. The VLAN tags are added to the packet by the applications or drivers at each endstation rather than by the switch.

vNIC bandwidth allocation and metering is only performed by the adapter itself. In such a case, Bandwidth management is only performed for the outgoing traffic on the adapter (server-to-switch). The incoming traffic (switch-to-server) uses the all available physical port bandwidth because there is no metering that is done on either the adapter or the switch side.

In vNIC2 mode, when storage protocols are enabled on the Emulex 10GbE Virtual Fabric Adapters, six vNICs (three per physical port) are Ethernet, and two vNICs (one per physical port) are either iSCSI or FCoE.

– pNIC mode

The adapter operates as a standard dual-port 10 Gbps Ethernet adapter, and it functions with any 10 GbE switch. In pNIC mode, with the Emulex FCoE/iSCSI License, the card operates in a traditional Converged Network Adapter (CNA) mode with two Ethernet ports and two storage ports (iSCSI or FCoE) available to the operating system.

SFP+ transceivers are not included with the server. You must order them separately. Table 3-34 lists the compatible transceivers.

Table 3-34 Transceiver ordering information

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

3.12.2 Optional adapters

Table 3-35 list the network adapters that are available for the x3850 X5.

Table 3-35 Network adapters

Part number	Feature code	Description	Maximum supported
10 Gb Ethernet			
49Y7910	A18Y	Broadcom NetXtreme II Dual Port 10GBaseT Adapter for IBM System x	7
42C1820	1637	Brocade 10 Gb CNA for IBM System x	7
49Y7950	A18Z	Emulex 10 GbE Virtual Fabric Adapter II for IBM System x	7
Standard	A148	Emulex 10 GbE Integrated Virtual Fabric Adapter II for IBM System x	1
95Y3751	A348	Emulex Dual Port VFAll Adapter and FCoE/iSCSI License for IBM System x	7
49Y7960	A2EC	Intel X520 Dual Port 10 GbE SFP+ Adapter for IBM System x	7
49Y7970	A2ED	Intel X540-T2 Dual Port 10GBaseT Adapter for IBM System x	7
00D9690	A3PM	Mellanox ConnectX-3 10 GbE Adapter for IBM System x	7
42C1800	5751	QLogic 10 Gb CNA for IBM System x	7
Converged Network Adapters (CNAs)			
42C1820	1637	Brocade 10 Gb CNA for IBM System x	7
42C1800	5751	QLogic 10 Gb CNA for IBM System x	7
1 Gb Ethernet			
90Y9370	A2V4	Broadcom NetXtreme I Dual Port GbE Adapter for IBM System x	7
90Y9352	A2V3	Broadcom NetXtreme I Quad Port GbE Adapter for IBM System x	7

Part number	Feature code	Description	Maximum supported
49Y4230	5767	Intel Ethernet Dual Port Server Adapter I340-T2 for IBM System x	7
49Y4240	5768	Intel Ethernet Quad Port Server Adapter I340-T4 for IBM System x	7
42C1780	2995	NetXtreme II 1000 Express Dual Port Ethernet Adapter	7
42C1750	2975	PRO/1000 PF Server Adapter by Intel	7
InfiniBand			
95Y3750	A2MY	Mellanox ConnectX-2 Dual-port QSFP QDR IB Adapter for IBM System x	1
00D9550	A3PN	Mellanox ConnectX-3 FDR VPI IB/E Adapter for IBM System x	7

Table 3-36 list the storage HBAs that are available for the x3850 X5.

Table 3-36 Storage adapters

Part number	Feature code	Description	Maximum supported
16 Gb Fibre Channel			
81Y1675	A2XV	Brocade 16 Gb FC Dual-port HBA for IBM System x	7
81Y1668	A2XU	Brocade 16 Gb FC Single-port HBA for IBM System x	7
81Y1662	A2W6	Emulex 16 Gb FC Dual-port HBA for IBM System x	7
81Y1655	A2W5	Emulex 16 Gb FC Single-port HBA for IBM System x	7
00Y3341	A3KX	QLogic 16 Gb FC Dual-port HBA for IBM System x	7
00Y3337	A3KW	QLogic 16 Gb FC Single-port HBA for IBM System x	7
8 Gb Fibre Channel			
46M6050	3591	Brocade 8 Gb FC Dual-port HBA for IBM System x	7
46M6049	3589	Brocade 8 Gb FC Single-port HBA for IBM System x	7
42D0494	3581	Emulex 8 Gb FC Dual-port HBA for IBM System x	7
42D0485	3580	Emulex 8 Gb FC Single-port HBA for IBM System x	7
42D0510	3579	QLogic 8 Gb FC Dual-port HBA for IBM System x	7
42D0501	3578	QLogic 8 Gb FC Single-port HBA for IBM System x	7
4 Gb Fibre Channel			
59Y1993	3886	Brocade 4 Gb FC Dual-port HBA for IBM System x	7
59Y1987	3885	Brocade 4 Gb FC Single-port HBA for IBM System x	7
42C2071	1699	Emulex 4 Gb FC Dual-Port PCI-E HBA for IBM System x	7
42C2069	1698	Emulex 4 Gb FC Single-Port PCI-E HBA for IBM System x	7
39R6527	3568	QLogic 4 Gb FC Dual-Port PCIe HBA for System x	7
39R6525	3567	QLogic 4 Gb FC Single-Port PCIe HBA for System x	7
SAS			

Part number	Feature code	Description	Maximum supported
46M0912	3876	IBM 6 Gb Performance Optimized HBA	4
46M0907	5982	IBM 6 Gb SAS HBA	7

Table 3-37 list the PCIe High IOPS storage adapters that are available for the x3850 X5.

Table 3-37 High IOPS SSD adapters

Part number	Feature codes	Description	Maximum supported
46M0877	0096	IBM 160 GB High IOPS SS Class Adapter	7
46M0898	1649	IBM 320 GB High IOPS MS Class Adapter	7
46M0878	0097	IBM 320 GB High IOPS SD Class Adapter	7
81Y4535	A1NE	IBM 320 GB High IOPS SLC Adapter	7
46C9078	A3J3	IBM 365 GB High IOPS MLC Mono Adapter	7
81Y4539	A1ND	IBM 640 GB High IOPS SLC Duo Adapter	5
46C9081	A3J4	IBM 785 GB High IOPS MLC Mono Adapter	7
90Y4377	A3DY	IBM 1.2 TB High IOPS MLC Mono Adapter	7
90Y4397	A3DZ	IBM 2.4 TB High IOPS MLC Duo Adapter	2

These lists are constantly updated and changed. To see the latest updates, visit the following web page:

http://www.ibm.com/systems/xbc/cog/x3850x5_7143/x3850x5_7143io.html

Tools, such as the COG or SSCT, contain information about supported part numbers. Many System x tools, including those tools that we mentioned, are on the following configuration tools web page:

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

See the ServerProven web page for a complete list of 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 were successfully tested by IBM with a System x server.

Another useful page on the ServerProven site is the list of vendors. On the home page for ServerProven, click the **industry leaders** link, as shown in Figure 3-30.

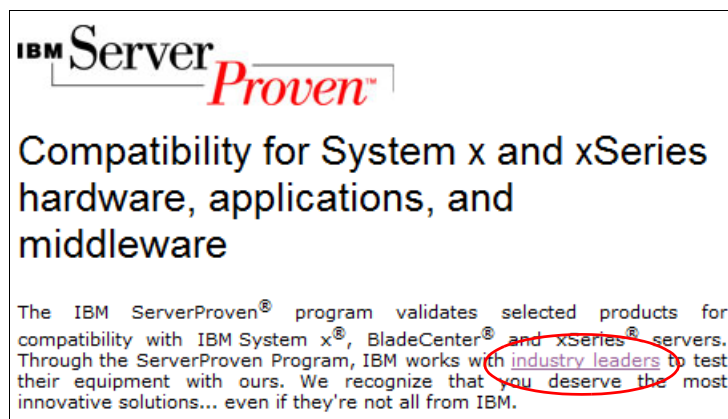


Figure 3-30 Link to vendor testing results

The resulting 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.

You can access this page directly at this web page:

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

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

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

3.13 Standard onboard features

Several standard features in the x3850 X5 are described.

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 offload 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 is characterized by the following environmental data:

- ▶ Heat output:
 - Minimum configuration: 734 BTU per hour (215 watts per hour)
 - Typical configuration: 2,730 BTU per hour (800 watts per hour)
 - Design maximum configuration:
 - 5,971 BTU per hour (1930 watts per hour) at 110 V ac
 - 6,739 BTU per hour (2150 watts per hour) at 220 V ac
- ▶ Electrical input: 100 - 127 V, 200 - 240 V, and 50 - 60 Hz
- ▶ Approximate input Kilovolt amperes (kVAs):
 - 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

The System x3850 X5 includes an integrated management module (IMM) that provides industry-standard IPMI 2.0-compliant Systems Management. You access the IMM through software compatible with IPMI 2.0 (xCAT, for example). You implement the IMM by using industry-leading firmware from OSA and applications with the IMM.

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 over the system management LAN connections. The operating system must support USB.

With video compression now built into the adapter hardware, the IMM 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 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 system board 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 Unified Extensible Firmware Interface

The x3850 X5 uses 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 compatible with an earlier version 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 web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5083207>

3.13.5 Integrated Trusted Platform Module

The Trusted Platform Module (TPM) in the x3850 X5 is compliant with TPM 1.2. This integrated security chip does 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 Remind button of the light path diagnostics panel for the TPM Physical Presence function.

For details about this technology, see the TCG TPM Main Specification at the following web page:

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, see the following web page:

<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. However, it 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 3-31 shows the light path diagnostics panel on the x3850 X5.



Figure 3-31 Light path diagnostics panel on the x3850 X5

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

<http://ibm.com/support/entry/portal/docdisplay?lnocid=MIGR-5084848>

3.14 Power supplies and fans of the x3850 X5 and MAX5

The power and cooling features of the x3850 X5 server and the MAX5 memory expansion unit are now described.

3.14.1 x3850 X5 power supplies and fans

The x3850 X5 includes one or two dual-rated power supplies as standard. (This feature is model-dependent). See 3.3, “Models” on page 53 for more information.

The power supplies are rated at the following values:

- ▶ 1975 watts at 220 V ac input
- ▶ 875 watts at 110 V ac input

The two power supplies are hot-swappable and redundant at 220 V ac only.

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 (Figure 3-32). Fans are redundant: there are 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

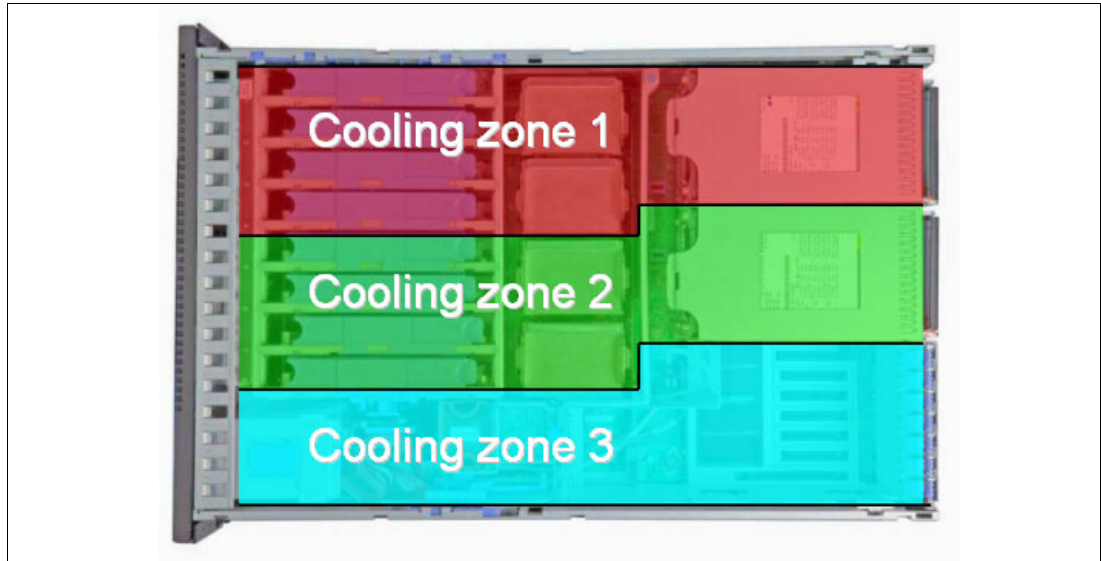


Figure 3-32 Cooling zones in the x3850 X5

Six strategically located hot-swap and redundant fans, which are combined with efficient airflow paths, provide highly effective system cooling for the x3850 X5 systems. This technology is known as *IBM Calibrated Vectors 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 correct 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 two-node complex. In this case, the QPI cables directly link the two servers.

3.14.2 MAX5 power supplies and fans

The MAX5 power subsystem consists of two hot-pluggable 675 W power supplies. The power subsystem is designed for N+N (fully redundant) operation and hot-swap replacement. MAX5 units have both power supplies installed as standard.

MAX5 has five redundant hot-swap fans, which are a all-in-one cooling zone. The IMM of the attached host controls the MAX5 fan speed, which is based on altitude and ambient temperature. In addition, a fan that is located inside each power supply cools the power modules.

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 preinstalled with VMware ESXi, as shown in Figure 3-33. However, all models of x3850 X5 support several USB keys as options, as listed in Table 3-38.

For more information about the USB keys, and to download the IBM customized version of VMware ESXi, visit the following web page:

<http://www.ibm.com/systems/x/os/vmware/esxi>

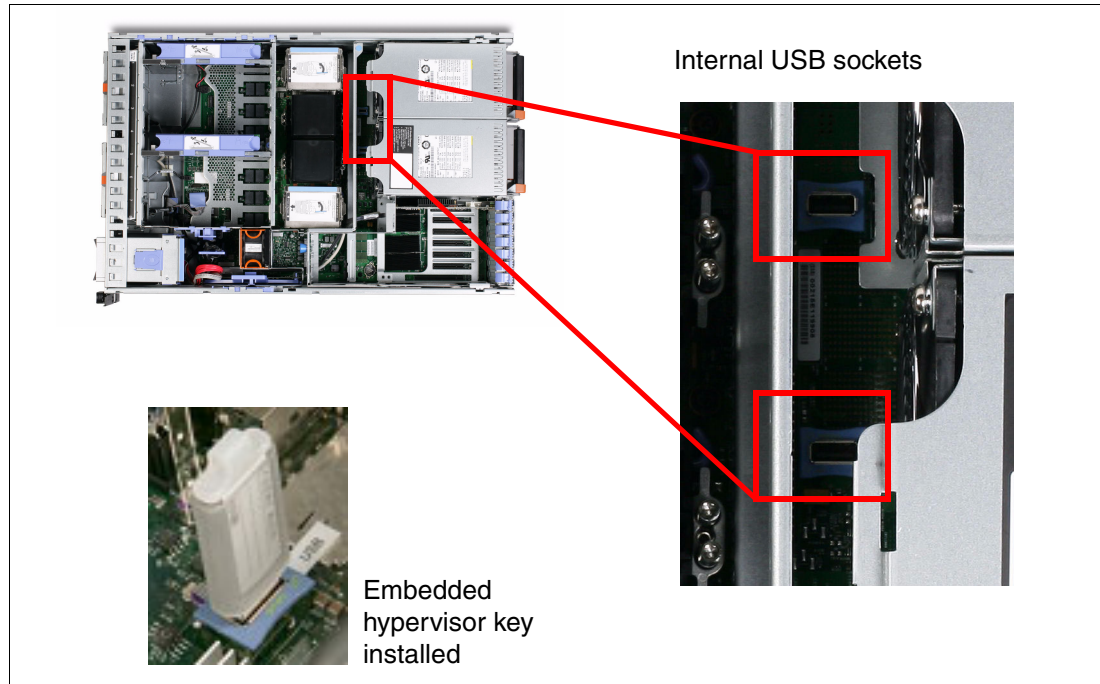


Figure 3-33 Location of internal USB ports for embedded hypervisor on the x3850 X5 and x3950 X5

Table 3-38 shows the VMware ESXi memory keys.

Table 3-38 VMware ESXi memory keys

Part number	Feature code	Description
41Y8298	A2G0	IBM Blank USB Memory Key for VMware ESXi Downloads
41Y8296	A1NP	IBM USB Memory Key for VMware ESXi 4.1 Update 1
41Y8300	A2VC	IBM USB Memory Key for VMware ESXi 5.0
41Y8307	A383	IBM USB Memory Key for VMware ESXi 5.0 Update 1
41Y8311	A2R3	IBM USB Memory Key for VMware ESXi 5.1

3.16 Operating system support

The following operating systems are supported by the x3850 X5 and x3950 X5:

- ▶ Microsoft Windows HPC Server 2008
- ▶ Microsoft Windows Server 2008 HPC Edition

- ▶ Microsoft Windows Server 2008 R2
- ▶ 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
- ▶ Microsoft Windows Server 2012
- ▶ Microsoft Windows Small Business Server 2008 Premium Edition
- ▶ Microsoft Windows Small Business Server 2008 Standard Edition
- ▶ Red Hat Enterprise Linux 5 Server with Xen x64 Edition
- ▶ Red Hat Enterprise Linux 5 Server x64 Edition
- ▶ Red Hat Enterprise Linux 6 Server x64 Edition
- ▶ Solaris 10 Operating System
- ▶ 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 4.1
- ▶ VMware ESXi 4.1
- ▶ VMware vSphere 5.0 (ESXi)
- ▶ VMware vSphere 5.1 (ESXi)

Check the ServerProven Operating System support page for the most up-to-date list:

<http://www.ibm.com/servers/eserver/serverproven/compat/us/nos/matrix.shtml>

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 you use 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 in.) for the front bezel (approximate)
- ▶ 2.5 cm (1 in.) for air flow (approximate)



IBM System x3690 X5

The x3690 X5 servers are powerful two-socket rack-mount servers with up to ten-core Intel Xeon processors. You can combine most models of the x3690 X5 servers with IBM MAX5 memory expansion for up to 2 TB of memory in a two-socket system.

The x3690 X5 server introduces a new design in this next generation of Enterprise X-Architecture (EXA) servers. Until now, EXA capabilities were reserved for the four-socket scalable systems. This server delivers innovation with enhanced reliability and availability features to enable optimal performance for databases, enterprise applications, and virtualized environments.

The following topics are described:

- ▶ 4.1, “Product features” on page 116
- ▶ 4.2, “Target workloads” on page 121
- ▶ 4.3, “Models” on page 122
- ▶ 4.4, “System architecture” on page 126
- ▶ 4.5, “MAX5” on page 126
- ▶ 4.6, “Scalability” on page 129
- ▶ 4.7, “Processor options” on page 130
- ▶ 4.8, “Memory” on page 131
- ▶ 4.9, “Storage” on page 146
- ▶ 4.10, “PCIe slots” on page 167
- ▶ 4.11, “Standard features” on page 173
- ▶ 4.12, “Power supplies” on page 177
- ▶ 4.13, “Integrated virtualization” on page 178
- ▶ 4.14, “Supported operating systems” on page 179
- ▶ 4.15, “Rack mounting” on page 180

4.1 Product features

The x3690 X5 is a 2U, two-socket, scalable system that offers over twice the memory capacity of current two-socket servers. This system has the following features:

- ▶ Support for two Intel Xeon E7 processors that offer up to ten cores per processor.
- ▶ Support for dual inline memory modules (DIMMs) with x4 dynamic random access memory (DRAM) modules. This support enables double device data correction (DDDC) and support for 32 GB DIMMs for a maximum of 2 TB per x3690 X5 with MAX5 V2.
- ▶ Memory implementation that uses high-speed PC3-10600 and PC3-8500 DDR3 memory technology at up to 1066 MHz bus speed.
- ▶ Up to 32 DIMMs in the base system (16 on the system board and 16 on an optional memory mezzanine). Plus, add an extra 32 DIMMs with an optional 1U MAX5 memory expansion unit, for a total of 64 DIMM sockets.
- ▶ Intel QuickPath Interconnect (QPI) technology for processor-to-processor connectivity and Intel 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, standard in all models.
- ▶ Emulex 10 Gb dual-port Ethernet adapter, standard on certain models, and optional on all other models.
- ▶ Memory ProteXion with Chipkill, memory mirroring, memory sparing, Intel SMI Lane Failover, SMI Packet Retry, and SMI Clock Failover.
- ▶ Up to 16 hot-swap 2.5-inch serial-attached SCSI (SAS) hard disk drives (HDDs) and up to 16 TB of maximum internal storage. Or, 16 hot-swap 2.5-inch solid-state drive (SSD) HDDs and up to 800 GB of storage, or up to 24 hot swap 1.8-inch SSD and four 2.5" drives for a maximum of 16 TB. The system includes (as standard) one 2.5-inch HDD backplane that can hold four drives, with an optional second and third backplane for an extra 12 drives. Adding more than two backplanes requires an extra SAS controller card.
- ▶ SAS-based internal storage with RAID-0, RAID-1, or RAID-10 to maximize throughput and ease of installation; other RAID levels are supported by optional RAID adapters.
- ▶ New eXFlash high-I/O operations per second (IOPS) solid-state storage technology for larger, faster databases. See 2.8, "IBM eXFlash" on page 38 for more information.
- ▶ A maximum of five PCI Express (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 by using riser card option 60Y0329.
 - One x16, two x8, and one x4 PCIe slot, which uses riser card option 60Y0331 for a 3/4 length adapter, or option 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 shows the x3690 X5 server with 16 hot-swap 2.5-inch SAS disk drives installed.



Figure 4-1 IBM System x3690 X5

The x3690 server has the following 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. Hot-swappable parts in this server include the HDDs, fans, and power supplies. Other parts require that the server is powered off before you remove that component.

4.1.1 System components

Figure 4-2 shows the components on the front of the system.

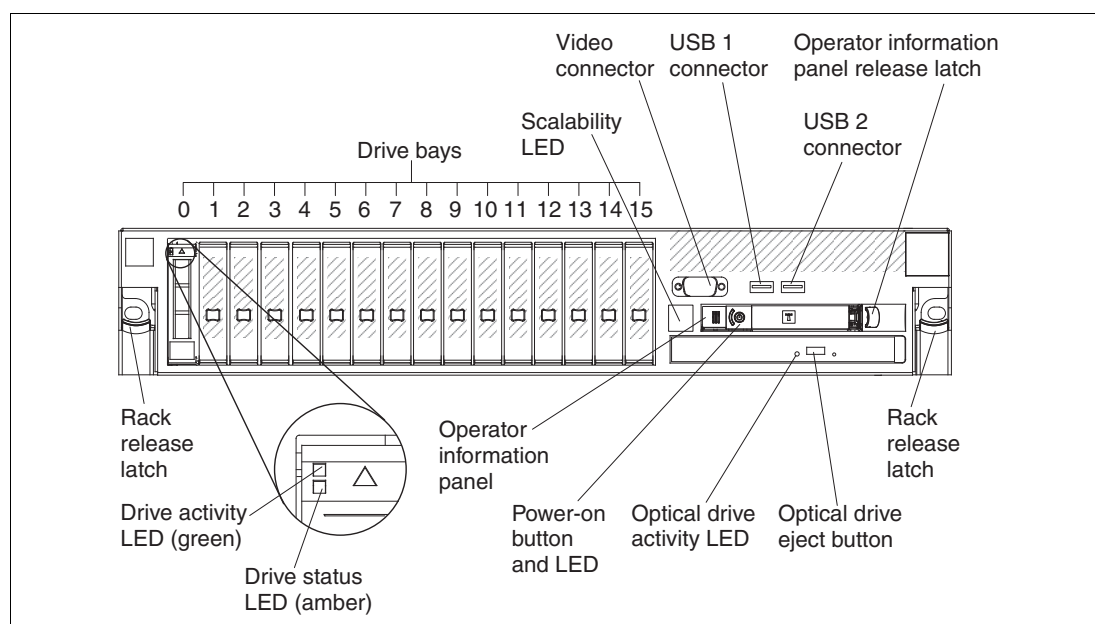


Figure 4-2 Front view of x3690 X5

Figure 4-3 shows the rear of the system.

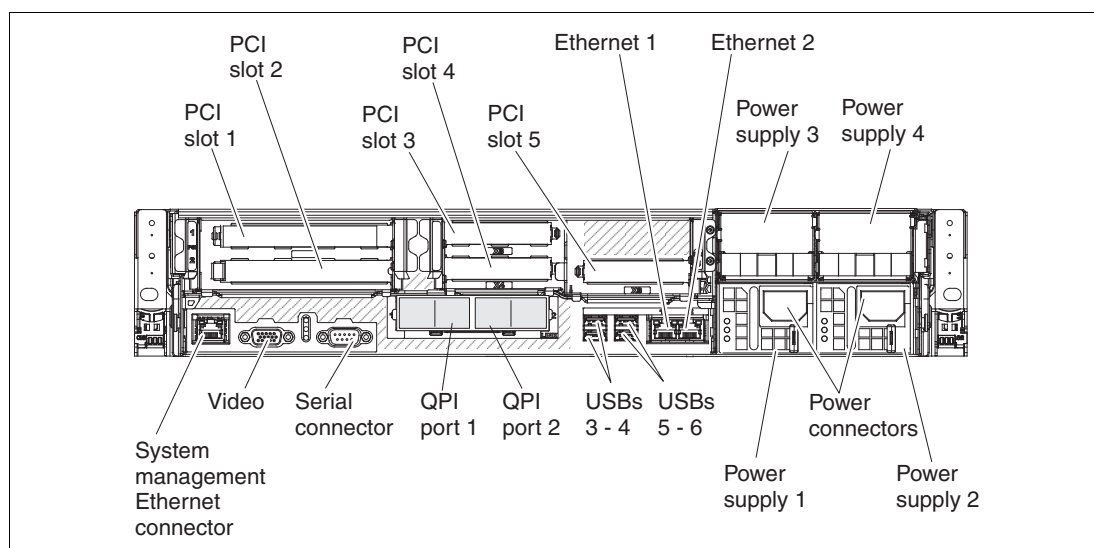


Figure 4-3 Rear view of x3690 X5

Figure 4-4 shows the system with the top cover removed.

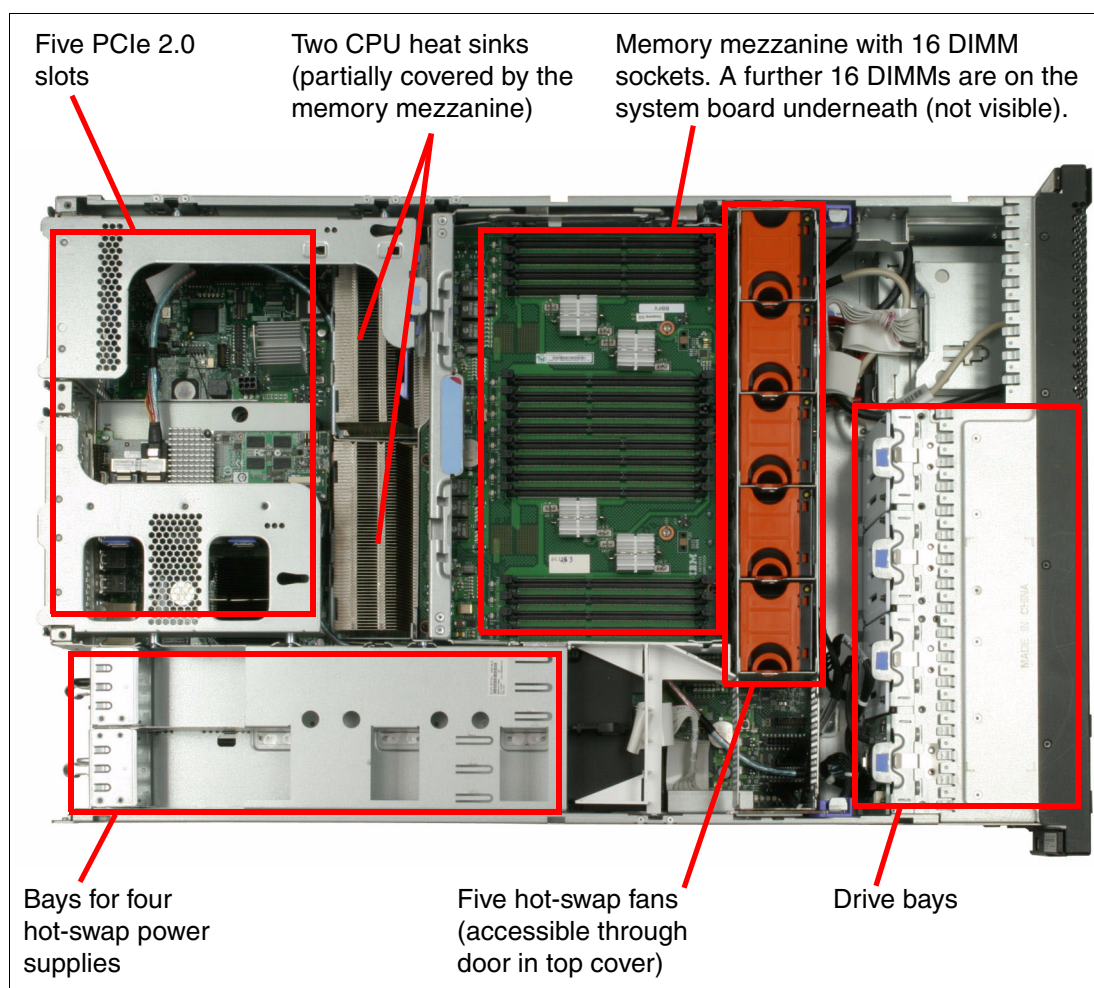


Figure 4-4 The x3690 X5 internals

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, 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 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 the expansion module to be powered off.

The second generation of the MAX5, the MAX5 V2, features newer versions of scalable memory buffers, which enable support for both 1.35 V DIMMs and 32 GB DIMMs.

Compatibility is summarized in Table 4-1.

Table 4-1 MAX5 compatibility

MAX5 model	x3690 X5 with E7 processors (machine type 7147)
IBM MAX5, 59Y6265	Supported
IBM MAX5 V2, 88Y6529	Supported

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 hardware 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, each of which operates at up to 6.4 GT/s, depending on the processors installed.

- ▶ Memory DIMMs:

- Minimum: 2 DIMMs, 4 GB
- Maximum: 32 DIMMs:
 - MAX5: up to 512 GB of memory with 16 GB DIMMs
 - MAX5 V2: up to 1 TB of memory with 32 GB DIMMs
- Type: PC3-10600, 1067 MHz, error correction code (ECC), DDR3 registered synchronous dynamic random access memory (SDRAM) DIMMs
- Sizes:
 - MAX5 supports 2 GB, 4 GB, 8 GB, and 16 GB DIMMs
 - MAX5 V2 supports 2 GB, 4 GB, 8 GB, 16 GB, and 32 GB DIMMs
- Low voltage (1.35 V) DIMM support with MAX5 V2

All DIMM sockets in the MAX5 are accessible regardless of the number of processors that are 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-watts (100 - 240 Volts AC 50-60 Hz auto-sensing)
- Two power supplies standard, full 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 inches)
- Depth: 724 mm (28.5 inches)
- Height: 44 mm (1.73 inches or 1U rack unit)
- Basic configuration: 12.8 kg (28.2 lb)
- Maximum configuration: 15.4 kg (33.9 lb)

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 you connect the MAX5 to an x3690 X5, the QPI ports are used. The EXA ports are unused.

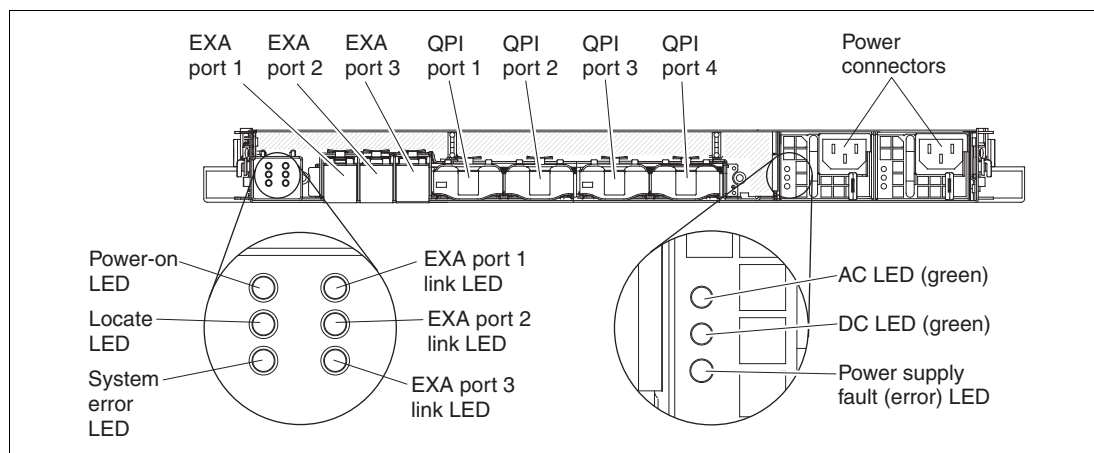


Figure 4-6 MAX5 connectors and LEDs

Figure 4-7 shows the internals of the MAX5, including the IBM EXA chip that acts as the interface to the QPI links from the x3690 X5.

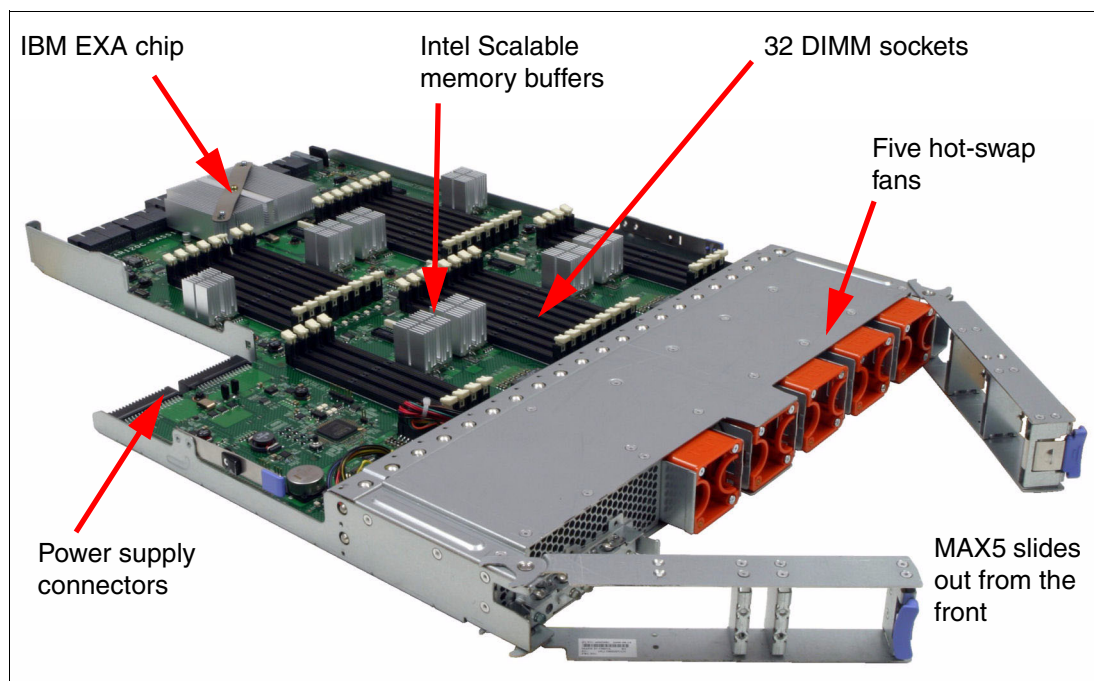


Figure 4-7 MAX5 memory expansion unit internals

For an in-depth look at the MAX5 offering, see 4.5, “MAX5” on page 126.

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 two-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 four-socket servers. This support is possible because they have more than double the memory of current two-socket, Intel Xeon E5 processor-based servers. The result can lead to client savings on hardware and also on software licensing. There are pre-configured “workload optimized” models for virtualization, see Table 4-3 on page 124.
- ▶ Database

The larger memory capacity of the x3690 X5 also offers leadership database performance. The x3690 X5 features the IBM eXFlash internal storage that uses SSDs to maximize the number of IOPS. Workload optimized models pre-configured for database serving and models tuned for SAP High-Performance Analytic Appliance (HANA) are available. See Table 4-3 on page 124 for more information about these models.

4.3 Models

In addition to the details in the tables in this chapter, each standard x3690 X5 model has the following specifications:

- ▶ The servers have 16 DIMM sockets on the system board. 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 167 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.

4.3.1 Base x3690 X5 models with Intel Xeon E7 series processors

Table 4-2 provides the standard models of the x3690 X5 that use the Intel Xeon E7 processor. The MAX5 memory expansion unit is optional on certain models, as indicated, but not supported on others.

Table 4-2 Base x3690 X5 models with Intel Xeon E7 series processors

Model	Processor (model, cores, core speed, L3 cache, memory speed, TDP power rating) (two max)	MAX5	RAM	Memory mezzanine	ServeRAID M1015 std	Disk bays	Disk drives	10 Gb Ethernet standard ^a	Optical drive	Power supplies (std/max)
7147-A1x	1x Xeon E7-2803 6C 1.73 GHz 18 MB 800 MHz 105 W	NS ^b	2x 4 GB	Opt	Opt	4x 2.5 in. 16 max	Opt	Opt	Opt	1 / 4
7147-A2x	1x Xeon E7-2820 8C 2.00 GHz 18 MB 978 MHz 105 W	NS ^b	2x 4 GB	Opt	Std	4x 2.5 in. 16 max	Opt	Opt	Opt	2 / 4

Model	Processor (model, cores, core speed, L3 cache, memory speed, TDP power rating) (two max)	MAX5	RAM	Memory mezzanine	ServeRAID M1015 std	Disk bays	Disk drives	10 Gb Ethernet standard ^a	Optical drive	Power supplies (std/max)
7147-A3x	1x Xeon E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	2x 4 GB	Opt	Std	4x 2.5 in. 16 max	Opt	Opt	Opt	2 / 4
7147-A5x	1x Xeon E7-2850 10C 2.00 GHz 24 MB 1066 MHz 130 W	Opt	2x 4 GB	Opt	Std	4x 2.5 in. 16 max	Opt	Opt	Opt	2 / 4
7147-A6x	1x Xeon E7-2860 10C 2.26 GHz 24 MB 1066 MHz 130 W	Opt	2x 4 GB	Opt	Std	4x 2.5 in. 16 max	Opt	Opt	Opt	2 / 4
7147-A7x	1x Xeon E7-2870 10C 2.40 GHz 30 MB 1066 MHz 130 W	Opt	2x 4 GB	Opt	Std	4x 2.5 in. 16 max	Opt	Opt	Opt	2 / 4
7147-C1x	1x Xeon E7-8837 8C 2.67 GHz 24 MB 1066 MHz 130 W	Opt	2x 4 GB	Opt	Std	4x 2.5 in. 16 max	Opt	Opt	Opt	2 / 4

a. Emulex 10 Gb Ethernet Adapter.

b. NS = not supported. The MAX5 is not supported on systems with E7-2803 or E7-2820 processors.

4.3.2 Workload-optimized x3690 X5 models with Xeon E7 series processors

Table 4-3 lists the workload-optimized models that are based on Intel Xeon E7 series processors. All of these models have four power supplies standard.

Table 4-3 Workload-optimized x3690 X5 models with Xeon E7 series processors

Model	Processor (model, cores, core speed, L3 cache, memory speed) (two max)	MAX5	Memory	Disk controllers	Disk bays	Disk drives Optical drive	Network
Database workload-optimized models							
7147-D3x ^a	2x E7-2860 10C 2.26 GHz 24 MB 1066 MHz	Opt	16x 4 GB Optional mezz.	2x M5015 (+Perf keys)	16x 1.8 in. / 24	16x 200 GB SSD	2x 1 Gb 2x 10 Gb
7147-D4x ^a	2x E7-2860 10C 2.26 GHz 24 MB 1066 MHz	Opt	16x 4 GB Optional mezz.	2x 6 Gb SSD HBA	16x 1.8 in. / 24	16x 200 GB SSD	2x 1 Gb 2x 10 Gb
SAP HANA workload-optimized models							
7147-HAx ^b	2x E7-2870 10C 2.40 GHz 30 MB 1066 MHz	NS ^c	8x 16 GB Standard mezz.	2x M5015 (+ Perf keys)	16x 1.8 in. / 24	10x 200 GB SSD Multiburner	6x 1 Gb 4x 10 Gb
7147-HBx ^b	2x E7-2870 10C 2.40 GHz 30 MB 1066 MHz	NS ^c	16x 16 GB Standard mezz.	2x M5015 (+ Perf keys)	16x 1.8 in. / 24	10x 200 GB SSD Multiburner	6x 1 Gb 4x 10 Gb
Virtualization workload-optimized models							
7147-F1x (VMware)	2x E7-2860 10C 2.26 GHz 24 MB 1066 MHz	Std	Server: 32x 4 GB MAX5: 32x 4 GB	1x M1015	4x 2.5 in. / 16	None	2x 1 Gb 2x 10 Gb
7147-F2x (RHEL)	2x E7-2860 10C 2.26 GHz 24 MB 1066 MHz	Std	Server: 32x 4 GB MAX5: 32x 4 GB	1x M1015	4x 2.5 in. / 16	None	2x 1 Gb 2x 10 Gb

a. The D3x and D4x models include one Emulex 10GbE Integrated Virtual Fabric Adapter II (no transceivers included).

b. The HAx and HBx models include two Emulex 10GbE Integrated Virtual Fabric Adapter II (each with two IBM 10GbE SW SFP+ Optical Transceivers) and one Intel Ethernet Quad Port Server Adapter I340-T4.

c. NS = Not supported. MAX5 is not currently certified for use with SAP HANA and is therefore not supported.

The following list provides information about these models:

- Models 7147-D3x, D4x: These models are designed for database applications and use SSDs for the best I/O performance.

Backplane connections for 16 1.8-inch SSDs are standard as are 16 200 GB high-performance SSDs. Model D4x includes two SSD host bus adapters. Model D3x includes two ServerRAID M5015 RAID controllers each with the ServerRAID M5000 Series Performance Accelerator Key.

- Models 7147-HAx, HBx: These models are optimized to run the SAP HANA solution.

HANA is an integrated, ready-to-run, hardware-software offering, featuring the new SAP In-Memory Computing Engine. These models include preinstalled software that consists of SUSE Linux Enterprise Server (SLES) for SAP, IBM General Parallel File System (GPFS), and the SAP HANA software stack. The models include two processors, 128 or 256 GB of memory, and a choice of either all eXFlash SSDs or a combination of solid state and spinning disk. They are designed for use in small to mid-sized SAP HANA configurations. H models also include a SATA Multiburner optical drive.

SAP HANA not supported by MAX5: MAX5 is not currently certified for use with SAP HANA and is therefore not supported.

- Model 7147-F1x: This model is designed for virtualization applications and includes VMware ESXi 4.1 Update 1 on an integrated bootable USB memory key.

The models come standard with the MAX5 V2 memory expansion unit and 256 GB of memory that is implemented by using cost-effective 4 GB memory DIMMs (128 GB in the server and 128 GB in the MAX5).

- Model 7147-F2x: This model is designed for Open Virtualization and includes Red Hat Enterprise Linux with the Red Hat Enterprise Virtualization Hypervisor (kernel-based virtual machine (KVM)). The software is not preinstalled.

The model comes standard with the MAX5 memory expansion unit and 256 GB of memory that is implemented by using cost-effective 4 GB memory DIMMs (128 GB in the server and 128 GB in the MAX5).

4.4 System architecture

Figure 4-8 shows a block diagram of the x3690 X5.

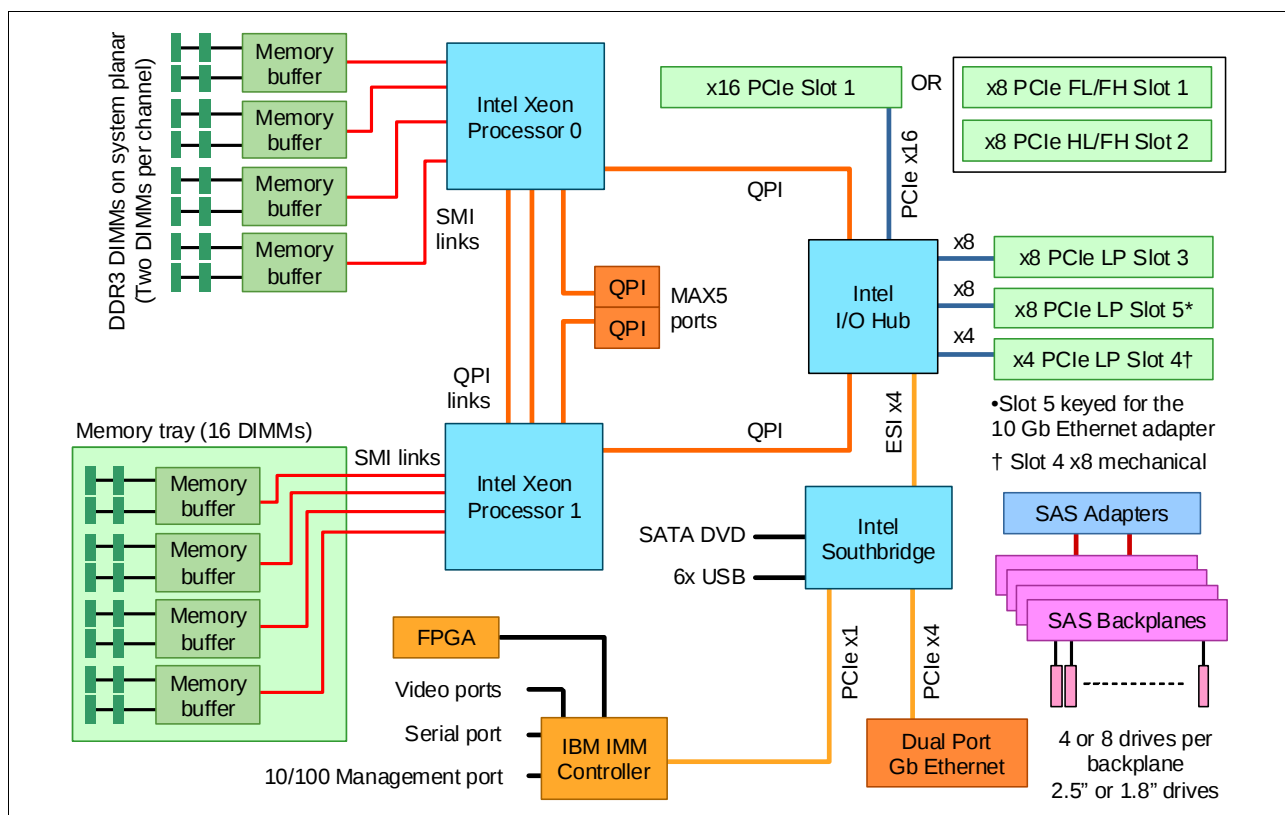


Figure 4-8 x3690 X5 block diagram

4.5 MAX5

As introduced in 4.1.2, “IBM MAX5 memory expansion unit” on page 119, 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 122. The MAX5 can also be ordered separately, as listed in Table 4-4.

There are two MAX5 options available.

- ▶ IBM MAX5 for System x, part number 59Y6265 (also known as MAX5 V1)
- ▶ IBM MAX5 V2 for System x, part number 88Y6529

When you order a MAX5, remember to order the appropriate cable kits as well. The MAX5 includes both power supplies as standard; therefore, no additional power supplies are needed.

Table 4-4 Ordering information for the IBM MAX5 for System x

Part number	Feature code	Description
59Y6265	4199	IBM MAX5 for System x
88Y6529	A19H	IBM MAX5 V2 for System x

Part number	Feature code	Description
60Y0332	4782	IBM High Efficiency 675W Power Supply (MAX5 V1 only, 59Y6265)
59Y6269	7481	IBM MAX5 to x3690 X5 Cable Kit (two cables)

There are two MAX5 models. Compatibility is shown in Table 4-5.

Table 4-5 MAX5 compatibility

MAX5 model	x3690 X5 with Intel Xeon E7 (machine type 7147)
MAX5, 59Y6265	Supported
MAX5 V2, 88Y6529	Supported

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

The eX5 chip connects directly through QPI links to both CPUs in the x3690 X5. And, it maintains a copy of the last-level cache of each CPU. This directory allows the eX5 chip to respond to cache update requests more quickly than the native processor implementation, improving performance. For more information about eX5 technology, see 2.1, “eX5 chip set” on page 10.

Figure 4-9 shows a block diagram of the MAX5.

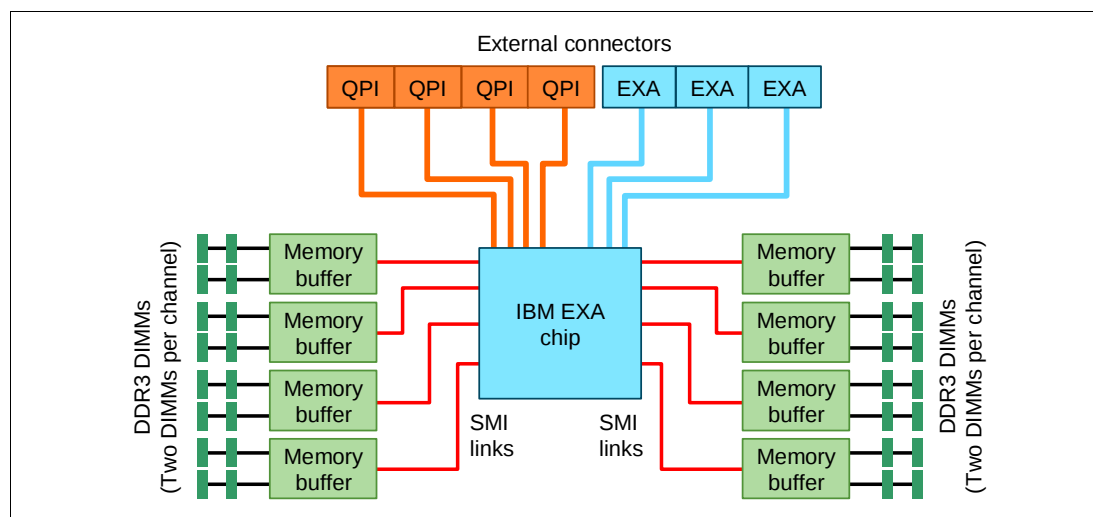


Figure 4-9 MAX5 block diagram

The MAX5 is connected to the x3690 X5 by using two cables. These cables connect 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 unused.

Figure 4-10 shows architecturally how a single-node x3690 X5 is connected to a MAX5.

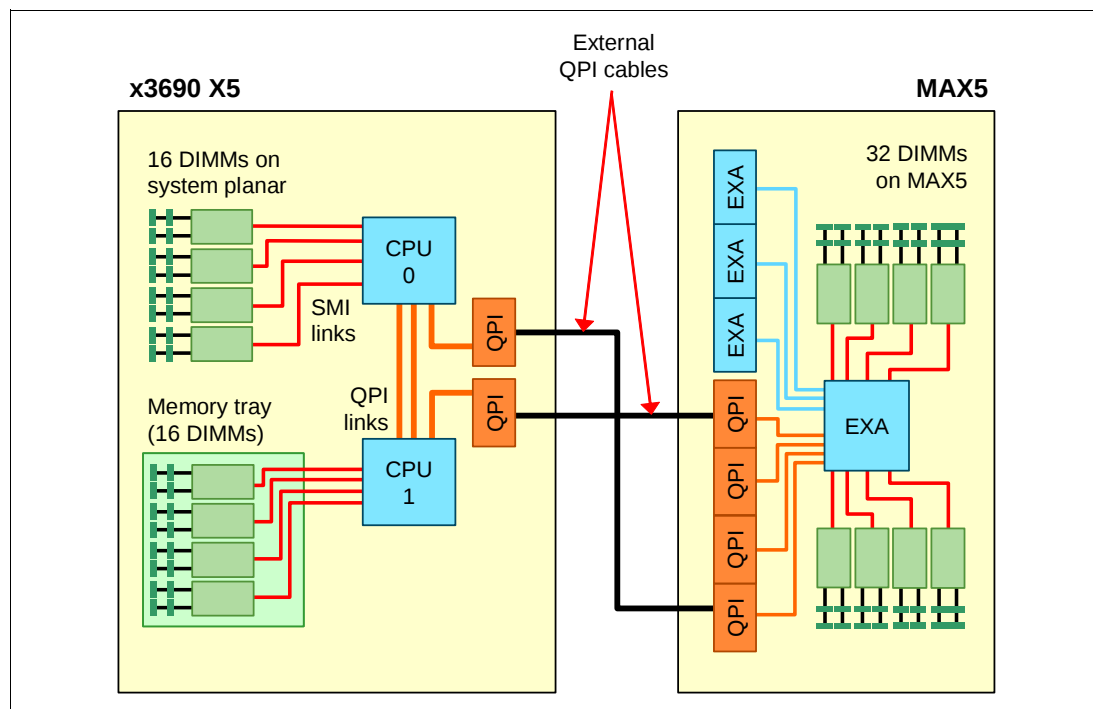


Figure 4-10 Connectivity of the x3690 X5 with a MAX5 memory expansion unit

As shown in Figure 4-10, the x3690 X5 attaches to the MAX5 by 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 that is stored in the cache of another CPU, the MAX5 not only has that same data stored in its own cache. It is also 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 2.2.6, “QuickPath Interconnect” on page 13.

Xeon E7-2803 processors: The Xeon E7-2803 processors do not support the use of the MAX5.

Connectivity of the MAX5 to the x3690 X5 is described in 4.6, “Scalability” on page 129.

For memory configuration information, see 4.8.4, “MAX5 memory population order” on page 139.

MAX5 V1 includes one power supply. The second power supply is optional (part 60Y0332) as listed in Table 4-4 on page 126 and provides redundancy. MAX5 V2 includes two power supplies so no additional power supplies are needed or available. MAX5 power supplies are hot-pluggable 675 W units. With two installed, the power subsystem is designed for N+N (fully redundant) operation and hot-swap replacement.

MAX5 has five redundant hot-swap fans, which are a all-in-one cooling zone. The integrated management module (IMM) of the attached host controls the MAX5 fan speed, which is based on altitude and ambient temperature. In addition, a fan that is located inside each power supply cools the power modules.

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.

4.6 Scalability

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 x3690 X5 does not support two-node configurations (with or without MAX5).

The MAX5 memory expansion unit allows the x3690 X5 to scale to an extra 32 DDR3 DIMM sockets.

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 4-6.

Table 4-6 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)

MAX5 processor support: The MAX5 is supported by either one or two processors that are installed in the x3690 X5. However, installing two processors and memory in every DIMM socket maximizes performance.

Figure 4-11 shows the connectivity between the IBM MAX5 and a single-node x3690 X5.

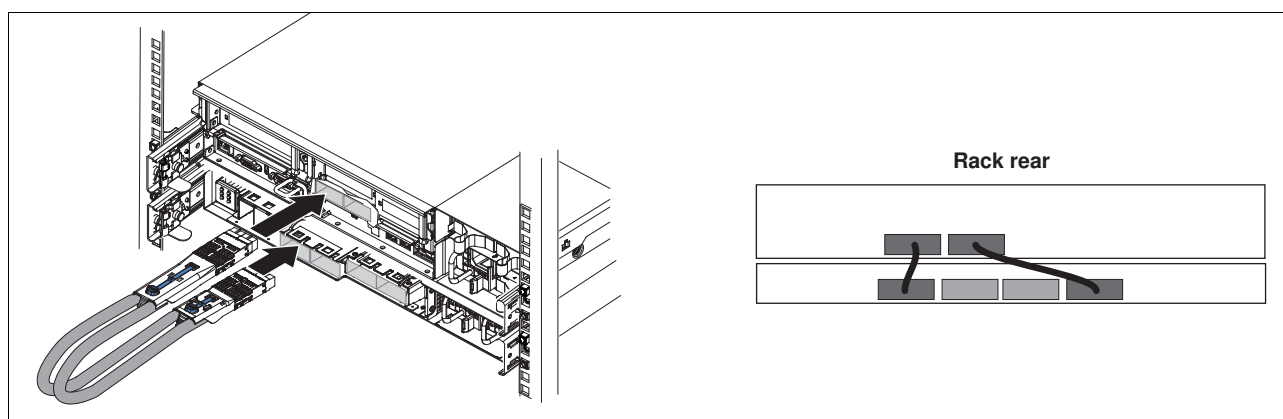


Figure 4-11 Connecting the MAX5 to a single-node x3690 X5

4.7 Processor options

Several Intel Xeon E7 processor options are available for the x3690 X5, machine type 7147, as listed in Table 4-7.

Table 4-7 x3690 X5 Intel Xeon E7 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										
88Y5663	A15U	E7-2870 10C	2.40 GHz	30 MB	6.4 GT/s	1066 MHz	130 W	Yes	Yes	Yes
88Y5664	A15V	E7-2860 10C	2.26 GHz	24 MB	6.4 GT/s	1066 MHz	130 W	Yes	Yes	Yes
88Y5720	A15Z	E7-2850 10C	2.00 GHz	24 MB	6.4 GT/s	1066 MHz	130 W	Yes	Yes	Yes
Standard processors										
88Y5665	A15W	E7-2830 8C	2.13 GHz	24 MB	6.4 GT/s	1066 MHz	105 W	Yes	Yes	Yes
88Y5666	A15X	E7-2820 8C	2.00 GHz	18 MB	5.86 GT/s	977 MHz	105 W	Yes	Yes	Yes
Basic processors										
88Y5662	A191	E7-4807 6C	1.86 GHz	18 MB	4.8 GT/s	800 MHz	95 W	Yes	No	Yes
88Y5667	A15Y	E7-2803 6C	1.73 GHz	18 MB	4.8 GT/s	800 MHz	105 W	Yes	No	No
High performance, low-power processor (L suffix)										
88Y5654	A15S	E7-8867L 10C	2.13 GHz	30 MB	6.4 GT/s	1066 MHz	105 W	Yes	Yes	Yes
Frequency optimized processor										
88Y5657	A15T	E7-8837 8C	2.67 GHz	24 MB	6.4 GT/s	1066 MHz	130 W	No	Yes	Yes

a. Thermal design power

b. Intel Hyper-Threading Technology

c. Intel Turbo Boost Technology

Xeon E7-2803 restriction: The Xeon E7-2803 does not support the use of the MAX5 memory expansion unit.

For more information about the processors that are used in the x3690 X5, see 2.2, “Intel Xeon processors” on page 10.

Xeon E7 processors (with two exceptions) support Intel Turbo Boost Technology, as indicated in Table 4-7 on page 130. When a CPU operates beneath its thermal and electrical limits, Turbo Boost dynamically increases the clock frequency of the processor by 133 MHz on short and regular intervals. This process is done until an upper limit is reached. See 2.2.5, “Turbo Boost Technology” on page 12 for more information.

Except for the E7-8837, all CPUs that are listed support Intel Hyper-Threading (HT) Technology. HT Technology is an Intel technology that is used to improve the parallelization of workloads. When HT is enabled in the Unified Extensible Firmware Interface (UEFI), the operating system treats each processor core as two independently addressable processing units. For more information, see 2.2.4, “Hyper-Threading Technology” on page 12.

All CPU options include a heat sink.

The x3690 X5 models include one CPU as standard. All five PCIe slots are usable, even with only one processor that is 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. For VMware, equal amounts of memory must be configured for each processor.

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 E7-8837 processor for CPU frequency-dependent workloads because it has the highest core frequency of the available processor models.
- ▶ The MAX5 is supported by either one or two processors that are installed in the x3690 X5. However, the recommendation is to have two processors that are installed and memory that is 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 the use of two processors with fewer cores or a lower-core frequency.

4.8 Memory

The x3690 X5 offers up to 32 DIMM sockets that are internal to the server chassis. Plus, there are an extra 32 DIMM sockets in the MAX5 memory expansion unit.

The following topics are covered:

- ▶ 4.8.1, “x3690 X5 memory options” on page 133
- ▶ 4.8.2, “MAX5 memory options” on page 134
- ▶ 4.8.3, “x3690 X5 memory population order” on page 136
- ▶ 4.8.4, “MAX5 memory population order” on page 139

- ▶ 4.8.5, “Memory balance” on page 140
- ▶ 4.8.6, “Mixing DIMMs and the performance effect” on page 141
- ▶ 4.8.7, “Memory mirroring” on page 142
- ▶ 4.8.8, “Memory sparing” on page 144
- ▶ 4.8.9, “Effect on performance when you use mirroring or sparing” on page 145

Implement the memory DIMMs that are internal to the x3690 X5 chassis:

- ▶ 16 DIMM sockets on the system board
- ▶ 16 DIMM sockets in an optional memory mezzanine

Tip: The memory mezzanine is referred to in the announcement letter as the *memory expansion card*. In the *Installation and User's Guide - IBM System x3690 X5*, it is referred to as the *memory tray*.

The memory mezzanine is an optional component and can be ordered as listed in Table 4-8.

Table 4-8 x3690 X5 memory mezzanine option part number

Option	Feature code	Description
81Y8926	A15H	IBM x3690 X5 16-DIMM Internal Memory Expansion for E7 processor-based servers, machine type 7147

Figure 4-12 shows the memory mezzanine and DIMMs in the x3690 X5.

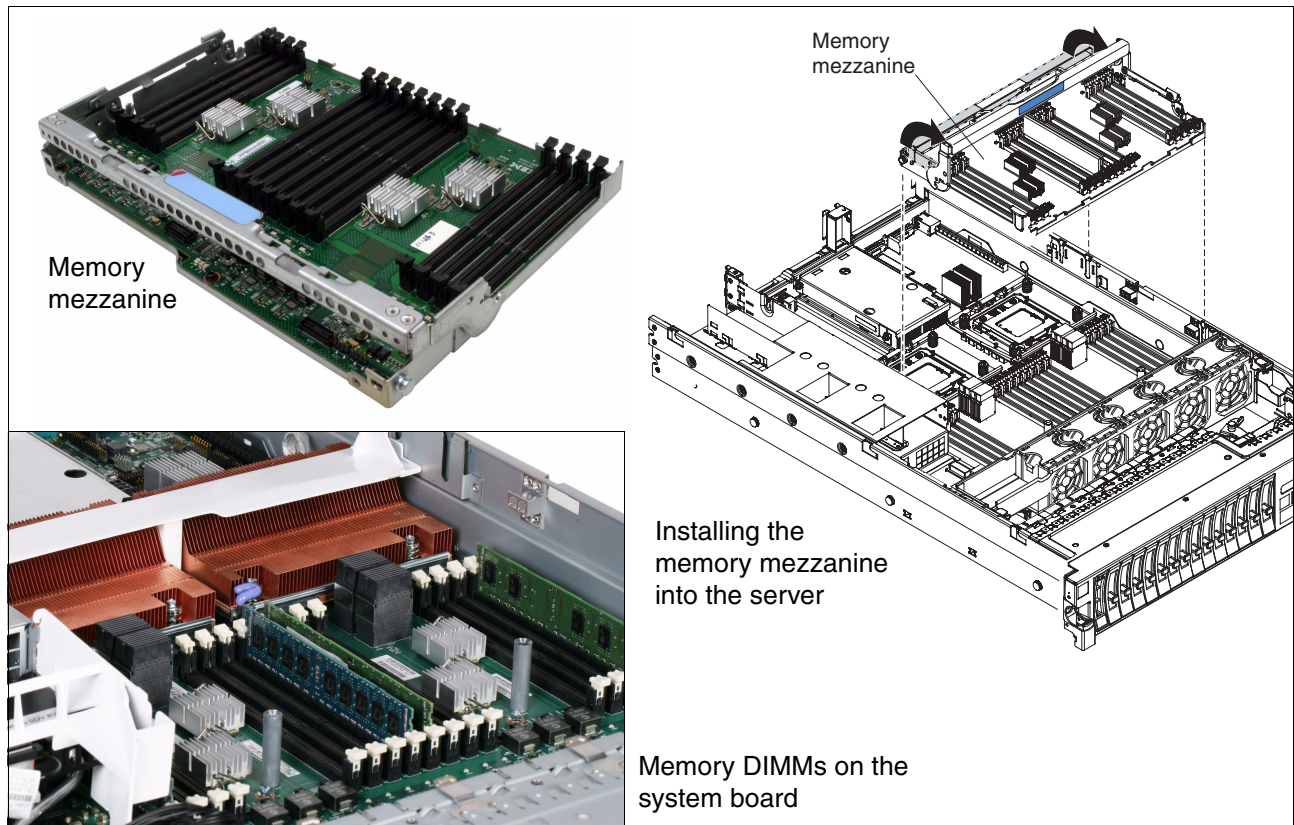


Figure 4-12 Location of the memory DIMMs

With these Intel 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 board.
- 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 (DPCs).

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. For more information about how this value is calculated, see 2.3.1, “Memory speed” on page 17.

The memory mezzanine is included in servers with Intel Xeon E7 processors, machine type 7147. This server supports DIMMS with x4 DRAM modules. Therefore, the server supports up to 1 TB of internal memory and uses low voltage (1.35 V) DIMMs to reduce energy consumption.

4.8.1 x3690 X5 memory options

The memory DIMM options for the x3690 X5 depend on which machine type you are configuring.

Table 4-9 shows the available memory options that are supported in the x3690 X5 server with Intel Xeon E7 processors, machine type 7147.

Table 4-9 Supported DIMMs for x3690 X5, machine type 7147 (E7 processors)

Part number	Feature code	Memory	Memory speed ^a	Ranks
44T1592 ^b	1712	2 GB (1x 2 GB), 1Rx8, PC3-10600 DDR3-1333	1333 MHz	Single x8
44T1481 ^b	3964	2 GB (1x 2 GB), 2Rx8, PC3-10600 DDR3-1333	1333 MHz	Dual x8
49Y1433 ^b	8934	2 GB (1x 2 GB), 2Rx8, PC3-10600 DDR3-1333	1333 MHz	Dual x8
44T1599 ^b	1713	4 GB (1x 4 GB), 2Rx8, PC3-10600 DDR3-1333	1333 MHz	Dual x8
46C7448 ^b	1701	4 GB (1x 4 GB), 4Rx8, PC3-8500 DDR3-1066	1066 MHz	Quad x8
46C7482 ^b	1706	8 GB (1x 8 GB), 4Rx8, PC3-8500 DDR3-1066	1066 MHz	Quad x8
46C7483 ^b	1707	16 GB (1x 16 GB), 4Rx4, PC3-8500 DDR3-1066	1066 MHz	Quad x4 ^c
49Y1407	8942	4 GB (1x 4 GB), 2Rx8, 1.35 V PC3L-10600 DDR3 1333 ^d	1333 MHz ^e	Dual x8
49Y1399	A14E	8 GB (1x 8 GB), 4Rx8, 1.35 V PC3L-8500 DDR3 1066 ^d	1066 MHz	Quad x8
49Y1400	8939	16 GB (1x 16 GB), 4Rx4, 1.35 V PC3L-8500 DDR3 1066 ^d	1066 MHz	Quad x4 ^c
49Y1563	A1QT	16 GB (1x16 GB), 2Rx4, 1.35 V PC3L-10600 DDR3 1333 ^d	1333 MHz	Dual x4

Part number	Feature code	Memory	Memory speed ^a	Ranks
90Y3101	A1CP	32 GB (1x 32 GB), 4Rx4, 1.35 V PC3L-8500 DDR3 1066 ^d	1066 MHz	Quad x4 ^c

- 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 the two values, that of the processor memory bus speed and of the DIMM memory bus speed.
- This part has been withdrawn from marketing.
- When DIMM slots are populated in hemisphere mode with DIMMs using x4 DRAM modules, DDDC can be enabled in the UEFI.
- When all DIMMs that are populated are low voltage (PC3L), the memory runs at 1.35 V. However, when mixed with 1.5 V DIMMs, they run at 1.5 V.
- Although 1333 MHz memory DIMMs are supported in the x3690 X5, the memory DIMMs run at a maximum speed of 1066 MHz.

Memory and DIMMs:

- ▶ Memory options must be installed in matched pairs. Single options cannot be installed. Therefore, listed options must be ordered in pairs.
- ▶ The maximum memory speed that is supported by the processors is 1066 MHz. DIMMs rated for 1333 MHz can be used, but operate at 1066 MHz.
- ▶ Mixing DIMMs sizes is supported, except for 16 GB and 32 GB DIMMs. However, mixed speed DIMMs operate at the speed of the slowest installed DIMM, and can affect performance.

4.8.2 MAX5 memory options

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. As described in 4.1.2, “IBM MAX5 memory expansion unit” on page 119, there are two MAX5 units available for the x3690 X5: MAX5 and MAX5 V2. The memory DIMM options that are supported in each unit are described.

DIMM memory options for MAX5 V2

Table 4-10 indicates the DIMM options that are supported in the MAX5 V2. When used in the MAX5 V2, the DIMMs have separate feature codes.

Table 4-10 DIMMs supported in MAX5 V2, 88Y6529

Part number	MAX5 V2 feature code	Description
44T1592	2429	2 GB MAX5 1x 2 GB 1Rx8 1.5 V PC3-10600 CL9 ECC DDR3 1333 MHz LP RDIMM
9Y1407	A1MH	4 GB MAX5 (1x 4 GB, 2 Gb, 2Rx8, 1.35 V) PC3L-10600R-999 LP ECC RDIMM
44T1599	2431	4 GB MAX5 1x 4 GB DualRankx8 PC310600 CL9 ECC DDR3 1333 MHz LP RDIMM
46C7482	2432	8 GB MAX5 1x 8 GB QuadRankx8 PC3-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
49Y1399	A1N7	8 GB MAX5 1x 8 GB, 4Rx8, 1.35 V PC3L-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
46C7483	2433	16 GB MAX5 1x 16 GB QuadRankx4 PC3-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
49Y1400	A1N8	16 GB MAX5 1x 16 GB 4Rx4 1.35 V PC3L-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
90Y3206	A1R2	32 GB MAX5 (1x 32 GB 4 Gb, 4Rx4, 1.35 V) PC3L-8500 DDR3-1066 MHz LP RDIMM

When DIMMs with x4 DRAM modules are used, DDDC is automatically enabled. For more information about DDDS, see “Redundant bit steering and double device data correction” on page 25.

Certain DIMMs listed in Table 4-10 on page 134 are low-voltage DIMMs (with “PC3L” in the description). When all DIMMs that are populated are low voltage (PC3L), the memory runs at 1.35 V. However, when mixed with 1.5 V DIMMs, all run at 1.5 V. MAX5 V2 (88Y6529) supports low-voltage DIMM operation. MAX5 (59Y6265) does not support low-voltage operation. If 1.35 V DIMMs are used in MAX5 (59Y6265), they run at 1.5 V.

Although 1333 MHz memory DIMMs are supported in MAX5 and MAX5 V2, the memory DIMMs run at a maximum speed of 1066 MHz. Actual memory speed depends on the processors that are installed in the attached server.

MAX5 V2 memory options: The 16 GB memory option, 46C7483, and the 32 GB memory option, 90Y3206, are supported in the MAX5 V2 only when they are the only type of memory (x4 DRAM) that is used in the MAX5 V2. No other memory options can be used in MAX5 V2 if one or both of these options are installed in the MAX5 V2.

Dual inline memory module options for MAX5

Table 4-11 indicates the DIMM options that are also supported in the MAX5, 59Y6265. When used in the MAX5, the DIMMs have separate feature codes.

Table 4-11 DIMMs supported in MAX5, 59Y6265

Part number	MAX5 feature code	Description
44T1592	2429	2 GB MAX5 1x 2 GB 1Rx8 1.5 V PC3-10600 CL9 ECC DDR3 1333 MHz LP RDIMM
44T1599	2431	4 GB MAX5 1x 4 GB DualRankx8 PC310600 CL9 ECC DDR3 1333 MHz LP RDIMM
9Y1407	A1MH	4 GB MAX5 (1x 4 GB, 2 Gb, 2Rx8, 1.35 V) PC3L-10600R-999 LP ECC RDIMM
46C7482	2432	8 GB MAX5 1x 8 GB QuadRankx8 PC3-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
49Y1399	A1N7	8 GB MAX5 1x 8 GB, 4Rx8, 1.35 V PC3L-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
46C7483	2433	16 GB MAX5 1x 16 GB QuadRankx4 PC3-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
49Y1400	A1N8	16 GB MAX5 1x 16 GB 4Rx4 1.35 V PC3L-8500 CL7 ECC DDR3 1066 MHz LP RDIMM
90Y3206	A1R2	32 GB MAX5 (1x 32 GB 4 Gb, 4Rx4, 1.35 V) PC3L-8500 DDR3-1066 MHz LP RDIMM

Notes:

- ▶ The 16 GB and 32 GB memory options are supported in the MAX5 only when they are the only type of memory (x4 DRAM) that is used in the MAX5. No other memory options can be used in the MAX5 if any of these options are installed in the MAX5.
- ▶ In the x3690 X5 with Intel Xeon E7 processors, DDDC, the Intel implementation of redundant bit steering (RBS) is supported. See “Redundant bit steering and double device data correction” on page 25 for details.
- ▶ The MAX5 memory expansion unit supports RBS, but only with x4 memory and not x8 memory. As shown in Table 4-10 on page 134 and Table 4-11, the 16 GB DIMM, part 46C7483, and the 32 GB DIMM, part 90Y3206, use x4 DRAM technology. RBS is automatically enabled in the MAX5 if all installed DIMMs are x4 DIMMs.

MAX5 memory as seen by the operating system

MAX5 can run in 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*.

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

4.8.3 x3690 X5 memory population order

Memory DIMM installation is key to maximizing system performance. How to install DIMMs is now described.

Figure 4-13 shows the slot numbering for DIMM installation.

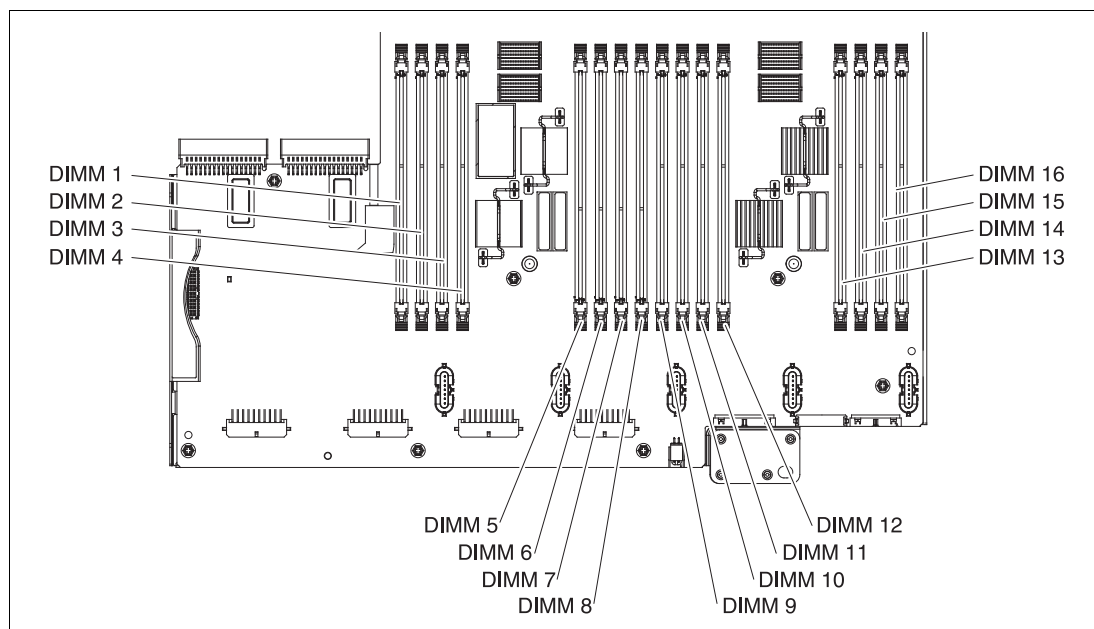


Figure 4-13 x3690 X5 system board showing memory DIMM locations

One or two processors without the memory mezzanine

If one processor is configured without a memory mezzanine, all of the memory of the system directly attaches to processor 1. If a second processor is installed, the system still accesses main memory through processor 1, likely resulting in performance degradation.

Tip: For performance reasons, install and populate the memory mezzanine if you install the second processor. VMware requires equal amounts of memory that is configured for each processor.

When the memory mezzanine is not installed, install the DIMMs in the order that is listed in Table 4-12. Only certain DIMM combinations allow the enablement of hemisphere mode. Hemisphere mode improves memory performance, as described in 2.3.5, “Hemisphere mode” on page 22.

Table 4-12 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 22.

Important: When you install and run VMware ESX on the x3690 X5, the operating system might fail to install. Or, the system might start with the following error message when the server memory configuration is not nonuniform memory access-compliant (NUMA-compliant):

“NUMA node 1 has no memory”

There are only three possible configurations that 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 that is installed in an attached MAX5 memory expansion is configured as non-pooled memory.

Two processors with memory mezzanine installed

With two processors that are installed in the system, distribute the memory evenly between both processors, maximizing system performance.

You must install a minimum of four DIMMs. Figure 4-14 shows the DIMM numbering on the memory mezzanine.

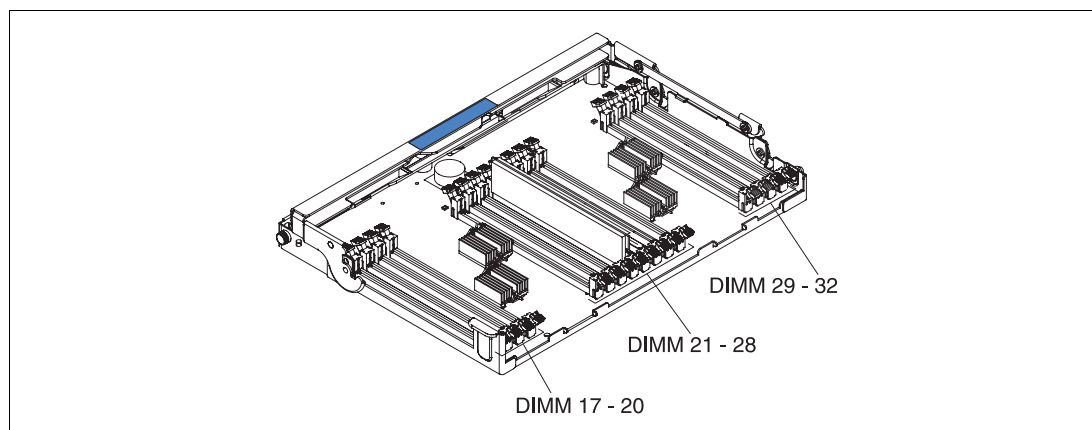


Figure 4-14 Memory mezzanine tray

Install the memory in the order that is listed in Table 4-13.

Table 4-13 DIMM installation: Two processors and the memory mezzanine installed

Number of DIMMs	Hemisphere mode ^a	Processor 1 (system board DIMMs)																Processor 2 (mezzanine DIMMs)															
		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	
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	

a. For more information about hemisphere mode and its importance, see 2.3.5, “Hemisphere mode” on page 22.

Tip: Table 4-13 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>

4.8.4 MAX5 memory population order

The memory that is 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 17, the QPI link speed of the installed processors dictates the maximum SMI link speed. This speed in turn dictates the memory speed.

Section 4.7, “Processor options” on page 130, summarizes the memory speeds for all of the models of Intel Xeon 7500 series CPUs.

Figure 4-15 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 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.

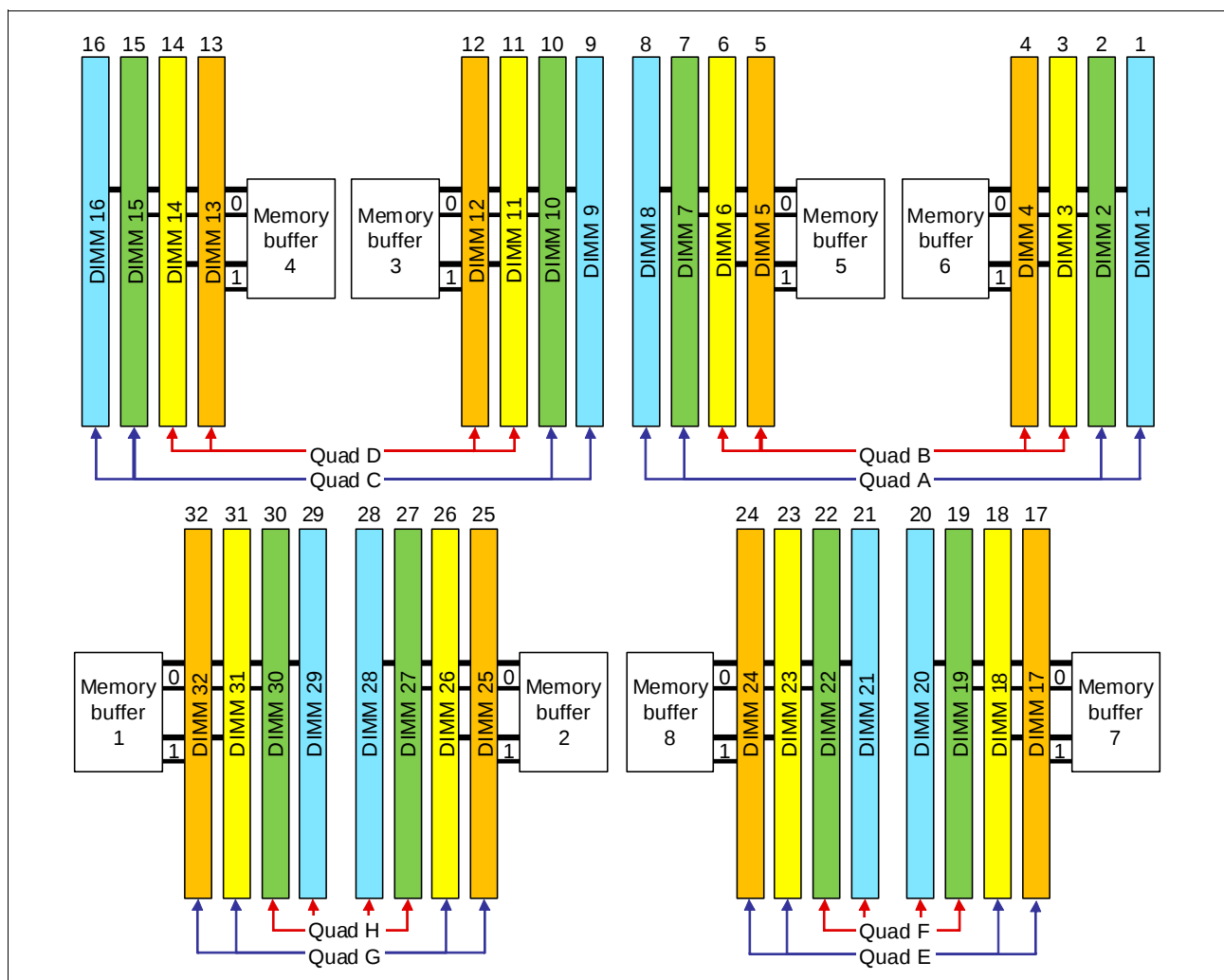


Figure 4-15 DIMM numbering on MAX5

Table 4-14 shows the population order of the MAX5 DIMM slots, ensuring that memory is balanced among the memory buffers. The colors in Table 4-14 match the colors in Figure 4-15 on page 139.

Table 4-14 DIMM installation sequence in the MAX5

DIMM pair	DIMM slot
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
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

4.8.5 Memory balance

The Xeon E7 series processors use a NUMA architecture, as described in 2.3.4, “Non-uniform memory access architecture” on page 21. Because NUMA is used, it is important to ensure that all memory controllers in the system are used by configuring all processors with memory. Populating all processors in an identical fashion is required by VMware and provides a balanced system.

Looking at Figure 4-16 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 memory. Therefore, threads running on Processor 1 have a longer latency to access memory as compared to threads on Processor 0.

This delay is because of 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, it is important to populate all processors with memory and follow the requirements necessary to ensure optimal interleaving and hemisphere mode.

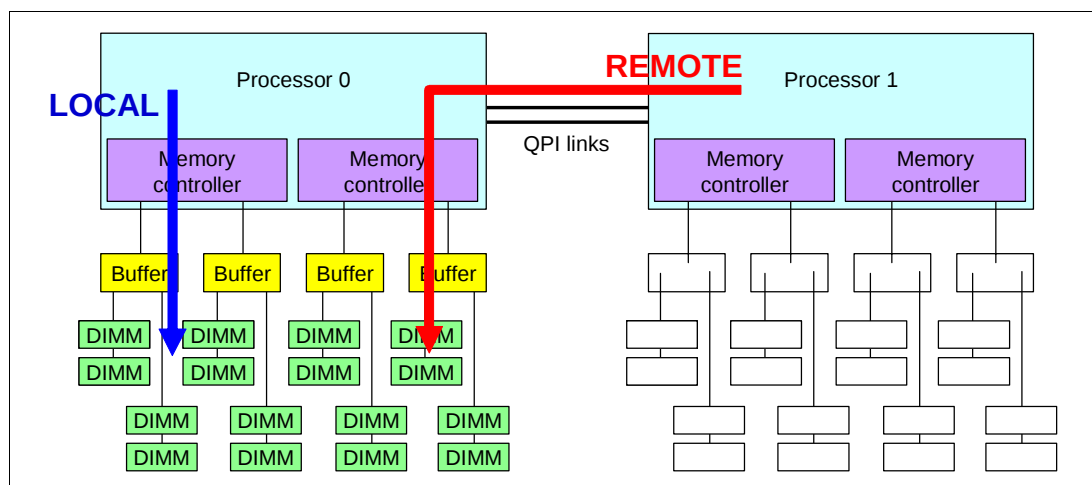


Figure 4-16 Memory latency when not spreading DIMMs across both processors

4.8.6 Mixing DIMMs and the performance effect

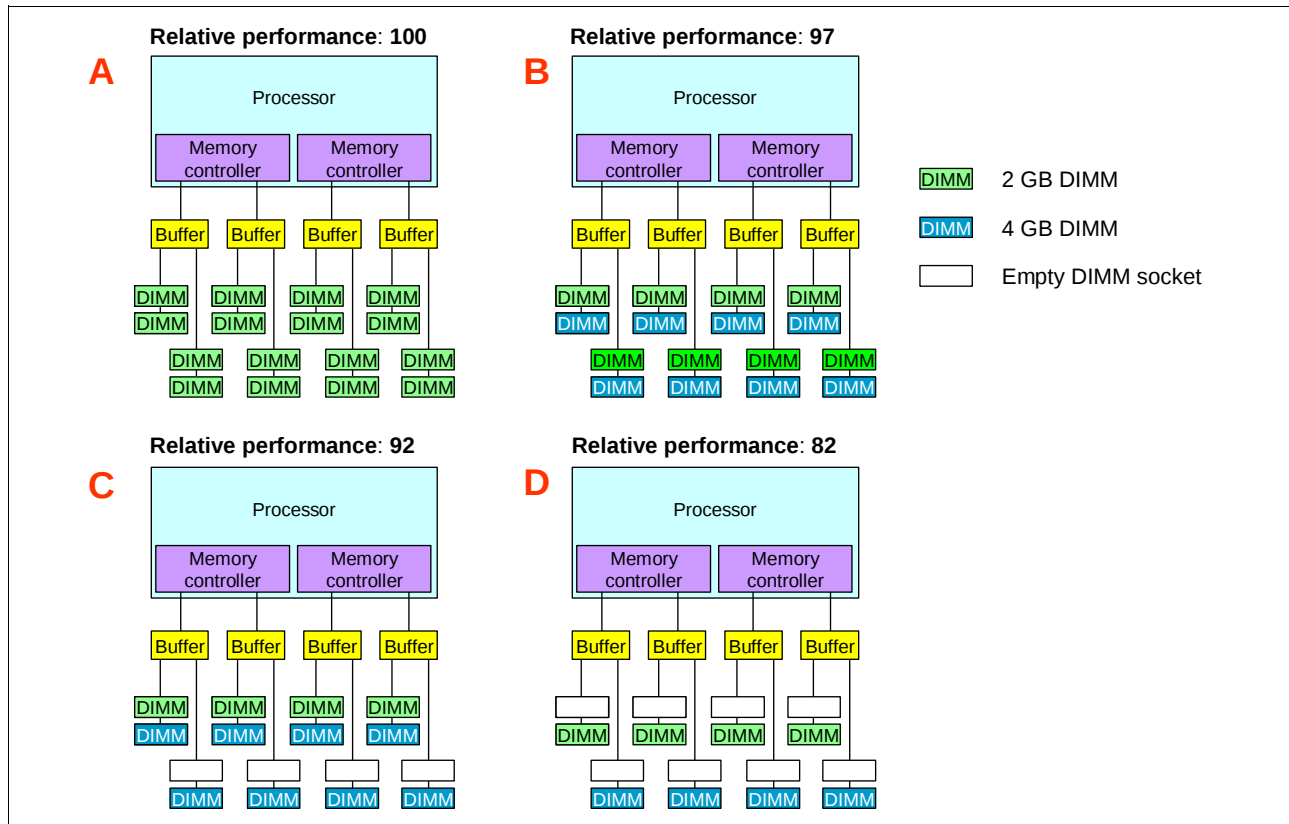
Using DIMMs of various capacities is supported for two reasons:

- ▶ Not all applications require the full memory capacity that a homogeneous memory population provides.
- ▶ Cost-saving requirements might dictate using a lower memory capacity for part of the DIMMs of the platform.

Figure 4-17 on page 142 illustrates the relative performance of three mixed memory configurations that are compared to a baseline of a fully populated memory configuration. Although these configurations use 4 GB (4R x8) and 2 GB (2R x8) DIMMs as specified, similar trends in this data are expected when you use other mixed DIMM capacities. In all cases, memory is populated in minimum groups of four to ensure that hemisphere mode is maintained.

- ▶ 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. However, 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 extra 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.



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.

Memory mirroring is supported by using x3690 X5 and MAX5. The DIMMs must be installed in sets of four for memory-mirroring mode in each server, memory tray, and iMAX5.

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 24.

DIMM installation for 3690 X5

Table 4-15 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 4-15 Mirror DIMM installation: Two processors and no memory mezzanine installed

Number of DIMMs	Processor 1 (system board 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 4-16 shows the DIMM population order for memory-mirroring mode without the mezzanine installed.

Table 4-16 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 4-17 lists the DIMM installation sequence for memory-mirroring mode when two processors and a memory mezzanine tray are installed in the server.

Table 4-17 Mirror DIMM installation: Two processors and memory mezzanine installed

Number of DIMMs	Processor 1 (system board 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 4-18 shows the DIMM population order for memory-mirroring mode with the mezzanine installed.

Table 4-18 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

Dual inline memory module installation: MAX5

Table 4-19 shows the installation guide for MAX5 memory mirroring.

Table 4-19 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	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	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	x	x

4.8.8 Memory sparing

Sparing provides a degree of redundancy in the memory subsystem, but not to the extent that mirroring does. For more information about memory sparing, see “Memory sparing” on page 24. 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 are being spared. 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 must be as large as the rank relative to the highest capacity DIMM that is being spared. 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.

These options are configured by using the UEFI during the boot sequence.

4.8.9 Effect on performance when you use mirroring or sparing

To understand the effect on performance of selecting various memory modes, we use as an example a system configured with x7560 processors and populated with sixty-four 4 GB quad-rank DIMMs.

Figure 4-18 shows the peak system-level memory throughput for various memory modes that are measured by using an IBM-internal memory load generation tool. As shown, there is a 50% decrease in peak memory throughput when you go from a normal (non-mirrored) configuration to a mirrored memory configuration.

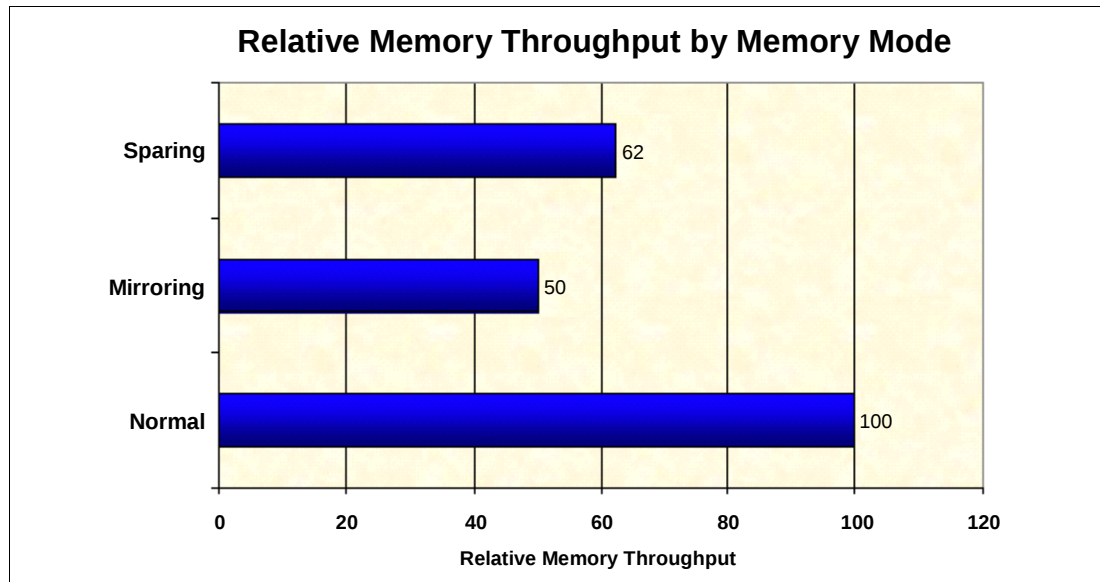


Figure 4-18 Relative memory throughput by memory mode

4.9 Storage

The x3690 X5 has internal capacity of up to sixteen 2.5-inch drives, as shown in Figure 4-19. The server supports 2.5-inch disk drives or SSDs, or 1.8-inch SSDs.



Figure 4-19 Front of the x3690 X5 with sixteen 2.5-inch drive bays

The following topics are covered:

- ▶ 4.9.1, “2.5-inch SAS drive support” on page 146
- ▶ 4.9.2, “IBM eXFlash and SSD 1.8-inch disk support” on page 151
- ▶ 4.9.3, “SAS and SSD controller summary” on page 155
- ▶ 4.9.4, “Battery backup placement” on page 159
- ▶ 4.9.5, “ServeRAID Expansion Adapter” on page 160
- ▶ 4.9.6, “Drive combinations” on page 161
- ▶ 4.9.7, “External direct-attach serial-attached SCSI storage” on page 166
- ▶ 4.9.8, “Optical drives” on page 167

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 SAS disk drives. These drives are connected to the server by 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 that supports four 2.5-inch SAS disks, as listed in 4.3, “Models” on page 122. More backplanes can be added to increase the supported number of SAS disks to 16 (using part number 60Y0381 for a 8x backplane and part number 60Y0339 for a 4x backplane). Certain models have extra backplanes as standard. See 4.3, “Models” on page 122 for details. The standard backplanes are installed in the leftmost sections.

Table 4-20 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 161.

Table 4-20 x3690 X5 hard disk 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, refer to Table 4-21.

As listed in Table 4-20, 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 disk drive backplane for 2.5-inch bays 1 - 8 (the left half of the drive bays in the server when viewed from the front). The *long* SAS cable is used for hard disk drive backplanes for 2.5-inch bays 9 - 16 (the right half of the drive bays when viewed from the front).

When you configure 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 4-21.

Table 4-21 x3690 X5 SAS cable options (not needed if you order 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

Cables: When you use the ServeRAID expansion 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 by using the long SAS cables, as shown in Table 4-21.

Figure 4-20 and Figure 4-22 on page 152 show the backplanes and their cable connections.

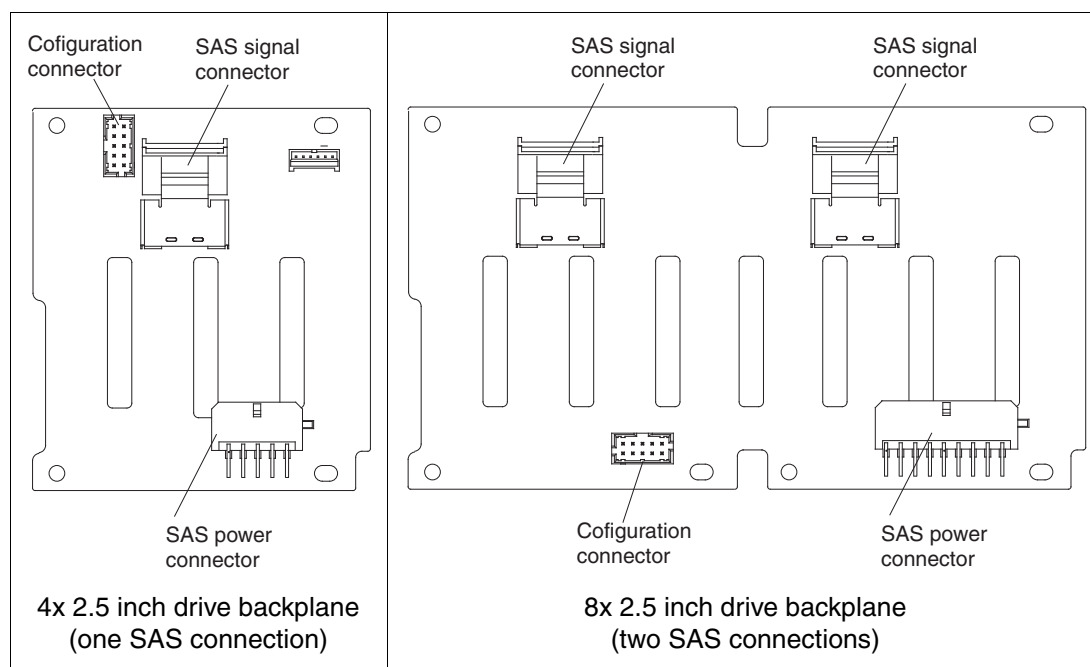


Figure 4-20 The 2.5-inch SAS backplanes (rear view)

Using 2.5-inch disk drives

Table 4-22 lists the 2.5-inch drives that are supported in the x3690 X5.

Table 4-22 Supported 2.5-inch drives

Part number	Feature	Description
2.5-inch SSD		
00W1125	A3HR	IBM 100GB SATA 2.5" MLC HS Enterprise SSD
49Y5839	A3AS	IBM 64GB SATA 2.5" MLC HS Enterprise Value SSD
90Y8648	A2U4	IBM 128GB SATA 2.5" MLC HS Enterprise Value SSD
90Y8643	A2U3	IBM 256GB SATA 2.5" MLC HS Enterprise Value SSD
49Y5844	A3AU	IBM 512GB SATA 2.5" MLC HS Enterprise Value SSD
43W7718	A2FN	IBM 200GB SATA 2.5" MLC HS SSD
49Y6129	A3EW	IBM 200GB SAS 2.5" MLC HS Enterprise SSD
49Y6134	A3EY	IBM 400GB SAS 2.5" MLC HS Enterprise SSD
49Y6139	A3F0	IBM 800GB SAS 2.5" MLC HS Enterprise SSD
2.5-inch 15K SAS hot-swap HDD		
90Y8926	A2XB	IBM 146GB 15K 6Gbps SAS 2.5" SFF G2HS HDD
42D0677	5536	IBM 146GB 15K 6Gbps SAS 2.5" SFF Slim-HS HDD
81Y9670	A283	IBM 300GB 15K 6Gbps SAS 2.5" SFF HS HDD

Part number	Feature	Description
2.5-inch 15K SAS hot-swap self-encrypting drives (SEDs)		
90Y8944	A2ZK	IBM 146GB 15K 6Gbps SAS 2.5" SFF G2HS SED
44W2294	5412	IBM 146GB 15K 6Gbps SAS 2.5" SFF Slim-HS SED
81Y9662	A3EG	IBM 900GB 10K 6Gbps SAS 2.5" SFF G2HS SED
2.5-inch 10K SAS hot-swap HDDs		
90Y8877	A2XC	IBM 300GB 10K 6Gbps SAS 2.5" SFF G2HS HDD
42D0637	5599	IBM 300GB 10K 6Gbps SAS 2.5" SFF Slim-HS HDD
90Y8872	A2XD	IBM 600GB 10K 6Gbps SAS 2.5" SFF G2HS HDD
49Y2003	5433	IBM 600GB 10K 6Gbps SAS 2.5" SFF Slim-HS HDD
81Y9650	A282	IBM 900GB 10K 6Gbps SAS 2.5" SFF HS HDD
2.5-inch 10K SAS hot-swap SEDs		
90Y8913	A2XF	IBM 300GB 10K 6Gbps SAS 2.5" SFF G2HS SED
44W2264	5413	IBM 300GB 10K 6Gbps SAS 2.5" SFF Slim-HS SED
90Y8908	A3EF	IBM 600GB 10K 6Gbps SAS 2.5" SFF G2HS SED
2.5-inch NL SAS hot-swap HDDs		
81Y9690	A1P3	IBM 1TB 7.2K 6Gbps NL SAS 2.5" SFF HS HDD
90Y8953	A2XE	IBM 500GB 7.2K 6Gbps NL SAS 2.5" SFF G2HS HDD
42D0707	5409	IBM 500GB 7200 6Gbps NL SAS 2.5" SFF Slim-HS HDD
2.5-inch NL SATA hot-swap HDDs		
81Y9730	A1AV	IBM 1TB 7.2K 6Gbps NL SATA 2.5" SFF HS HDD
81Y9722	A1NX	IBM 250GB 7.2K 6Gbps NL SATA 2.5" SFF HS HDD
81Y9726	A1NZ	IBM 500GB 7.2K 6Gbps NL SATA 2.5" SFF HS HDD
2.5-inch SATA hot-swap HDDs		
None ^a	5414	IBM 500GB 7200 SATA 2.5" SFF Slim-HS HDD

a. Available by special bid or CTO only

Self-encrypting drives (SEDs) are also an available option as listed in Table 4-22 on page 148. SEDs provide cost-effective advanced data security with Advanced Encryption Standard (AES) 128-disk encryption. To use 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, or a ServeRAID M5016 controller. SEDs can be used in place of non-SEDs, although the data is not encrypted. See "Controller options with 2.5-inch drives" on page 151 for details.

For more information about SEDs, see the IBM Redbooks Product Guide, *Self-Encrypting Drives for IBM System x*, TIPS0761, available at this web page:

<http://www.ibm.com/redbooks/abstracts/tips0761.html>

Single 500 GB SATA drive

The x3690 X5 optionally supports the IBM x3690 X5 Single Serial Advanced Technology Attachment (SATA) HDD Bay. This bay 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 4-23) is installed in the rightmost HDD bay, closest to the information panel and encompassing drive bays 12 - 15. See Figure 4-21. No additional drives can be used in bays 12 - 14 because the bays are covered by a filler panel.

Single SATA drive: The single SATA drive uses the same connector on the system board as the DVD-ROM drive. Therefore, the DVD-ROM drive cannot be installed when you use the SATA drive.

Table 4-23 x3690 X5 Single SATA HDD Bay kit

Option	Feature code	Description
60Y0333	9284	IBM x3690 X5 Single SATA HDD Bay kit

Figure 4-21 shows the location of the single SATA drive in the x3690 X5.

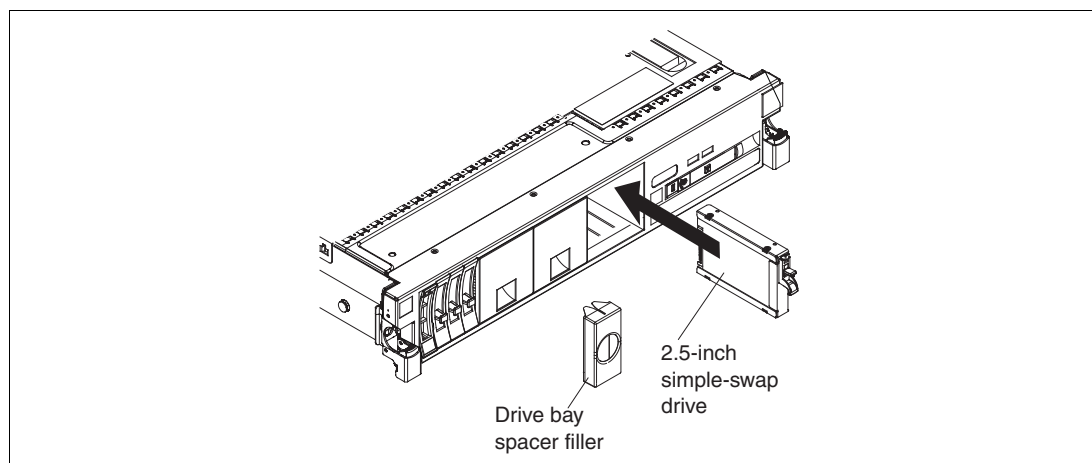


Figure 4-21 Location of the single SATA drive in the x3690 X5

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 cable of the optical drive from the system board connector.
3. Plug the cable from the single SATA HDD bay into the connector on the system board.
4. Install the SATA drive into drive bay 15.

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

SSD support: As listed in Table 4-22 on page 148, the 2.5-inch 200 GB SSD is also supported by the standard SAS backplane and the SAS backplane option. The part numbers are 60Y0381 and 60Y0339. The 2.5-inch SSD is not compatible with the 1.8-inch SSD eXFlash backplane, 60Y0360.

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 2.5-inch SAS backplane.

Controller options with 2.5-inch drives

Table 4-24 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 122.

Table 4-24 RAID controllers that are compatible with SAS backplane and SAS disk drives

Part number	Feature code	Description
60Y0309	4164	IBM x3690 X5 RAID Expansion Adapter
46M0831	0095	ServeRAID M1015 SAS / SATA Controller (standard on most models; see 4.3, “Models” on page 122)
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 ^b	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 ^b	A10C	IBM ServeRAID M5000 Performance Accelerator Key: Adds Cut Through I/O (CTIO) for SSD fast path optimization on ServeRAID M5014, M5015, and M5025 controllers.
90Y4304	A2NF	ServeRAID M5016 SAS/SATA Controller

a. The battery is not included with the ServeRAID M5015 if ordered by using the feature code, but it is not needed if you use only SSDs.

b. Only one key is supported in each controller, either the Advance Feature Key or the Performance Accelerator Key.

4.9.2 IBM eXFlash and SSD 1.8-inch disk support

IBM eXFlash is the name of the feature of the x3690 X5 that offers high-performance 1.8-inch SSDs through optimized eXFlash SSD backplanes and SSD controllers.

IBM eXFlash SSD offerings

IBM eXFlash is available as an option on all models. However, workload-optimized models of the x3690 X5 include IBM eXFlash SSD backplanes that support eight 1.8-inch SSDs. 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 4-25. The eight 1.8-inch drive bays require the same physical space as four 2.5-inch SAS hard disk bays. A single eXFlash backplane requires two SAS x4 input cables and one custom power/configuration cable (shipped standard). Up to three SSD backplanes and 24 SSDs are supported in the x3690 X5 chassis.

For more information about eXFlash and SSD, including a brief overview of the benefits from using eXFlash, see 2.8, “IBM eXFlash” on page 38.

Table 4-25 x3690 X5 hard disk drive backplanes

Part number	Feature code	Backplane	Drives supported	SAS cables included
60Y0360	9281	IBM eXFlash 8x 1.8-inch HS SAS SSD Backplane	Eight 1.8-inch solid-state drives	2 short, 2 long

Figure 4-22 shows the 8x 1.8-inch SSD backplane with its two SAS connectors.

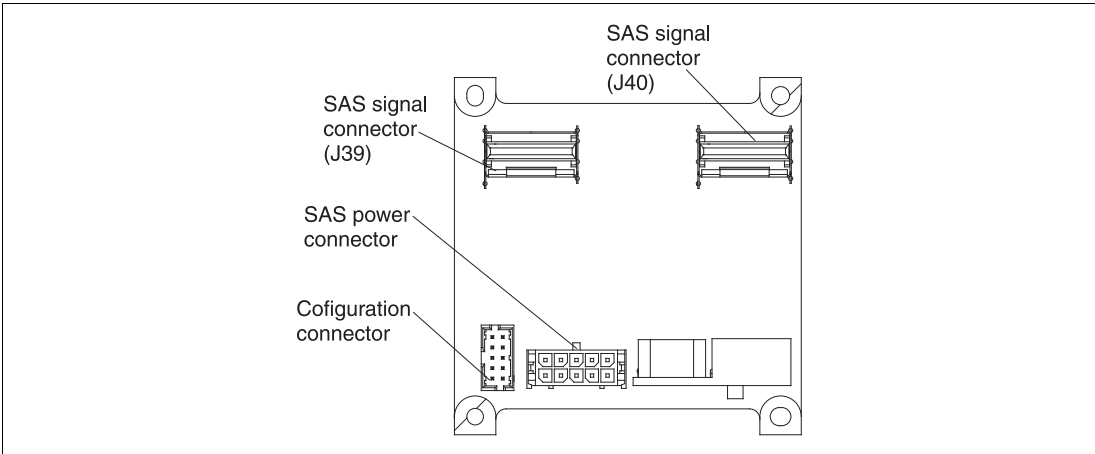


Figure 4-22 8x 1.8-inch SSD backplane (rear view)

Table 4-26 lists the supported 1.8-inch SSDs.

Table 4-26 x3690 X5 1.8-inch SSD options

Part number	Feature code	Description
00W1120	A3HQ	IBM 100 GB SATA 1.8" MLC Enterprise SSD
49Y6119	A3AN	IBM 200 GB SATA 1.8" MLC Enterprise SSD
49Y6124	A3AP	IBM 400® GB SATA 1.8" MLC Enterprise SSD
49Y5834	A3AQ	IBM 64 GB SATA 1.8" MLC Enterprise Value SSD
00W1222	A3TG	IBM 128 GB SATA 1.8" MLC Enterprise Value SSD
00W1227	A3TH	IBM 256 GB SATA 1.8" MLC Enterprise Value SSD
49Y5993	A3AR	IBM 512 GB SATA 1.8" MLC Enterprise Value SSD
43W7726	5428	IBM 50 GB SATA 1.8" MLC SSD
43W7746	5420	IBM 200 GB SATA 1.8" MLC SSD

The failure rate of SSDs is low because, in part, the drives have no moving parts. The SSD is an enterprise grade multi-level cell (eMLC) with NAND flash chips. They also include discrete capacitors to assure there is enough energy to fully commit writes in the event of a power disruption, data error checking and correction circuitry, I/O path error checking and correction circuitry, and thermal monitoring and reporting. Wear-levelling algorithms are used, and cell usage counts are recorded and reported. Using these technologies, the use of RAID redundancy might not always be necessary. Therefore, in certain cases, RAID level 0 might be an acceptable solution.

Enterprise Value SSDs and Enterprise SSDs have similar read and write IOPS performance. However, the key difference between them is their endurance, that is, how long they can perform write operations because SSDs have a finite number of program and erase cycles. Enterprise Value SSDs have a better cost/IOPS ratio but lower endurance when compared to Enterprise SSDs.

For more information about Enterprise SSDs and Enterprise Value SSDs, see the IBM Redbooks Product Guide *IBM SATA 1.8-inch and 2.5-inch MLC Enterprise Value SSDs* at the following website:

<http://www.redbooks.ibm.com/abstracts/tips0879.html>

Table 4-27 lists the controllers that support SSDs.

Table 4-27 Controllers that are supported by the eXFlash SSD backplane option

Part number	Feature code	Description
46M0912	3876	IBM 6 Gb Performance Optimized HBA (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 fast path optimization on ServeRAID M5014, M5015, and M5025 controllers.
90Y4304 ^b	A2NF	ServeRAID M5016 SAS/SATA Controller

a. Add the Performance Accelerator Key to the ServeRAID M5015 or M5014 for use with SSDs.

b. The ServeRAID M5016 includes Performance Accelerator Key functionality.

Important: When you use M5000 series controllers with only SSD drives, disable write-back caching to reduce latency. In a mixed SSD and HDD environment, use battery-backed cache.

If you 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. Disable the cache by going into

the MegaRAID web BIOS configuration utility and disabling Disk Cache and Default Write, as shown in Figure 4-23.

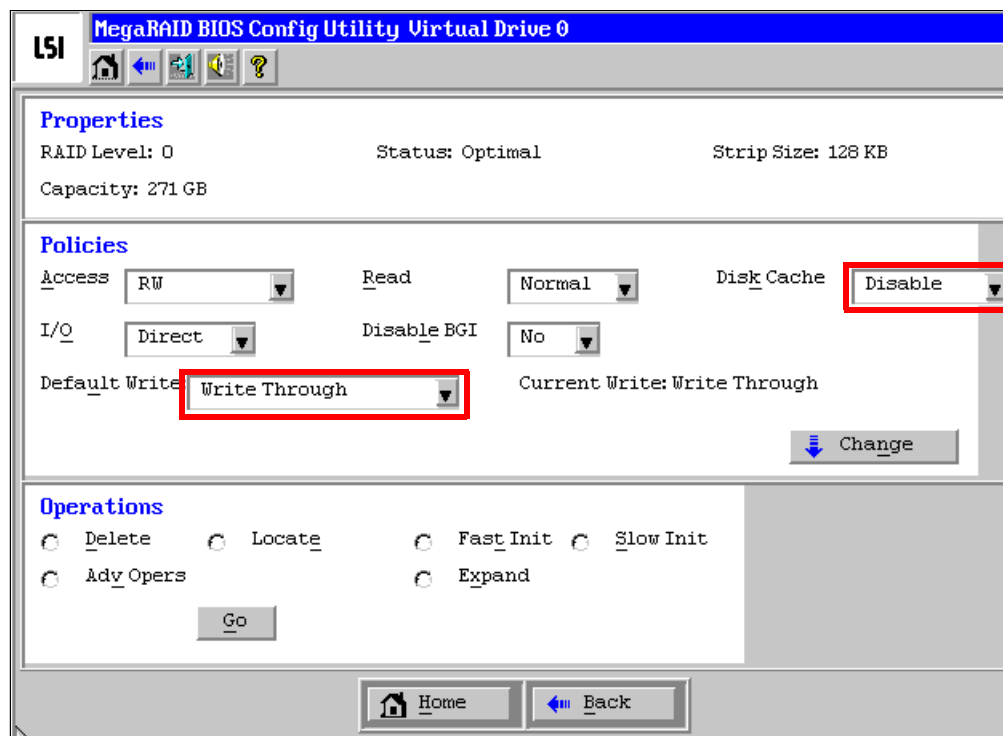


Figure 4-23 Disabling battery cache on controller in MegaRAID web BIOS

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 M5xxx 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 Product Guide *ServeRAID M5000 Series Performance Accelerator Key for IBM System x*, available at this web page:

<http://www.ibm.com/redbooks/abstracts/tips0799.html>

4.9.3 SAS and SSD controller summary

This section provides details for the features of each controller card and what they offer. Table 4-28 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 122 for more information.

Table 4-28 Disk controllers that are compatible with the x3690 X5

Part number	Feature code	Name	Supports 2.5-inch SAS backplane	Supports eFlash SSD backplane	Write-cache protection	Cache	RAID support
46M0831	0095	ServeRAID M1015	Yes	Yes	No	None	0, 1, 10, 5, and 50 ^a
46M0916	3877	ServeRAID M5014	Yes	Yes	Optional Battery	256 MB ^b	0, 1, 10, 5, 50, 6, 60 ^c
46M0829	0093	ServeRAID M5015	Yes	Yes	Battery ^d	512 MB ^b	0, 1, 10, 5, 50, 6, 60 ^c
46M0912	3876	6 Gb Performance Optimized HBA	No	Yes	No	None	No
46M0969	3889	ServeRAID B5015 SSD	No	Yes	No	None	1 and 5
90Y4304	A2NF	ServeRAID M5016	Yes	Yes	Capacitor	1 GB ^b	0, 1, 10, 5, 50, 6, 60

a. M1015 support for RAID-5 and RAID-50 requires the M1000 Advanced Feature Key (46M0832, feature code 9749).

b. Disable write caching in SSD-only implementations

c. M5014 and M5015 support for RAID-6 and RAID-60 requires the M5000 Advanced Feature Key (46M0930, feature code 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.

ServeRAID M5014 and M5015 Controller

The ServeRAID M5014 and M5015 adapters 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 you use all SSD drives. If you use a controller in a mixed environment with SSD and SAS, you can order and use a battery and the Performance Enablement Key.

- ▶ Support for RAID levels 0, 1, 5, 10, and 50 (RAID-6 and RAID-60 support with the optional M5000 Advanced Feature Key).
- ▶ Connection of up to 32 SAS or SATA drives.
- ▶ Support for SAS and SATA drives, 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 important when you use 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: The Performance Accelerator Key uses the same features as the Advanced Feature Key. However, the Performance Accelerator Key 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 web page:

<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 RoC controller
- ▶ x8 PCI Express 2.0 host interface

- ▶ Support for RAID levels 0, 1, and 10 (and also RAID levels 5 and 50 with the 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
- ▶ Compliance with DDF COD
- ▶ S.M.A.R.T. support

RAID-5, RAID-50, and SED technology are optional upgrades to the ServeRAID M1015 adapter, with the addition of the ServeRAID M1000 Series Advanced Feature Key. The part number is 46M0832, feature 9749.

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

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

IBM 6 Gb Performance Optimized host bus adapter

The IBM 6 Gb Performance Optimized host bus adapter (HBA) is an ideal host bus adapter 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 6 Gb Performance Optimized HBA 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 533 MHz
- ▶ UEFI support

For more information, see *IBM 6 Gb Performance Optimized HBA*, TIPS0744, available at the following web page:

<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 has the part number 46M0912. Do not confuse this variant 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 millions of instructions per second (MIPS) processing cores
- ▶ High performance contention-free architecture
- ▶ Up to four ServeRAID B5015 adapters that are supported in a system
- ▶ Support for up to four arrays or logical volumes

For more information, see *ServeRAID B5015 SSD Controller*, TIPS0763, available at the following web page:

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

Important: The ServeRAID B5015 SSD Controller does not use Megaraid. This controller is listed in power-on self-test (POST) and UEFI as a PMC-SIERRA card. The controller also uses maxRAID Storage Manager for management.

ServeRAID M5016 controller

The ServeRAID M5016 adapters have the following specifications:

- ▶ Eight internal 6 Gbps SAS/SATA ports
 - ▶ Two Mini-SAS internal connectors (SFF-8087)
 - ▶ Throughput per port of 6 Gbps
 - ▶ A 800 MHz dual-core PowerPC processor with an LSI SAS2208 6 Gbps RoC controller
 - ▶ PCI Express x8 Gen 2 host interface
 - ▶ Onboard data cache of 1 GB (DDR3 running at 1333 MHz)
 - ▶ CacheVault technology to protect data in cache in case of critical power or server failure
- CacheVault flash cache protection uses NAND flash memory that is powered by a supercapacitor to protect data that is stored in the controller cache. This module eliminates the need for a lithium-ion battery that is commonly used to protect DRAM cache memory on PCI RAID controllers.

To avoid the possibility of data loss or corruption during a power or server failure, CacheVault technology transfers the contents of the DRAM cache to NAND flash (CacheVault flash module (CVFM)) by using power from the CacheVault power module (CVPM). After the power is restored to the M5016 RAID controller, CacheVault technology transfers the contents of the NAND flash back to the DRAM, which is then flushed to disk.

- ▶ Supports RAID levels 0, 1, 5, 6, 10, 50, and 60
- ▶ Connects to up to 128 SAS or SATA drives
- ▶ Intermix of SAS and SATA drives are supported, but the mixing of SAS and SATA drives in the same RAID array is not supported
- ▶ Supports up to 64 logical volumes
- ▶ Supports LUN sizes up to 64 TB
- ▶ Configurable stripe size up to 1 MB
- ▶ Compliant with DDF COD
- ▶ S.M.A.R.T. support
- ▶ 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 Product Guide *ServeRAID M5016 SAS/SATA Controller*, TIPS0847, which is available at the following web page:

<http://www.redbooks.ibm.com/abstracts/tips0847.html>

4.9.4 Battery backup placement

When you install RAID adapters that include batteries, the RAID batteries must be remotely located to prevent them 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 4-29 lists the kit to order for a remote battery cable.

Table 4-29 Remote battery cable order

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 4-24 shows these components.

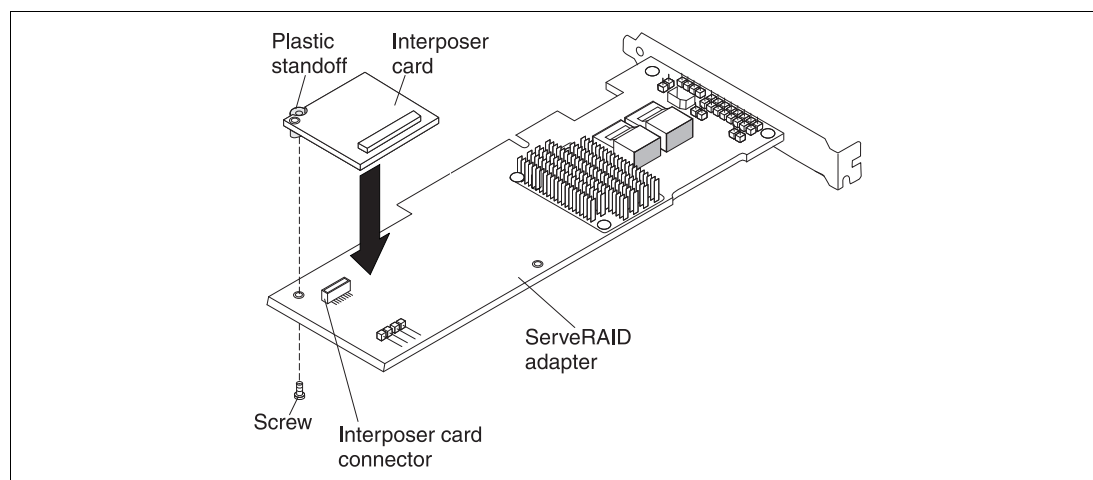


Figure 4-24 Remote battery cable kit

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 one is installed, or to the air baffle, which is

in place of the mezzanine. Figure 4-25 shows how the battery trays are installed in the memory mezzanine. Each battery tray can hold two batteries.

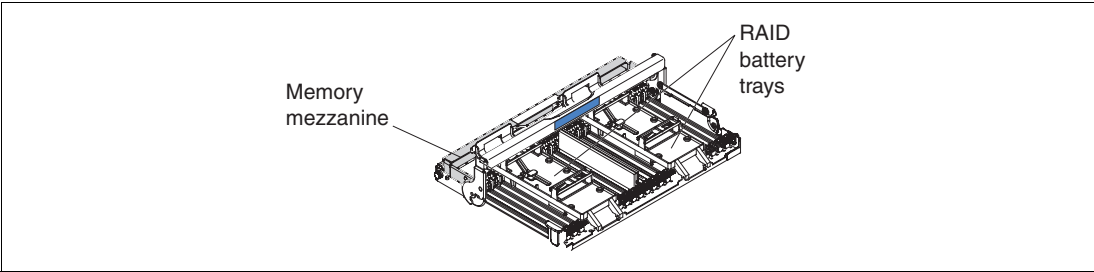


Figure 4-25 RAID battery trays on the memory mezzanine

4.9.5 ServeRAID Expansion Adapter

The ServeRAID Expansion Adapter, also known as the *IBM x3690 X5 RAID Expansion Adapter* or *IBM 4x4 Drive Backplane ServeRAID Expansion Adapter*, is a SAS expander. With the adapter, you can create RAID arrays of up to 16 drives and across up to four backplanes. Table 4-30 shows the ordering information.

Table 4-30 ServeRAID Expansion Adapter order

Option	Feature code	Description
60Y0309	4164	ServeRAID Expansion Adapter

The card, which is shown in Figure 4-26, has two input connectors that you connect to a supported RAID controller. Plus, the card has four output connectors to go to each backplane, which allow up to 16 drives to be connected.

Important: You can use only the 2.5-inch hot-swap drive backplanes with this adapter (see Table 4-25).

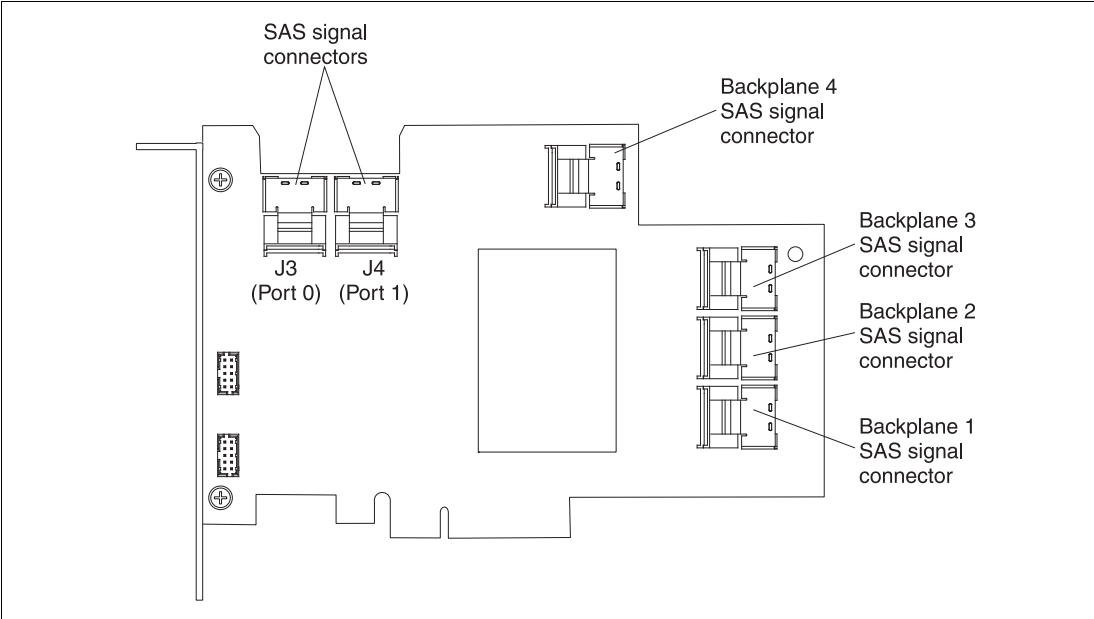


Figure 4-26 ServeRAID Expansion Adapter

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
- ▶ ServeRAID M5016 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 supports various combinations of drives. All combinations are supported. However, not all of these configurations are orderable in a CTO configuration.

Mixing 2.5-inch and 1.8-inch drives and backplanes:

- ▶ 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 you mix 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 viewed from the front of the server).

Configuration with four drives

Figure 4-27 shows a four-drive configuration that uses one 4x HDD backplane. This configuration uses one SAS cable.

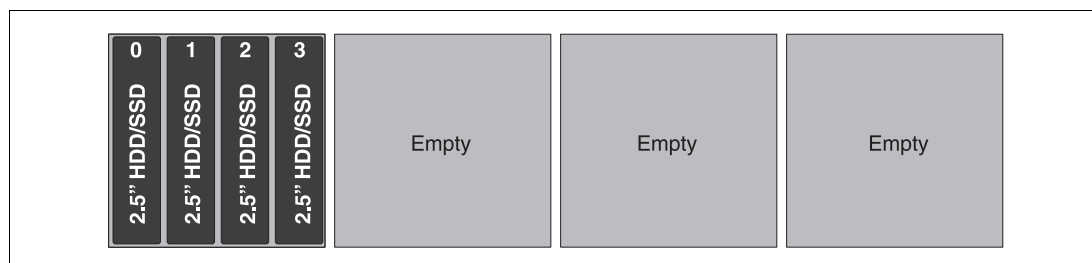


Figure 4-27 x3690 with one IBM 4x 2.5-inch HS SAS HDD backplane

Configurations with eight drives

Figure 4-28 shows a configuration using two 4x HDD backplanes. This configuration requires two SAS cables.

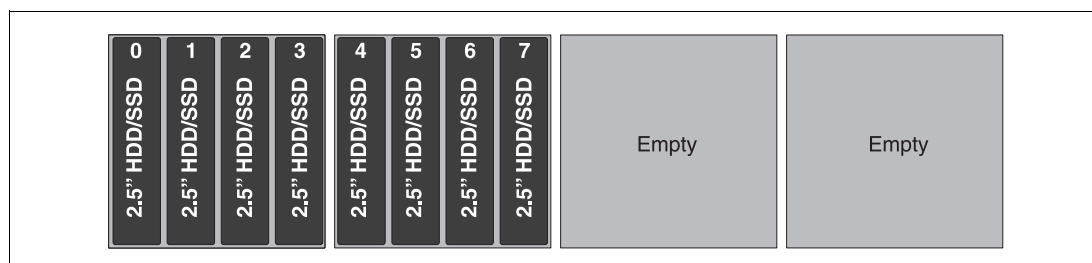


Figure 4-28 x3690 with two IBM 4x 2.5-inch HS SAS HDD backplanes

Figure 4-29 shows a configuration that uses one 8x HDD backplane instead of two 4x HDD backplanes. Two SAS cables are needed.

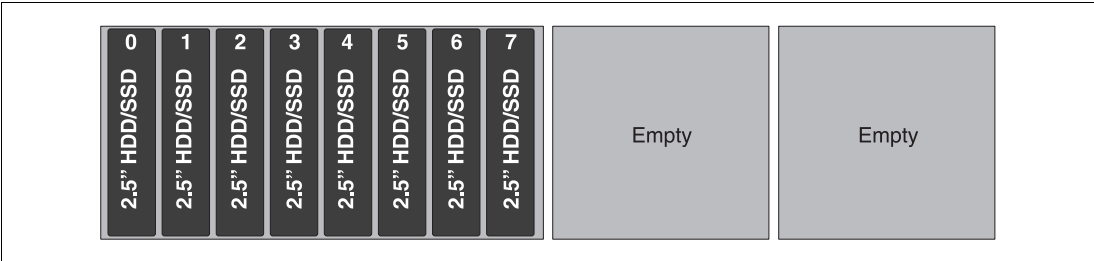


Figure 4-29 x3690 X5 with one IBM 8x 2.5-inch HS SAS HDD backplane

Figure 4-30 illustrates a configuration by using 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.

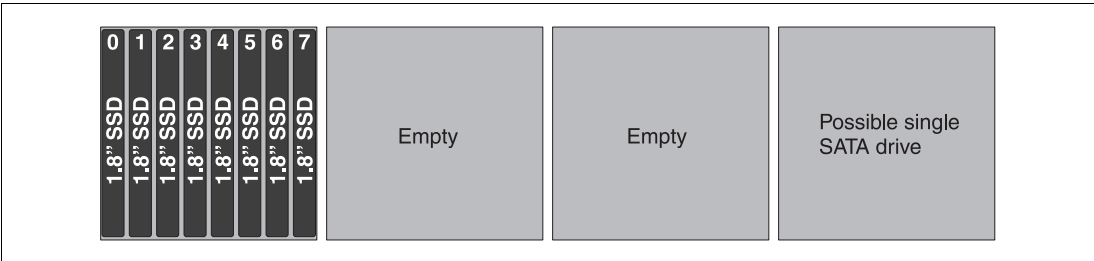


Figure 4-30 x3690 with one IBM eXFlash 8x 1.8-inch HS SAS SSD backplane

Configurations with 12 drives

Figure 4-31 shows three 4x HDD backplanes. This configuration requires three SAS cables.

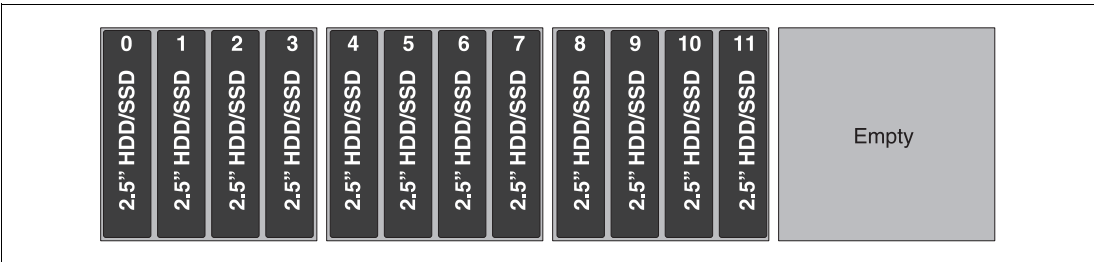


Figure 4-31 x3690 with three IBM 4x 2.5-inch HS SAS HDD backplanes

Figure 4-32 shows one 8x and one 4x HDD backplane, resulting in 12 drives. This configuration also requires three SAS cables.

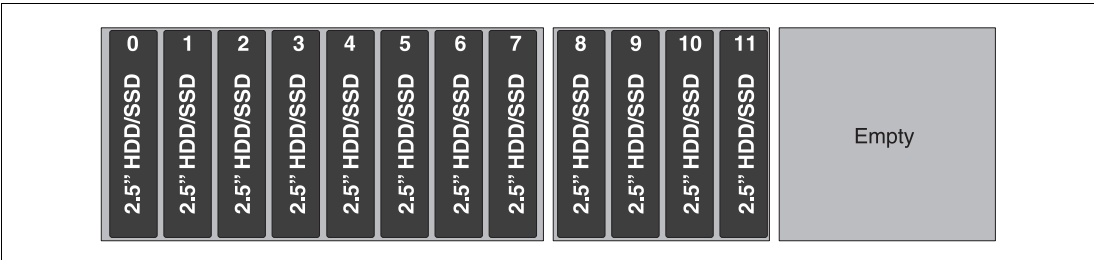


Figure 4-32 x3690 with one 8x 2.5-inch HS SAS HDD and one 4x 2.5-inch HS SAS HDD backplane

Figure 4-33 shows a mixture of 2.5-inch HDDs and 1.8-inch SSDs. This configuration requires three SAS cables.

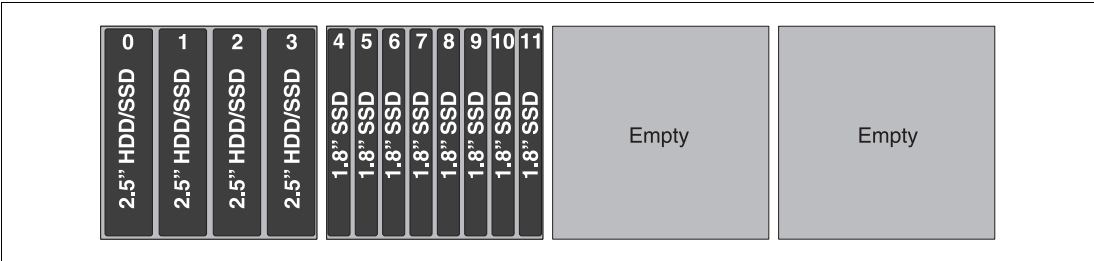


Figure 4-33 x3690 with one 8x 2.5-inch backplane and one eXFlash 8x 1.8-inch SSD backplane

Configurations with 16 drives

Figure 4-34 shows a full sixteen 2.5-inch drive configuration. It requires four SAS cables.

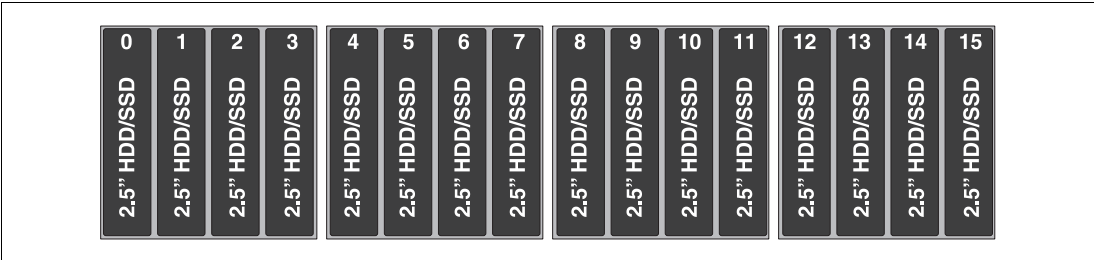


Figure 4-34 x3690 with four IBM 4x 2.5-inch HS SAS HDD backplanes

Figure 4-35 also shows a sixteen 2.5-inch drive configuration. Four SAS cables are required.

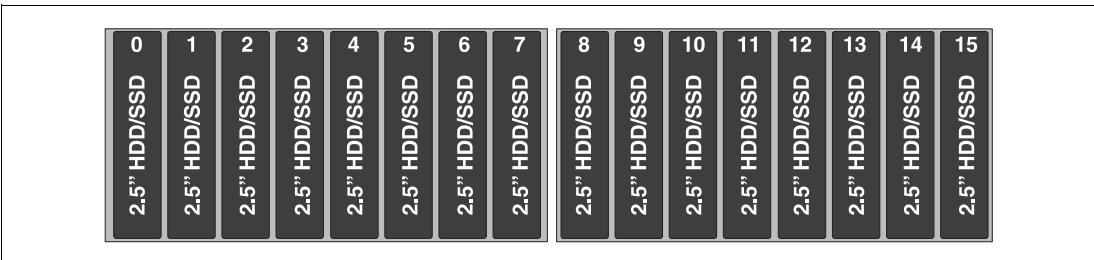


Figure 4-35 x3690 with two IBM 8x 2.5-inch HS SAS HDD backplanes

Figure 4-36 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 4-36 x3690 with one 8x 2.5-inch backplane and two 4x 2.5-inch backplanes

Figure 4-37 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.

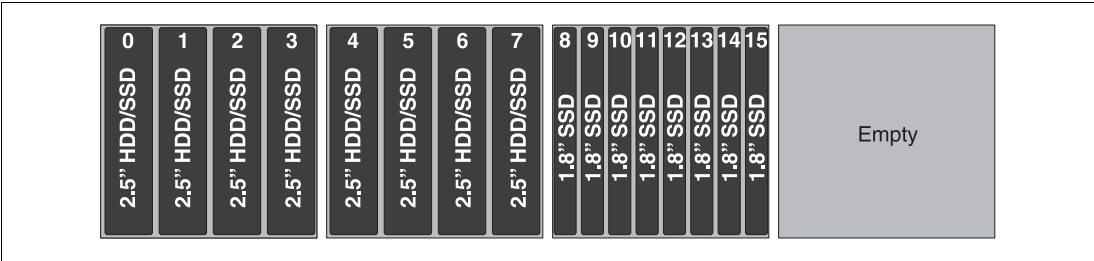


Figure 4-37 x3690 with two 4x 2.5-inch backplanes and one IBM eXFlash 8x 1.8-inch SSD backplane

Figure 4-38 shows two 8x eXFlash backplanes. Using these two backplanes requires four SAS cables. Figure 4-38 also shows the use of a single SATA drive.

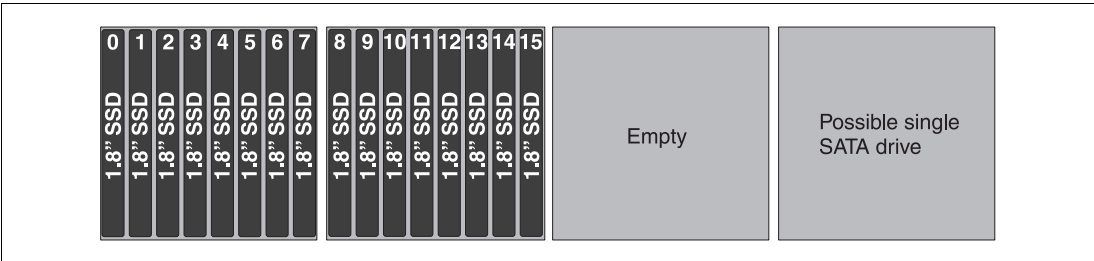


Figure 4-38 x3690 X5 with two IBM eXFlash 8x 1.8-inch SSD backplanes and a single SATA drive

Configurations with 20 drives

Figure 4-39 shows a full complement of drives that uses three 4x backplanes and one 8x eXFlash backplane. You can also achieve this configuration with one 8x backplane, one 4x, and one 8x eXFlash. Either configuration uses five SAS cables.

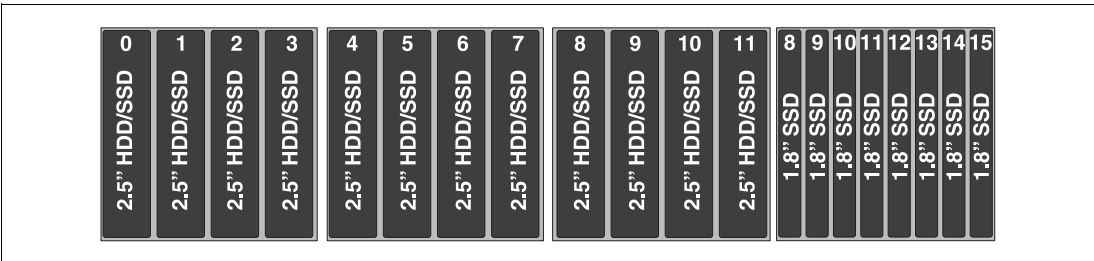


Figure 4-39 x3690 X5 with three 4x 2.5-inch backplanes and one eXFlash 8x 1.8-inch SSD backplane

Figure 4-40 shows one 4x backplane and two 8x eXFlash backplanes. Five SAS cables are needed.

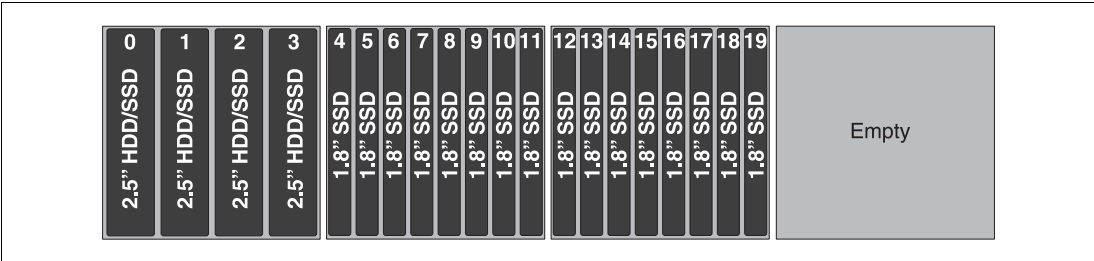


Figure 4-40 x3690 X5 with one 4x 2.5-inch backplane and two eXFlash 8x 1.8-inch SSD backplanes

Configurations with 24 drives

Figure 4-41 shows two 4x backplanes and two 8x eXFlash backplanes. The 8x backplane can be used here instead of two 4x backplanes. Six SAS cables are required.

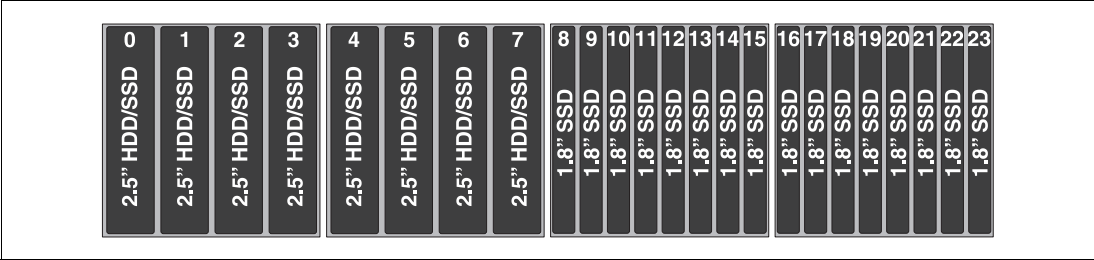


Figure 4-41 x3690 X5 with two 4x 2.5-inch backplanes and two IBM eXFlash 8x 1.8" SSD backplanes

Figure 4-42 shows the maximum number of 8x eXFlash backplanes supported in an x3690 X5. This configuration requires six SAS cables. Figure 4-42 also shows the use of a single SATA drive.

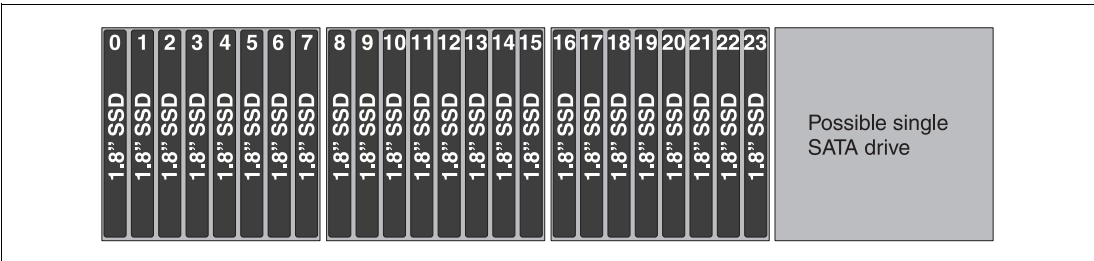


Figure 4-42 x3690 X5 with three IBM eXFlash 8x 1.8-inch SSD backplanes and a possible SATA drive

Important: A configuration of 32 each 1.8-inch drives is not supported.

4.9.7 External direct-attach serial-attached SCSI storage

The x3690 X5 supports the ServeRAID M5025 for external SAS storage connectivity. The M5025 offers two external SAS ports to connect to external storage. Table 4-31 lists the cards, support cables, and feature keys.

Table 4-31 External ServeRAID card

Part number	Feature code	Description
46M0830	0094	IBM ServeRAID M5025 SAS/SATA Controller
39R6531	3707	IBM 3 m SAS external cable for ServeRAID M5025 to an EXP2512 (1747 HC1) or EXP2524 (1747 HC2)
39R6529	3708	IBM 1 m 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 that are implemented through two four-lane (x4) connectors.
- ▶ Two mini-SAS external connectors (SFF-8088).
- ▶ Throughput of 6 Gbps per SAS port.
- ▶ An 800-MHz PowerPC processor with LSI SAS2108 6 Gbps RoC controller.
- ▶ PCI Express 2.0 x8 host interface.
- ▶ A 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 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 support, 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.
- ▶ Compliance with DDF 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.
 - 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.
 - 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, available at the following web page:

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

4.9.8 Optical drives

An optical drive is optional. Table 4-32 lists the supported part numbers.

Table 4-32 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 drive: 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 you use the SATA drive.

4.10 PCIe slots

The x3690 X5 provides five PCIe 2.0 slots for add-in cards. Figure 4-43 shows the location of the slots as viewed from the rear of the server.

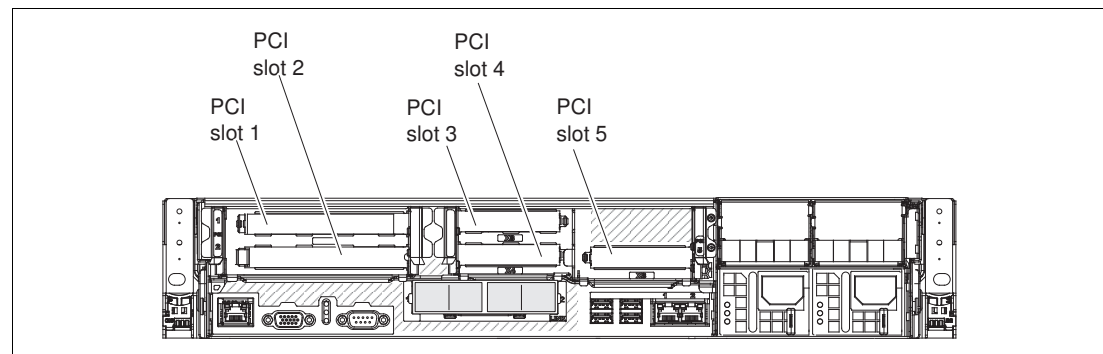


Figure 4-43 x3690 X5 PCIe slots

These slots are connected to the system board through two riser cards, both of which are installed as standard. Figure 4-44 shows the locations of the two riser cards in the server.

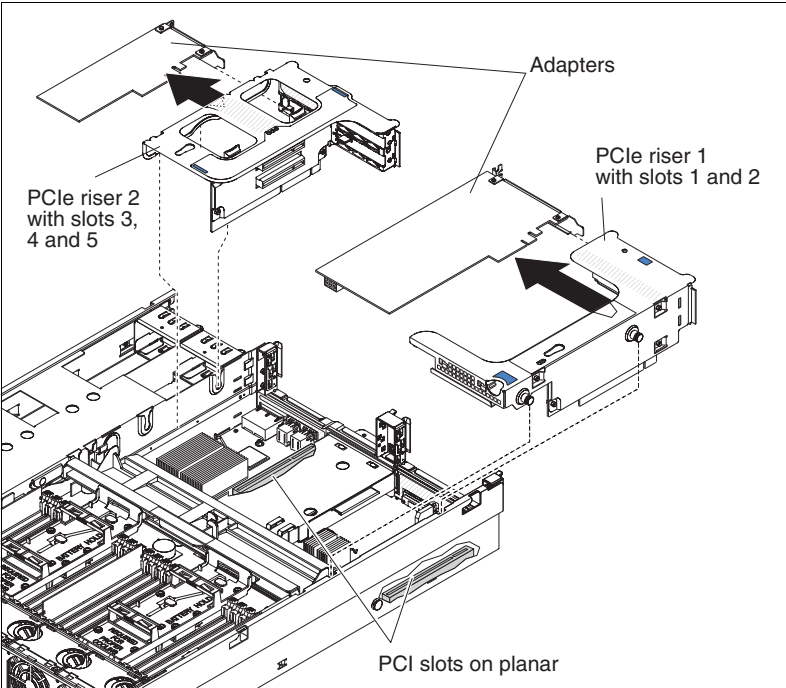


Figure 4-44 Location of the PCIe riser cards in the server

4.10.1 Riser 1

Standard x3690 X5 models contain the following riser slots:

- ▶ Slot 1, PCIe 2.0 x8 full height, full length slot
- ▶ Slot 2, PCIe 2.0 x8 full height, half length slot

Slot 1 contains an installed 2 x8 riser card (60Y0329, feature 9285).

The 2 x8 riser card can be replaced by one of two other riser cards as listed in Table 4-33. One riser card, 60Y0331, offers a single 3/4-length PCI Express x16 expansion slot. The other riser card option, 60Y0337, offers a single full-length PCI Express x16 expansion slot. These x16 expansion slots are suitable for graphics processing unit (GPU) adapters. Extra power for the adapter is available from an onboard power connector if needed. The riser card that offers a full-length slot (60Y0337) cannot be used if the memory mezzanine is installed in the server.

Table 4-33 lists the riser card options for riser slot 1. Only one of the risers that are listed in the table can be installed in the server at a time.

Table 4-33 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 contains a 3 x8 riser card that is installed in all standard models except for model 7148-ARx (see 4.3, “Models” on page 122). The 3 x8 riser card contains slots for the following cards:

- ▶ 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 10 Gb Ethernet adapter is installed in this slot if the adapter is part of the server configuration.

Adapter size: 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 4-34 lists the 3 x8 riser card option.

Table 4-34 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.

Emulex 10 GbE Virtual Fabric Adapter: The Emulex 10 GbE Virtual Fabric Adapter that is standard in most models is installed in slot 5. See 4.10.3, “Emulex 10 Gb Ethernet Adapter” on page 169 for details about the adapter.

4.10.3 Emulex 10 Gb Ethernet Adapter

As described in 4.3, “Models” on page 122, some models include either the Emulex 10 GbE Integrated Virtual Fabric Adapter or the Emulex 10 GbE Integrated Virtual Fabric Adapter II. The card is installed in PCIe slot 5. Slot 5 is a nonstandard x8 slot and is slightly longer than normal. It accepts both standard PCIe adapters and the Integrated 10 Gb Ethernet Adapter.

The integrated 10 GbE adapter is a custom version of the equivalent adapter available as a System x option. The Emulex 10 GbE Integrated Virtual Fabric Adapter II (feature code A148) has the same features and functions as the Emulex 10 GbE Virtual Fabric Adapter II for IBM System x (part number 49Y7950).

The Integrated 10 Gb Ethernet Adapter in the x3690 X5 is customized with a special type of connector that is 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.

Only the x3850 X5 and the x3690 X5 have slots that are compatible with the custom-built Integrated 10 Gb Ethernet Adapter that is shown in Figure 4-45.

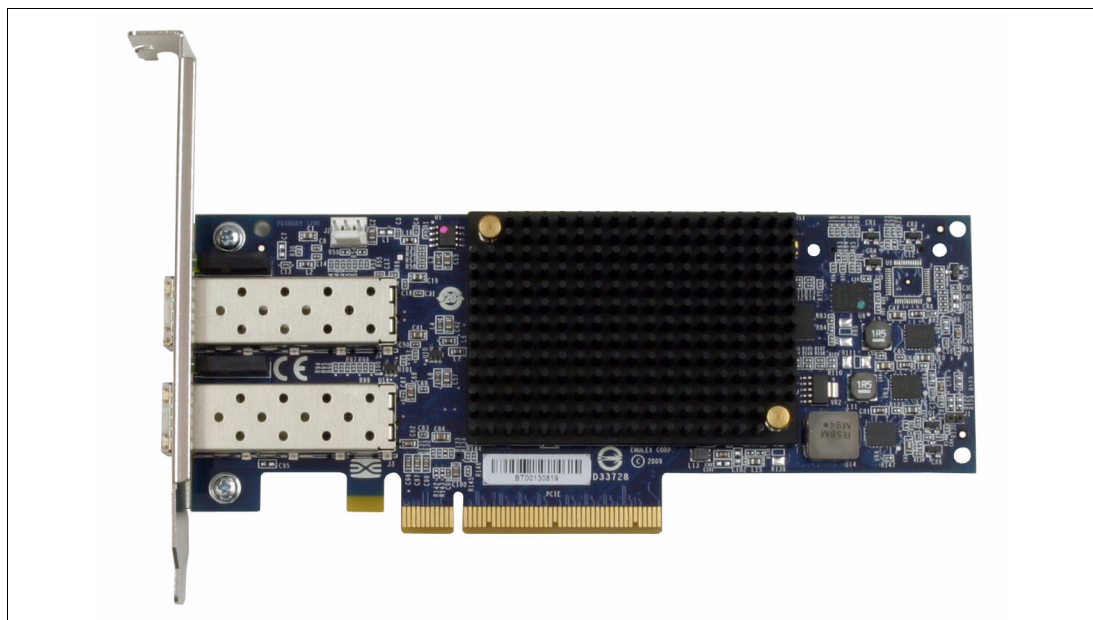


Figure 4-45 Integrated 10 Gb Ethernet Adapter with blue circuit board and longer connector

General details about this card can be found in *Emulex 10GbE Virtual Fabric Adapter II and III family for IBM System x*, which is available at the following web page:

<http://www.redbooks.ibm.com/abstracts/tips0844.html>

Fibre Channel over Ethernet and Internet Small Computer System Interface

upgrades: The Emulex 10 GbE Virtual Fabric Adapter II card supports the Internet Small Computer System Interface (iSCSI) hardware initiator or Fibre Channel over Ethernet (FCoE) feature upgrade. The part number for this upgrade is 49Y4265, feature code 2436.

The Emulex 10 Gb Ethernet Adapter for x3690 X5 includes the following features:

- ▶ Dual-channel, 10 Gbps Ethernet controller
- ▶ Near line-rate 10 Gbps performance
- ▶ Two small form-factor pluggable+ (SFP+) empty cages to support either of the following items:
 - SFP+ short range (SR) link is with SFP+ SR Module with LC connectors
 - SFP+ twinaxial copper link is with SFP+ direct attached copper module and cable

Transceivers: Servers that include the Emulex 10 Gb Ethernet Adapter do not include transceivers. You must order transceivers separately if needed, as listed in Table 4-35 on page 171.

- ▶ TCP/IP stateless offloads
- ▶ TCP chimney offload
- ▶ Basis on Emulex OneConnect technology
- ▶ FCoE support as a future feature entitlement upgrade

- ▶ Hardware parity, cyclic redundancy check (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
- ▶ Virtual local area network (VLAN) insertion and extraction
- ▶ Support for jumbo frames up to 9000 bytes
- ▶ Preboot Execution 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

The Emulex 10 Gb Ethernet Adapter is interoperable with BNT 10 Gb Top of Rack (ToR) switch for FCoE functions and Nexus 5000, Brocade 10 Gb Ethernet switches for NIC/FCoE.

SFP+ transceivers are not included with the server and must be ordered separately. Table 4-35 lists compatible transceivers.

Table 4-35 Transceiver ordering information

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

4.10.4 I/O adapters

Table 4-36 shows the current list in the *Configuration and Options Guide* (COG) at the time of writing this paper. See the following web page:

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

Table 4-36 Available I/O adapters for the x3690 X5

Part number	Feature code	Product description
1 Gb Ethernet		
90Y9370	A2V4	Broadcom NetXtreme I Dual Port GbE Adapter
90Y9352	A2V3	Broadcom NetXtreme I Quad Port GbE Adapter
49Y4230	5767	Intel Ethernet Dual Port Server Adapter I340-T2
49Y4240	5768	Intel Ethernet Quad Port Server Adapter I340-T4
42C1780	2995	NetXtreme II 1000 Express Dual Port Ethernet Adapter
None	1485	NetXtreme II 1000 Express G Ethernet Adapter - PCIe
49Y4220	5766	NetXtreme II 1000 Express Quad Port Ethernet Adapter
42C1750	2975	PRO/1000 PF Server Adapter

Part number	Feature code	Product description
10 Gb Ethernet		
49Y7910	A18Y	Broadcom NetXtreme II Dual Port 10GBaseT Adapter
42C1820	1637	Brocade 10 Gb CNA
None ^a	A148	Emulex 10GbE Integrated Virtual Fabric Adapter II
49Y7950 ^b	A18Z	Emulex 10GbE Virtual Fabric Adapter II
49Y4274 ^b	5715	Emulex VFA II FCoE/iSCSI License (Feature on Demand upgrade for 49Y7950)
95Y3751 ^b	A348	Emulex Dual Port VFAll Adapter & FCoE/iSCSI License
49Y7960	A2EC	Intel X520 Dual Port 10GbE SFP+ Adapter
49Y7970	A2ED	Intel X540-T2 Dual Port 10GBaseT Adapter
81Y9990	A1M4	Mellanox ConnectX-2 Dual Port 10GbE Adapter
42C1800	5751	QLogic 10Gb CNA
16 Gb Fibre Channel HBA		
81Y1675	A2XV	Brocade 16Gb FC Dual-port HBA
81Y1668	A2XU	Brocade 16Gb FC Single-port HBA
81Y1662	A2W6	Emulex 16Gb FC Dual-port HBA
81Y1655	A2W5	Emulex 16Gb FC Single-port HBA
00Y3341	A3KX	QLogic 16Gb FC Dual-port HBA
00Y3337	A3KW	QLogic 16Gb FC Single-port HBA
8 Gb Fibre Channel HBA		
46M6050	3591	Brocade 8Gb FC Dual-port HBA
46M6049	3589	Brocade 8Gb FC Single-port HBA
42D0494	3581	Emulex 8Gb FC Dual-port HBA
42D0485	3580	Emulex 8Gb FC Single-port HBA
42D0510	3579	QLogic 8Gb FC Dual-port HBA
42D0501	3578	QLogic 8Gb FC Single-port HBA
4 Gb Fibre Channel HBA		
59Y1993	3886	Brocade 4Gb FC Dual-port HBA
59Y1987	3885	Brocade 4Gb FC Single-port HBA
42C2071	1699	Emulex 4Gb FC Dual-Port PCI-E HBA
42C2069	1698	Emulex 4Gb FC Single-Port PCI-E HBA
39R6527	3568	QLogic 4Gb FC Dual-Port PCIe HBA
39R6525	3567	QLogic 4Gb FC Single-Port PCIe HBA

Part number	Feature code	Product description
SAS HBA		
46M0912	3876	IBM 6Gb Performance Optimized HBA
46M0907	5982	IBM 6Gb SAS HBA
High IOPS SSD adapters		
46C9078	A3J3	IBM 365GB High IOPS MLC Mono Adapter
46C9081	A3J4	IBM 785GB High IOPS MLC Mono Adapter
46M0877	0096	IBM 160GB High IOPS SS Class SSD PCIe Adapter
46M0878	0097	IBM 320GB High IOPS SD Class SSD PCIe Adapter
46M0898	1649	IBM 320GB High IOPS MS Class SSD PCIe Adapter
90Y4377	A3DY	IBM 1.2TB High IOPS MLC Mono Adapter
90Y4397	A3DZ	IBM 2.4TB High IOPS MLC Duo Adapter

- a. Provided standard with most models of x3690 X5, orderable with CTO systems.
- b. Maximum of three if Emulex 10 GbE Integrated Virtual Fabric Adapter II for IBM System x (FC A148) is not installed. Otherwise, there is a maximum of two.

4.11 Standard features

The standard onboard features of the x3690 X5 are described. The following topics are covered:

- ▶ 4.11.1, “Integrated management module” on page 173
- ▶ 4.11.2, “Ethernet subsystem” on page 174
- ▶ 4.11.3, “USB subsystem” on page 174
- ▶ 4.11.4, “Integrated Trusted Platform Module” on page 175
- ▶ 4.11.5, “Light path diagnostics” on page 175
- ▶ 4.11.6, “Cooling” on page 175
- ▶ 4.12, “Power supplies” on page 177

4.11.1 Integrated management module

The x3690 X5 contains the Vitesse VSC452 integrated management module (IMM). The module combines the baseboard management controller (BMC), video controller, and Remote Supervisor Adapter (RSA) II and concurrent keyboard video mouse (KVM) functions into a single chip.

The VSC452 includes the following major features:

- ▶ 300 MHz 32-bit processor
- ▶ BMC I/O, including I²C and general-purpose I/Os
- ▶ Matrox G200 Video core
- ▶ DDR2-250 MHz memory controller
- ▶ USB 2.0 configurable peripheral device
- ▶ Avocent digital video compression

The IMM includes the following system management features:

- ▶ Environmental monitor with fan speed control for temperature, voltages, fan failure, power supply failure, and power backplane failure indicators
- ▶ Light path indicators to report system errors, and failure of fans, power supplies, CPU, or voltage regulator module (VRM)
- ▶ 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 (IPMI) Specification V2.0 and Intelligent Platform Management Bus (IPMB) support
- ▶ Serial Over LAN (SOL)
- ▶ Active Energy Manager
- ▶ Power and 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 includes the following features:

- ▶ TCP offload 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 are 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 are on the front of the server as shown in Figure 4-2 on page 117, and four are on the rear of the server, as shown in Figure 4-3 on page 118.

The server also has two internal USB ports, which are on riser card 2, as shown in Figure 4-49 on page 178. One of these internal ports is used for the integrated hypervisor key. The other internal port is available for other USB devices.

See 4.13, “Integrated virtualization” on page 178 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. The TPM provides the hardware support for the Trusted Computing Group (TCG) specification. For more information about the TCG specification, go to the following web page:

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 4-46 shows the location of the light path diagnostics panel on the x3690 X5.



Figure 4-46 x3690 X5 light path diagnostics panel

Light path diagnostics can alert the user of 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 web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5085206>

4.11.6 Cooling

The x3690 X5 includes the following fans:

- ▶ Five hot-swappable fans that are in the front portion of the chassis
- ▶ Power supply internal fans that are 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 to the power supplies, and Fan 5 is nearest to the operator information panel. Figure 4-47 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 4-47 shows the location of the x3690 X5 fans.

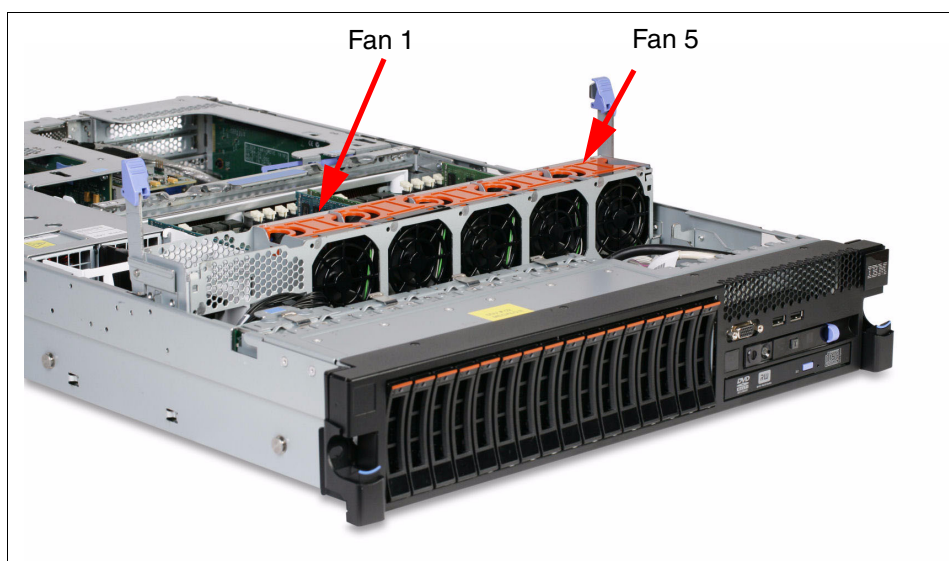


Figure 4-47 x3690 X5 fans

Figure 4-48 shows the top of the server and the hot-swap fan access panel.



Figure 4-48 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

The power subsystem of the x3690 X5 and the MAX5 is described.

4.12.1 x3690 X5 power subsystem

The x3690 X5 power subsystem consists of up to four hot-pluggable 675 W auto-sensing power supplies. The modules are independently powered by ac power cords.

Most standard models have one power supply as standard; workload-optimized models have more. See 4.3, “Models” on page 122 for details. One power supply is sufficient when the total power budget is less than 675 W. 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 675 W, installing a second power supply provides redundancy. To install a second power supply, use the IBM High Efficiency 675 W 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 675 W 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 4-37 shows the part numbers for the power supply options.

Table 4-37 IBM 675 W Redundant Power Supply Kit for x3690 X5

Option	Feature code	Description	Use
60Y0332	4782	IBM High Efficiency 675 W Power Supply	Power supply 2
60Y0327	Various ^a	IBM 675 W 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 675 W Redundant Power Supply Kit, option 60Y0327, includes the following items:

- ▶ Two 675 W power supplies
- ▶ Two Y-cord power cables (2.8 m, 10 A / 200 - 250 V, 2xC13 to IEC 320-C14)
- ▶ Two power cables (2.8 m, 10 A / 100 - 250 V, 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 board.

Table 4-38 lists the ac power input requirements.

Table 4-38 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	7.8 A root mean square (RMS)
High range	180 V ac	265 V ac	200 - 240 V ac 50/60 Hz	3.8 A RMS

4.12.2 MAX5 power subsystem

The MAX5 power subsystem consists of two hot-pluggable 675 W power supplies. The power subsystem is designed for N+N (fully redundant) operation and hot-swap replacement. MAX5 units have both power supplies installed as standard.

MAX5 has five redundant hot-swap fans, all within one cooling zone. The IMM of the attached host controls the MAX5 fan speed, which is based on altitude and ambient temperature. In addition, a fan that is located inside each power supply cools the power modules.

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 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. To enable the embedded hypervisor function, an internal USB connector on the x8 low profile PCI riser card (Figure 4-49) is reserved to support one USB flash drive. See Table 4-39 on page 179 for details.

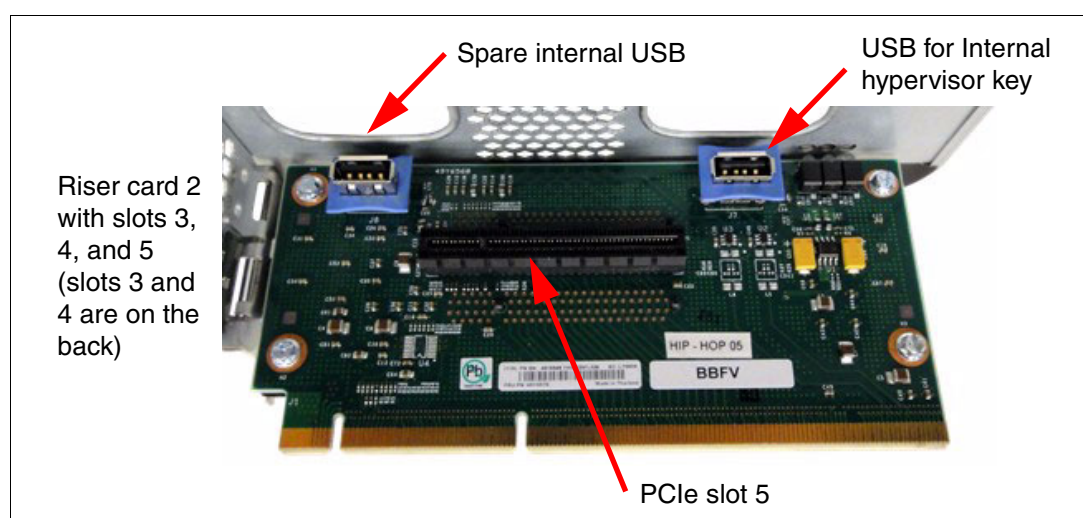


Figure 4-49 Low profile x8 riser card with hypervisor flash USB connector

The IBM USB Memory Key for virtualization is included in the virtualization-optimized models that are listed in 4.3, “Models” on page 122. However, it can be added to any x3690 X5 system.

For more information about the USB keys, and to download the IBM customized version of VMware ESX4.1 Update 1, visit the following web page:

<http://www.ibm.com/systems/x/os/vmware/esxi>

Table 4-39 shows the USB key options for the embedded hypervisor.

Table 4-39 USB key for the embedded hypervisor

Option	Feature code	Description
41Y8298	A2G0	IBM Blank USB Memory Key for VMware ESXi Downloads
41Y8296	A1NP	IBM USB Memory Key for VMware ESXi 4.1 Update 1
41Y8300	A2VC	IBM USB Memory Key for VMWare ESXi 5.0
41Y8307	A383	IBM USB Memory Key for VMware ESXi 5.0 Update1
41Y8311	A2R3	IBM USB Memory Key for VMWare ESXi 5.1

For more information and setup instructions for VMware ESXi software, see the *VMware ESXi Embedded and vCenter Server Setup Guide*, available at the following web page:

http://www.vmware.com/pdf/vsphere4/r41/vsp_41_esxi_e_vc_setup_guide.pdf

4.14 Supported operating systems

The following operating systems are supported by the x3690 X5:

- ▶ Microsoft Windows Server 2008 R2
- ▶ 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
- ▶ Microsoft Windows Server 2012
- ▶ Red Hat Enterprise Linux 5 Server with Xen x64 Edition
- ▶ Red Hat Enterprise Linux 5 Server 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 4.1
- ▶ VMware ESXi 4.0
- ▶ VMware ESXi 4.1
- ▶ VMware vSphere 5
- ▶ VMware vSphere 5.1

Check the ServerProven Operating System support page for the most up-to-date list of supported operating systems:

<http://www.ibm.com/servers/eserver/serverproven/compat/us/nos/matrix.shtml>

Notes:

- ▶ 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 web page:

<http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/matrix.shtml>

- ▶ MAX5 requires VMware ESX 4.1 or later.
- ▶ 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 or 1.75 inches). The MAX5 memory expansion unit is an extra 1U high unit. Both devices are 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



IBM BladeCenter HX5

The IBM BladeCenter HX5 blade server showcases the eX5 architecture and technology in a blade form factor. The server is introduced and its features and options are described.

The following topics are described:

- ▶ 5.1, “Introduction” on page 182
- ▶ 5.2, “Comparison between HS23 and HX5” on page 184
- ▶ 5.3, “Target workloads” on page 185
- ▶ 5.4, “Chassis support” on page 186
- ▶ 5.5, “HX5 models” on page 187
- ▶ 5.6, “System architecture” on page 192
- ▶ 5.7, “Speed Burst Card” on page 193
- ▶ 5.8, “IBM MAX5 and MAX5 V2 for HX5” on page 194
- ▶ 5.9, “Scalability” on page 196
- ▶ 5.10, “Processor options” on page 199
- ▶ 5.11, “Memory” on page 201
- ▶ 5.12, “Storage” on page 212
- ▶ 5.13, “BladeCenter PCI Express Gen 2 Expansion Blade II” on page 217
- ▶ 5.14, “I/O expansion cards” on page 219
- ▶ 5.15, “Standard onboard features” on page 222
- ▶ 5.16, “Integrated virtualization” on page 223
- ▶ 5.17, “Partitioning capabilities” on page 225
- ▶ 5.18, “Operating system support” on page 225

5.1 Introduction

The IBM BladeCenter HX5 supports up to two Intel Xeon E7 processors (machine type 7873). The HX5 supports up to 40 dual inline memory modules (DIMMs) with the addition of the MAX5 memory expansion blade.

Figure 5-1 shows the following three configurations:

- ▶ Single-wide HX5 with two processor sockets and 16 DIMM sockets.
- ▶ Two-node HX5 with a total of four processors and 32 DIMM sockets.
- ▶ An HX5 with an attached MAX5 memory expansion blade that contains two processors and a total of 40 DIMM sockets, 16 in the HX5 server and 24 in the MAX5.

MAX5 connectivity: MAX5 can connect only to a single HX5 server.

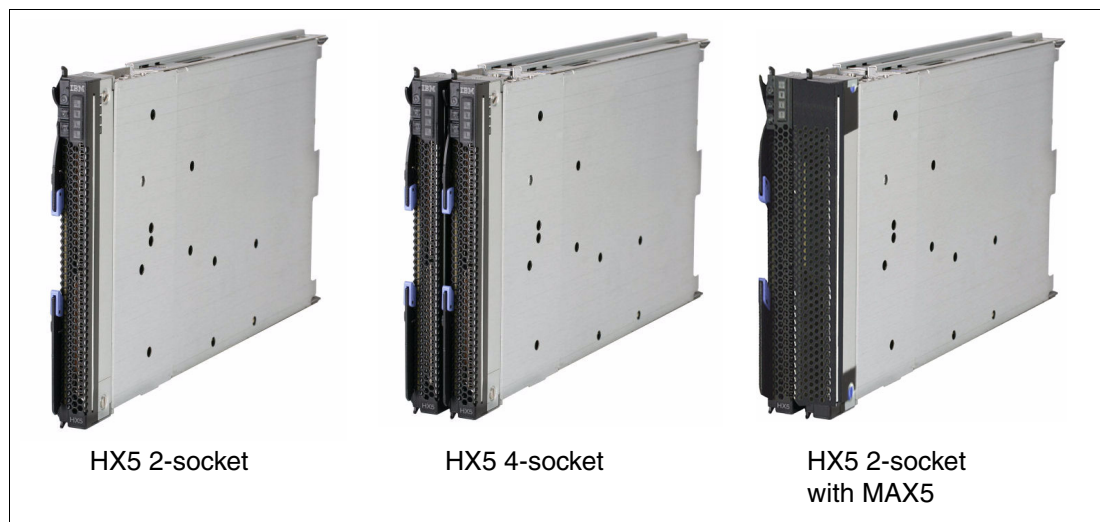


Figure 5-1 IBM BladeCenter HX5 blade server configurations

Table 5-1 lists the features of the HX5 with Intel Xeon E7 processors, machine type 7873.

Table 5-1 Features of the HX5 type 7873

Features	HX5 two-socket	HX5 four-socket	HX5 two-socket with MAX5
Form factor	30 mm (one-wide)	60 mm (two-wide)	60 mm (two-wide)
Maximum number of processors	Two	Four	Two
Processor options	Intel Xeon E7-8800, E7-4800 and E7-2800: Six-core, eight-core, and ten-core	Intel Xeon E7-8800 and E7-4800: Six-core, eight-core, and ten-core	Intel Xeon E7-8800, E7-4800 and E7-2800: Six-core, eight-core, and ten-core
Cache	18 MB, 24 MB, or 30 MB shared between cores (processor-dependent)		
Memory speed	1066, 978, or 800 MHz (processor SMI link speed-dependant)		
Memory (DIMM slots/maximum)	16 DIMM slots/256 GB, maximum of 16 GB DIMMs	32 DIMM slots/512 GB, maximum of 16 GB DIMMs	40 DIMM slots/640 GB, maximum of 16 GB DIMMs
Memory type	DDR3 error correction code (ECC) Very Low Profile (VLP) Registered DIMMs		
Internal storage	Optional 1.8-inch solid-state drives (SSDs); non-hot-swap (require an extra SSD carrier)		

Features	HX5 two-socket	HX5 four-socket	HX5 two-socket with MAX5
Maximum number of drives	Two	Four	Two
Maximum internal storage	Up to 800 GB by using two 400 GB SSDs	Up to 1.6 TB by using four 400 GB SSDs	Up to 800 GB by using two 400 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.

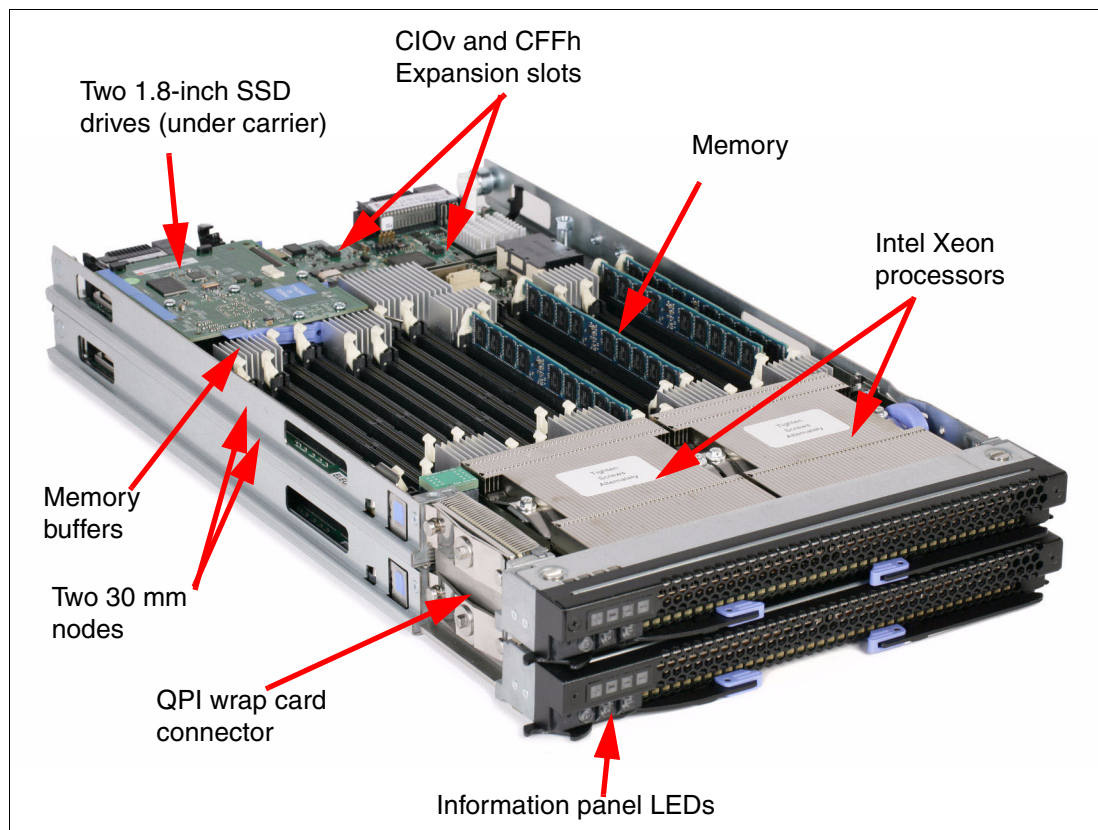


Figure 5-2 Layout of HX5 (showing a two-node, four-socket configuration)

The MAX5 memory expansion blade, which is shown in Figure 5-3, 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.

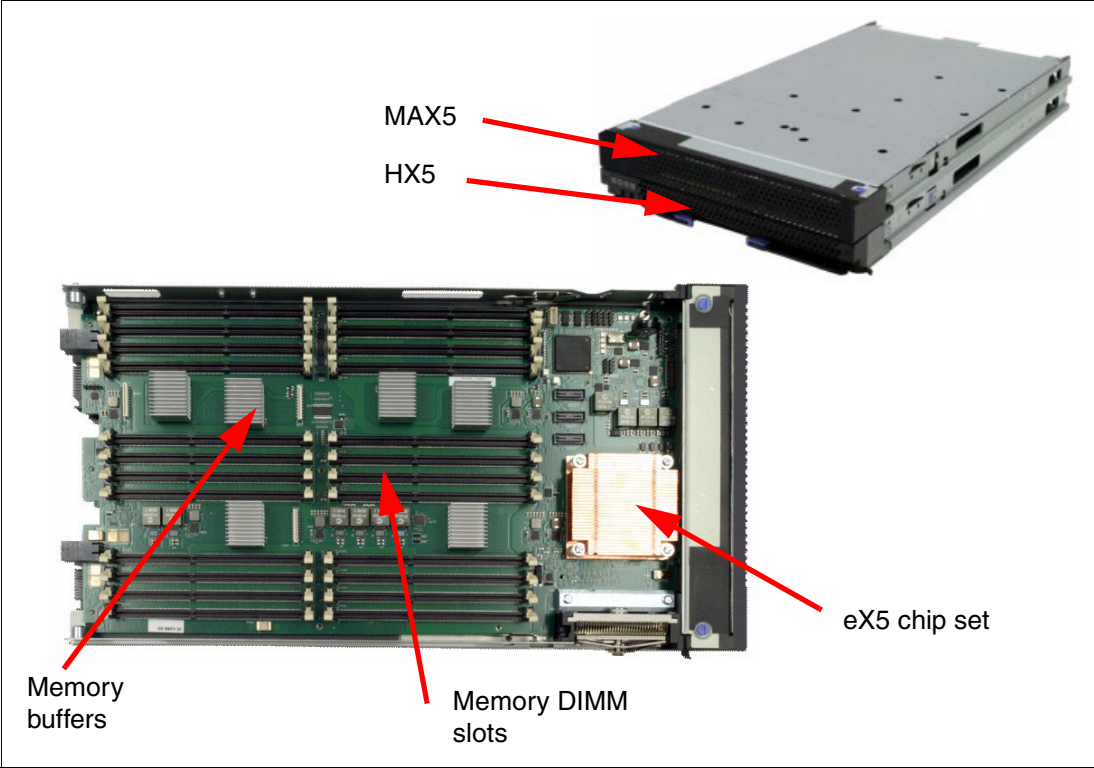


Figure 5-3 MAX5 for IBM BladeCenter

5.2 Comparison between HS23 and HX5

The BladeCenter HS23 is a general-purpose, two-socket blade server. Table 5-2 compares the HS23 with the HX5 offerings.

Table 5-2 HX5 compared to HS23

Feature	HS23	HX5	HX5 with MAX5
Form factor	30 mm blade (one-wide)	30 mm blade (one-wide) 60 mm blade (two-wide)	60 mm blade (two-wide)
Processor	Intel Xeon E5 processor	Intel Xeon E7 processor	Intel Xeon E7 processor
Maximum number of processors	Two	30 mm blade: two 60 mm blade: four	Two
Number of cores	Four, six, or eight cores	Six, eight, or ten cores	Six, eight, or ten cores
Cache	10 MB, 15 MB, or 20 MB	12 MB, 18 MB, 24 MB, or 30 MB	
Memory speed	Up to 1600 MHz	Up to 1066 MHz	Up to 1066 MHz
DIMMs per channel	Two	One	HX5: One MAX5: Two

Feature	HS23	HX5	HX5 with MAX5
DIMM sockets	16	30 mm: 16 60 mm: 32	40
Maximum installable RAM	512	30 mm: 512 GB 60 mm: 1 TB	1.25 TB
Memory type	DDR3 ECC VLP RDIMMs	DDR3 ECC VLP RDIMMs	DDR3 ECC VLP RDIMMs
Internal disk drives	Two hot-swap 2.5-inch drive (SAS, SATA, or SSD)	Two or four non-hot-swap 1.8-inch SSDs (requires the SSD Expansion Card)	Two non-hot-swap 1.8-inch SSDs (requires the SSD Expansion Card)
I/O expansion	▶ One CIOv ▶ One CFFh	Per 30 mm blade: ▶ One CIOv ▶ One CFFh	Per 60 mm blade: ▶ One CIOv ▶ One CFFh
Serial-attached SCSI (SAS) controller	Onboard LSI SAS2004	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
Onboard Ethernet	Emulex BE3	Broadcom 5709S	Broadcom 5709S
Chassis supported ^a	▶ BladeCenter E ▶ BladeCenter H ▶ BladeCenter S ▶ BladeCenter HT	▶ BladeCenter H ▶ BladeCenter S ▶ BladeCenter HT	

a. Important: There might be restrictions on the number of blades, and required types of power supplies and cooling modules to support these blades in the chassis listed. For BladeCenter E, the advanced management module (AMM) is required. For more information, see the IBM BladeCenter Interoperability Guide at this website:

<http://www.redbooks.ibm.com/big>

5.3 Target workloads

The HX5 is designed for business-critical workloads, such as databases and virtualization.

Virtualization provides many benefits, including improved physical resource usage, 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 must be tracked by a company or department.

Virtualization and server consolidation can provide the following benefits:

- ▶ Reducing the rate of physical server proliferation
- ▶ Simplifying infrastructure
- ▶ Improving manageability
- ▶ Lowering the total cost of IT, including power and cooling costs

The HX5 two-socket and HX5 four-socket are strong database systems. They are ideal upgrade candidates for database workloads already on a blade. The multi-core 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. In-memory databases such as SolidDB and SAP HANA are viable options given the memory capacity of these servers.

5.4 Chassis support

The HX5 is supported in BladeCenter chassis S, H, and HT, as listed in Table 5-3.

Table 5-3 HX5 compatibility with BladeCenter (BC) chassis

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	Yes
HX5+MAX5 server	No	Yes	Yes ^a	Yes	Yes

a. One-node and two-node HX5 configurations with 130 W processors are not supported in chassis with standard cooling modules. See Table 5-4.

The number of HX5 servers that are supported in each chassis depends on the thermal design power of the processors that are used in the HX5 servers. Table 5-4 lists the HX5 servers, using the following conventions:

- ▶ A green 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 BladeCenter H).
- ▶ A yellow 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 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	CPU TDP ^a	Maximum number of HX5 supported in each chassis						
		BC-S (8886) (6 bays)	BC-H (14 bays)				BC-HT DC (8740) (12 bays) ^b	BC-HT AC (8750) (12 bays) ^b
			2900 W power supplies		2980 W power supplies ^c			
			Std. blower	Enh. blower ^d	Std. blower	Enh. blower ^d		
HX5 single-wide	130 W	4	None ^e	10	None ^e	12	6	8
	105 W	5	14	14	14	14	10	10
	95 W	5	14	14	14	14	10	10
HX5 + MAX5 double-wide	130 W	2	6	6	7	7	4	5
	105 W	2	7	7	7	7	5	5
	95 W	2	7	7	7	7	5	5

a. TDP = Thermal Design Power.

b. Support shown is for non-Network Equipment Building System (NEBS) (Enterprise) environments.

c. IBM BladeCenter H 2980 W AC Power Modules, 68Y6601 (standard in 4Tx, 5Tx, and 9xx BC-H chassis models; optional with all other BC-H chassis models).

d. IBM BladeCenter H Enhanced Cooling Modules, 68Y6650 (standard in 4Tx, 5Tx, and 9xx BC-H chassis models; optional with all other BC-H chassis models).

e. Not supported.

5.5 HX5 models

The available HX5 models are described and are grouped by the following types:

- ▶ 5.5.1, “Base models of machine type 7873” on page 187
- ▶ 5.5.2, “Two-node models of machine type 7873” on page 189
- ▶ 5.5.3, “Workload optimized models of machine type 7873” on page 190

5.5.1 Base models of machine type 7873

Table 5-5 shows the base models of the BladeCenter HX5 type 7873 (with Intel Xeon E7 processors), with and without the MAX5 memory expansion blade.

If the MAX5 is attached, you cannot also attach the two-node scalability kit to form a two-node configuration. The reverse is also true: forming a two-node configuration precludes the use of the MAX5. Models with E7-2800 series processors do not support forming a two-node configuration.

Table 5-5 Base models of HX5 type 7873

Model	Processor (qty, model, cores, core speed, L3 cache, memory speed) (two max)	MAX5 ^a	Two-node ^a	Standard memory	Memory speed ^b	Standard networking ^c	Storage
Models with optional MAX5							
7873-B1x	1x E7-4807 6C 1.86 GHz 18 MB 800 MHz 95 W	Opt	Opt	2x 4 GB	800 MHz	2x 1 Gb	Optional
7873-B2x	1x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	Opt	2x 4 GB	1066 MHz	2x 1 Gb	Optional
7873-B3U ^d	1x E7-4850 10C 2.00 GHz 24 MB 1066 MHz 130 W	Opt	Opt	2x 8 GB	1066 MHz	2x 1 Gb	Optional
7873-C1x	1x E7-8837 8C 2.67 GHz 24 MB 1066 MHz 130 W	Opt	Opt	2x 4 GB	978 MHz	2x 1 Gb	Optional
7873-D1x	1x E7-8867L 10C 2.13 GHz 30 MB 1066 MHz 105 W	Opt	Opt	2x 4 GB	1066 MHz	2x 1 Gb	Optional
7873-E1x	2x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	Opt	16x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E1)	Optional
7873-E3x	2x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	Opt	16x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-F1x	1x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	Opt	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E1)	Optional
7873-F2x	1x E7-4870 10C 2.40 GHz 30 MB 1066 MHz 130 W	Opt	Opt	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E1)	Optional
7873-F4x	1x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	Opt	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-F5x	1x E7-4870 10C 2.40 GHz 30 MB 1066 MHz 130 W	Opt	Opt	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-H1x	1x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	Opt	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional

Model	Processor (qty, model, cores, core speed, L3 cache, memory speed) (two max)	MAX5 ^a	Two-node ^a	Standard memory	Memory speed ^b	Standard networking ^c	Storage
Models with standard MAX5							
7873-A1x	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Std	NS	4x 4 GB	1066 MHz	2x 1 Gb	Optional
7873-A2x	2x E7-2860 10C 2.26 GHz 24 MB 1066 MHz 130 W	Std	NS	4x 4 GB	1066 MHz	2x 1 Gb	Optional
7873-A3x	2x E7-2870 10C 2.40 GHz 30 MB 1066 MHz 130 W	Std	NS	4x 4 GB	1066 MHz	2x 1 Gb	Optional
7873-E2x	2x E7-2860 10C 2.26 GHz 24 MB 1066 MHz 130 W	Std	NS	HX5: 16x 8 GB MAX5: 12x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E1)	Optional
7873-E4x	2x E7-2860 10C 2.26 GHz 24 MB 1066 MHz 130 W	Std	NS	HX5: 16x 8 GB MAX5: 12x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-F3x	2x E7-4807 6C 1.86 GHz 18 MB 800 MHz 95 W	Std	NS	2x 4 GB	800 MHz	2x 1 Gb + 2x 10 Gb (E1)	Optional
7873-F6x	2x E7-4807 6C 1.86 GHz 18 MB 800 MHz 95 W	Std	NS	4x 4 GB	800 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-G1x	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Std	NS	40x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E1)	Optional
7873-H2x	1x E7-4870 10C 2.40 GHz 30 MB 1066 MHz 130 W	Std	NS	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-H3x	2x E7-4807 6C 1.86 GHz 18 MB 800 MHz 95 W	Std	NS	4x 4 GB	800 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional

a. The HX5 supports either a MAX5 or can be expanded to two nodes through the two-node scalability kit. However, both the MAX5 and two-node scalability are not supported at the same time. Several models have the MAX5 standard (88Y6128), and other models have the two-node scalability kit standard (46M6975). Models with E7-2800 series processors do not support two-node configurations.

b. With Xeon E7 processors, the memory speed in the HX5 and the MAX5 are the same.

c. All models contain an onboard two-port Gigabit Ethernet controller. Several models also include an extra 10 Gb Expansion Card that is installed in the CFFh expansion slot, as follows:

- (E1) Emulex 10 GbE Virtual Fabric Adapter Advanced
- (E2) Emulex 10 GbE Virtual Fabric Adapter Advanced II

d. Model B3U is available in the US only.

5.5.2 Two-node models of machine type 7873

Table 5-6 lists Intel Xeon E7-based models that are used in a two-node configuration. For these models, order one model with the two-node scalability kit and another with the same processor without the scalability kit. For example, order model 7873-BAx and 7873-BHx together. These models do not support the use of a MAX5.

Table 5-6 Two-node models of HX5 type 7873

Model	Processor (qty, model, cores, core speed, L3 cache, memory speed) (two max)	MAX5 ^a	Two node ^a	Standard memory	Memory speed ^b	Standard networking ^c	Storage
Models for two-node configurations							
7873-BAx	1x E7-4807 6C 1.86 GHz 18 MB 800 MHz 95 W	NS	Std ^d	2x 4 GB	800 MHz	2x 1 Gb	Optional
7873-BHx	1x E7-4807 6C 1.86 GHz 18 MB 800 MHz 95 W	NS	Connect to BAx	2x 4 GB	800 MHz	2x 1 Gb	Optional
7873-BBx	1x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	NS	Std ^d	2x 4 GB	1066 MHz	2x 1 Gb	Optional
7873-BJx	1x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	NS	Connect to BBx	2x 4 GB	1066 MHz	2x 1 Gb	Optional
7873-CAx	1x E7-8837 8C 2.67 GHz 24 MB 1066 MHz 130 W	NS	Std ^d	2x 4 GB	978 MHz	2x 1 Gb	Optional
7873-CHx	1x E7-8837 8C 2.67 GHz 24 MB 1066 MHz 130 W	NS	Connect to CAx	2x 4 GB	978 MHz	2x 1 Gb	Optional
7873-DAx	1x E7-8867L 10C 2.13 GHz 30 MB 1066 MHz 105 W	NS	Std ^d	2x 4 GB	1066 MHz	2x 1 Gb	Optional
7873-DHx	1x E7-8867L 10C 2.13 GHz 30 MB 1066 MHz 105 W	NS	Connect to DAx	2x 4 GB	1066 MHz	2x 1 Gb	Optional
7873-FAx	1x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	NS	Std ^d	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E1)	Optional
7873-FHx	1x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	NS	Connect to FAx	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E1)	Optional
7873-FDx	1x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	NS	Std ^d	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-FMx	1x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	NS	Connect to FDx	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-FEx	1x E7-4870 10C 2.40 GHz 30 MB 1066 MHz 130 W	NS	Std ^d	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-FNx	1x E7-4870 10C 2.40 GHz 30 MB 1066 MHz 130 W	NS	Connect to FEx	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-FBx	1x E7-4870 10C 2.40 GHz 30 MB 1066 MHz 130 W	NS	Std ^d	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E1)	Optional
7873-FJx	1x E7-4870 10C 2.40 GHz 30 MB 1066 MHz 130 W	NS	Connect to FBx	2x 4 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E1)	Optional

- a. The HX5 supports either a MAX5 or can be expanded to two nodes through the two-node scalability kit. However, both the MAX5 and two-node scalability are not supported at the same time. Several models have the MAX5 standard (88Y6128) and other models have the two-node scalability kit standard (46M6975). Models with E7-2800 series processors do not support two-node configurations.
- b. With Xeon E7 processors, the memory speed in the HX5 and the MAX5 are the same.
- c. All models contain an onboard two-port Gigabit Ethernet controller. Several models also include an extra 10 Gb Expansion Card that is installed in the CFFh expansion slot, as follows:
 - (E1) Emulex 10 GbE Virtual Fabric Adapter Advanced
 - (E2) Emulex 10 GbE Virtual Fabric Adapter Advanced II
- d. The two-node scalability kit, 46M6975, is standard with these models to enable two-node scalability, as described in 5.9.2, “Two-node HX5 configuration” on page 197. Order the partner model with the same processor as listed to build a two-node configuration.

5.5.3 Workload optimized models of machine type 7873

Table 5-7 on page 191 lists the workload-optimized models of the HX5. These models are pre-configured and pre-tested models that are targeted at specific workloads.

Database models

The 7873-G2x and G4x that are listed in Table 5-7 on page 191 are transactional database-optimized models and include the following features in addition to standard HX5 features:

- ▶ Eight 8 GB memory DIMMs for a total of 64 GB of available memory.
- ▶ One BladeCenter PCIe Gen 2 Expansion Blade. See 5.13, “BladeCenter PCI Express Gen 2 Expansion Blade II” on page 217, for details.
- ▶ Two IBM 320 GB High IOPS SD Class SSD PCIe Adapters (PCI Express (PCIe) form factor) installed in the BladeCenter PCIe Gen 2 Expansion Blade.
- ▶ One dual port Emulex 10 GbE Virtual Fabric Adapter Advanced or Emulex 10 GbE Virtual Fabric Adapter Advanced II.

BladeCenter Foundations for Cloud models

The 7873-9xx models are part of an IBM BladeCenter Foundation for Cloud configuration. IBM BladeCenter Foundation for Cloud provides a comprehensive, converged solution that brings together the hardware, software, and services that are needed to quickly establish a robust virtualized environment. With the addition of select software, IBM BladeCenter Foundation for Cloud can easily be extended to a private cloud environment.

For details about IBM BladeCenter Foundation for Cloud, see the following web page:

<http://ibm.com/systems/bladecenter/solutions/virtualization/integratedcloudplatform>

zBX models

Table 5-7 also lists the models that can be installed in the IBM zEnterprise® BladeCenter Extension (zBX) or within a traditional IBM BladeCenter chassis. These HX5 models are configured with Fibre Channel and Ethernet Networking options, making them easy to order, configure, and deploy.

For more information about zBX, see Chapter 4 of *IBM BladeCenter Products and Technology*, SG24-7523, available at the following web page:

<http://www.redbooks.ibm.com/abstracts/sg247523.html>

Table 5-7 Workload-optimized models of the HX5 type 7873

Model	Processor (qty, model, core speed, cores, L3 cache, memory speed) (2 max)	MAX5 ^a	Two node ^a	Standard memory	Memory speed ^b	Standard networking and HBAs ^c	Storage
WOS database models (standard PCIe SSD adapters in IBM BladeCenter PCIe Gen 2 Expansion Blade)							
7873-G2x	2x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	Opt	8x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E1)	2x 320 GB PCIe SSD ^d
7873-G4x	2x E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	Opt	8x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	2x 320 GB PCIe SSD ^d
WOS models for IBM BladeCenter Foundation for Cloud (optional or standard MAX5)							
7873-91x	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	NS	16x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-92x	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	NS	16x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (Q)	Optional
7873-93x	2x E7-8867L 10C 2.13 GHz 30 MB 1066 MHz 105 W	Std	NS	40x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (E2)	Optional
7873-94x	2x E7-8867L 10C 2.13 GHz 30 MB 1066 MHz 105 W	Std	NS	40x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (Q)	Optional
WOS models for zEnterprise BladeCenter Extension (zBX)							
7873-A4x	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	NS	8x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (B)+ 2x 8 Gb FC	2x 50 GB MLC SSD ^e
7873-A5x	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	NS	16x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (B)+ 2x 8 Gb FC	2x 50 GB MLC SSD ^e
7873-A6x	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	NS	8x 8 GB + 8x 16 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (B)+ 2x 8 Gb FC	2x 50 GB MLC SSD ^e
7873-A7x	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	NS	16x 16 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (B)+ 2x 8 Gb FC	2x 50 GB MLC SSD ^e
7873-AAx	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	NS	8x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (B2) 2x 8 Gb FC	2x 100 GB MLC SSD ^f

Model	Processor (qty, model, core speed, cores, L3 cache, memory speed) (2 max)	MAX5 ^a	Two node ^a	Standard memory	Memory speed ^b	Standard networking and HBAs ^c	Storage
7873-ABx	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	NS	16x 8 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (B2) 2x 8 Gb FC	2x 100 GB MLC SSD ^f
7873-ACx	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	NS	8x 8 GB+ 8x 16 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (B2) 2x 8 Gb FC	2x 100 GB MLC SSD ^f
7873-ADx	2x E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Opt	NS	16x 16 GB	1066 MHz	2x 1 Gb + 2x 10 Gb (B2) 2x 8 Gb FC	2x 100 GB MLC SSD ^f

a. The HX5 supports either a MAX5 or can be expanded to two nodes with the two-node scalability kit. However, both the MAX5 and two-node scalability are not supported at the same time. Several models have the MAX5 standard (88Y6128) and other models have the two-node scalability kit standard (46M6975). Models with E7-2800 series processors do not support two-node configurations.

b. With Xeon E7 processors, the memory speed in the HX5 and the MAX5 are the same.

c. All models contain an onboard two-port Gigabit Ethernet controller. Several models also include an extra 10 Gb Expansion Card that is installed in the CFFh expansion slot, as follows:

- (B) Broadcom 10 Gb Gen2 two-port Ethernet Expansion Card (CFFh)
- (B2) Broadcom 2-port 10 Gb Virtual Fabric Adapter for IBM BladeCenter
- (E1) Emulex 10 GbE Virtual Fabric Adapter Advanced
- (E2) Emulex 10 GbE Virtual Fabric Adapter Advanced II
- (Q) QLogic two-port 10 Gb Converged Network Adapter (CFFh)

d. Model 7873-G2x and G4x include the 30 mm IBM BladeCenter PCIe Gen 2 Expansion Blade. The combined server is 60 mm wide (double-wide) and occupies two blade bays in the chassis. The Expansion Blade contains two IBM 320 GB High IOPS SD Class SSD PCIe Adapters.

e. Models 7873-A4x, A5x, A6x, and A7x include two IBM 50 GB SATA 1.8" MLC solid-state drives (SSDs) plus the SSD Expansion Card for IBM BladeCenter HX5.

f. Models 7873-AAx, ABx, ACx, and ADx include two IBM 100 GB SATA 1.8" MLC Enterprise SSDs plus the SSD Expansion Card for IBM BladeCenter HX5.

5.6 System architecture

The Intel Xeon E7 processors in the HX5, machine type 7873, contain up to ten cores with up to 30 MB shared L3 cache. Common technologies include Hyper-Threading, several with Turbo Boost, four QuickPath Interconnect (QPI) links, and an integrated memory controller with four buffered system memory interconnect (SMI) for eight memory channels per processor.

The HX5 two-socket server has the following system architecture features as standard:

- ▶ Two 1567-pin land grid array (LGA) processor sockets
- ▶ Intel 7510 "Boxboro" I/O Hub
- ▶ Intel ICH10 south bridge
- ▶ Eight Intel Scalable Memory Buffers, each with two memory channels
- ▶ One DIMM per memory channel
- ▶ Sixteen 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 provides a block diagram of the X5 functional components.

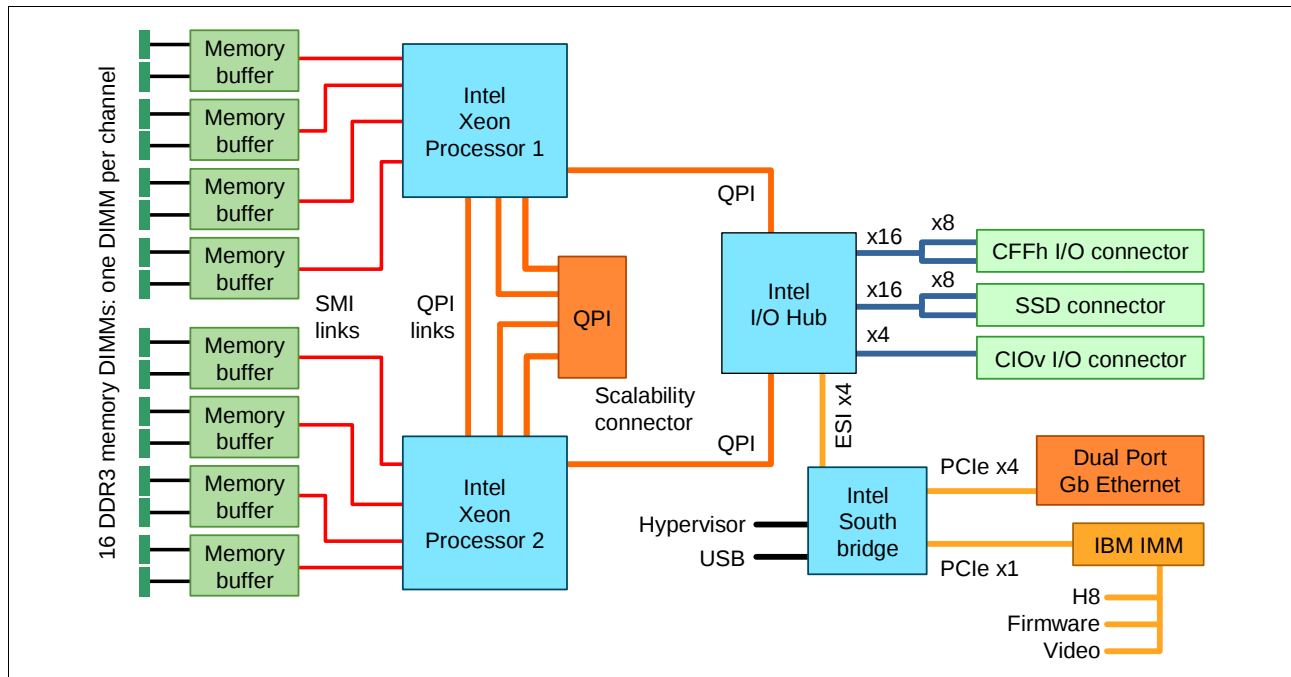


Figure 5-4 HX5 block diagram

5.7 Speed Burst Card

To increase performance in a two-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 two-socket blades or a MAX5 and routes them back to the processors on the same blade. Table 5-8 lists the ordering information.

Table 5-8 HX5 1-Node Speed Burst Card

Part number	Feature code	Description
59Y5889	1741	IBM HX5 1-Node Speed Burst Card

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 two-node configuration. A separate card is available for a two-node configuration, as described in 5.9, “Scalability” on page 196.

Figure 5-5 shows a block diagram of the Speed Burst Card attachment to the system.

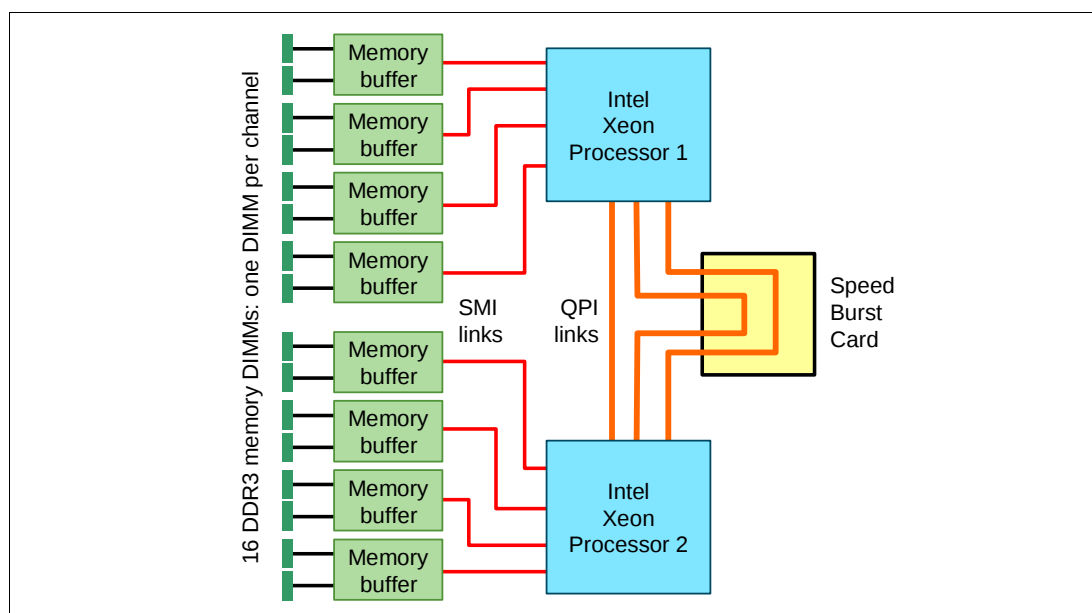


Figure 5-5 HX5 1-Node Speed Burst Card block diagram

Figure 5-6 shows where the Speed Burst Card is installed on the HX5.

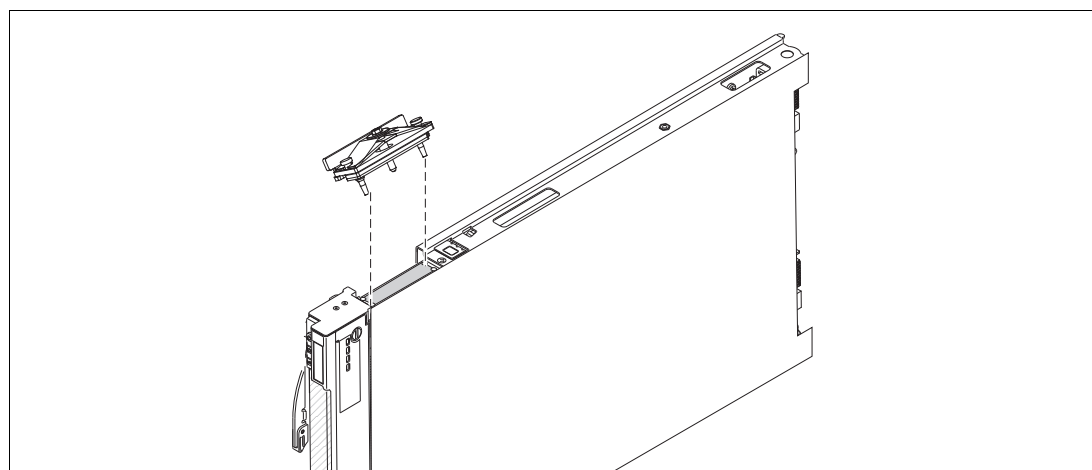


Figure 5-6 Installing the Speed Burst Card

5.8 IBM MAX5 and MAX5 V2 for HX5

The IBM MAX5 for BladeCenter (Figure 5-3 on page 184) is a memory expansion blade that attaches to a single HX5 two-socket blade server.

There are two versions of the MAX5:

- ▶ The original MAX5 that was released with the first HX5 machine type 7872 in March 2010. The part number is 46M6973.
- ▶ MAX5 V2, which was released in May 2011 along with the HX5 machine type 7873. The part number is 88Y6128.

MAX5 compatibility is shown in Table 5-9.

Table 5-9 MAX5 compatibility

	HX5 with Intel Xeon E7 (machine type 7873)
MAX5, 46M6973	Supported ^a
MAX5 V2, 88Y6128	Supported ^a

a. Connecting a MAX5 or MAX5 V2 is not supported by servers that have the E7-2820 or E7-2803 processors installed.

The MAX5 and MAX5 V2 have the following system architecture features:

- ▶ IBM EXA memory controller
- ▶ 24 DIMM slots/six memory buffers
- ▶ Four DIMM slots per memory buffer (two per channel)
- ▶ Support for VLP DDR3 memory
- ▶ Ability to attach to a single HX5 by using the IBM HX5 MAX5 1-Node Scalability Kit, 59Y5877, as described in “HX5 with MAX5” on page 198
- ▶ Communication with the processors on the HX5 by using high-speed QPI links
- ▶ Support for low voltage (1.35 V) DIMM modules (operate at 1.5 V in MAX5)

The MAX5 is standard with certain models (Table 5-5 on page 187), and as an option for other models (see Table 5-10).

MAX5 is standard with certain models, as listed in 5.5, “HX5 models” on page 187. For other models, MAX5 is available as an option, as listed in Table 5-10.

Table 5-10 IBM MAX5 for BladeCenter

Part number	Feature code	Description
88Y6128	A16N	IBM MAX5 V2 for BladeCenter
46M6973	1740	IBM MAX5 for BladeCenter
59Y5877	1742	IBM HX5 MAX5 1-Node Scalability Kit

Table 5-11 lists the memory DIMM options compatibility information for both the MAX5 and MAX5 V2.

Table 5-11 Memory options for the MAX5 and MAX5 V2

Part number	Feature code	Description	Supports MAX5 46M6973	Supports MAX5 V2 88Y6128
44T1596	1908	4 GB PC3-10600 CL9 ECC VLP (2Rx8, 1.5 V, 2 Gbit)	Yes	No
46C7499	1917	8 GB PC3-8500 CL7 ECC VLP (4Rx8, 1.5 V, 2 Gbit)	Yes	No
46C0560	A0WX	2 GB (1x2 GB, 1Rx8, 1.35 V) PC3-10600 CL9 ECC DDR3 1333 MHz VLP RDIMM	No	Yes
46C0564	A0WZ	4 GB (1x 4 GB, 2Rx8, 1.35 V) PC3-10600 CL9 ECC DDR3 1333 MHz VLP RDIMM	No	Yes
46C0570	A17Q	8 GB (1x8 GB, 4Rx8, 1.35 V) PC3-8500 CL7 ECC DDR3 1066 MHz VLP RDIMM	No	Yes
46C0599	2422	16 GB (1x16 GB, 4Rx8, 1.35 V) PC3-10600 CL9 ECC DDR3 1066 MHz VLP RDIMM	No	Yes
00D5008	A3KN	32 GB (1x32 GB, 4Rx4, 1.35 V) PC3L-10600 CL9 ECC DDR3 1333 MHz VLP RDIMM	No	Yes

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. See Figure 5-14 on page 207 for a diagram of these power domains.

5.9 Scalability

We now describe how the MAX5 connects to the HX5 and how the HX5 can be expanded to increase the number of processors and the number of memory DIMMs. The memory options and rules are explained in 5.11, “Memory” on page 201.

The HX5 blade architecture allows for a number of scalable configurations, including the use of a MAX5 memory expansion blade. However, the blade currently supports three configurations:

- ▶ A single HX5 server with two processor sockets. This server is a standard 30 mm blade, which is known as a *single-wide* server or *single-node* server.
- ▶ Two HX5 servers that are connected to form a single image four-socket server. This server is a 60 mm blade, which is known as a *double-wide* server or *two-node* server.
- ▶ A single HX5 server with two processor sockets, plus a MAX5 memory expansion blade that is attached to it, resulting in a 60 mm blade configuration. This configuration is sometimes referred to as a *1-node+MAX5* configuration.

Each configuration is described in the following sections. The supported BladeCenter chassis for each configuration is listed in 5.4, “Chassis support” on page 186.

The HX5 also supports the attachment of one or more of the IBM BladeCenter PCIe Gen 2 Expansion Blade units. Additionally, the HX5 supports adding a MAX5 for forming a two-node configuration. See 5.13, “BladeCenter PCI Express Gen 2 Expansion Blade II” on page 217 for details.

5.9.1 Single HX5 configuration

The single-node 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 that are installed, ensure that the server has the Speed Burst Card installed for maximum performance, as described in 5.7, “Speed Burst Card” on page 193. This card is not required but is important for top performance.

5.9.2 Two-node HX5 configuration

In the two-node configuration, the two HX5 servers are physically connected and a two-node scalability card is attached to the side of the blades. This configuration provides the path for the QPI scaling.

Each node can have one or two processors installed (that is, two-node configurations with a total of two processors each, or four processors). All installed processors must be identical, however.

The two servers are connected by using a two-node scalability card, as shown in Figure 5-7. The scalability card is immediately next to the processors and provides a direct connection between the processors in the two nodes.

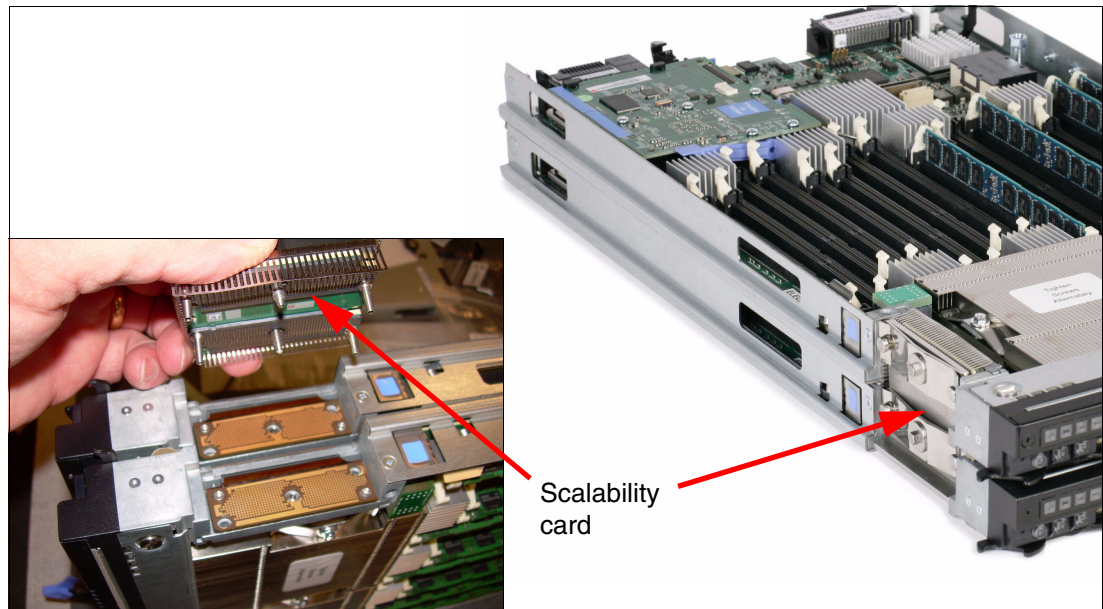


Figure 5-7 Two-node HX5 with two-node scalability card indicated

The two-node configuration consists of two connected HX5 servers. This configuration uses two blade slots and has the two-node scalability card attached. The scaling is done through QPI scaling. The two-node scalability card is not included with the server and must be ordered separately, as listed in Table 5-12.

Table 5-12 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 two-node scalability card, plus the necessary hardware to physically attach the two HX5 servers to each other.

Figure 5-8 shows the functional configuration of a two-node HX5 and the location of the HX5 two-node scalability card.

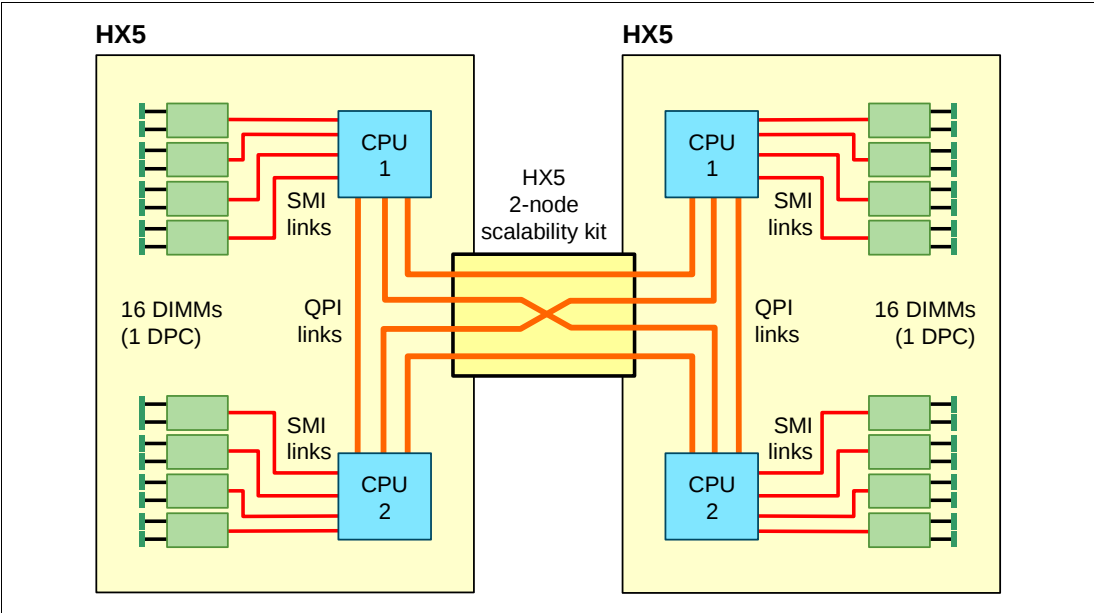


Figure 5-8 Block diagram of a two-node HX5

5.9.3 HX5 with MAX5

In the HX5 and MAX5 configuration, the HX5 and MAX5 units connect through a one-node MAX5 scalability card, which provides QPI scaling. See Figure 5-9.



Figure 5-9 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. This kit is similar in physical appearance to the 2-Node Scalability Kit that was shown in Figure 5-7 on page 197. Table 5-13 lists the ordering information.

Table 5-13 HX5 1-Node Scalability Kit

Part number	Feature code	Description
59Y5877	1742	IBM HX5 MAX5 1-Node Scalability Kit

Figure 5-10 shows the block diagram that depicts the configuration of the single-node HX5 with MAX5.

Single node: The MAX5 can be connected only to a single HX5 server. A configuration of two MAX5 units that are connected to a two-node HX5 is not supported.

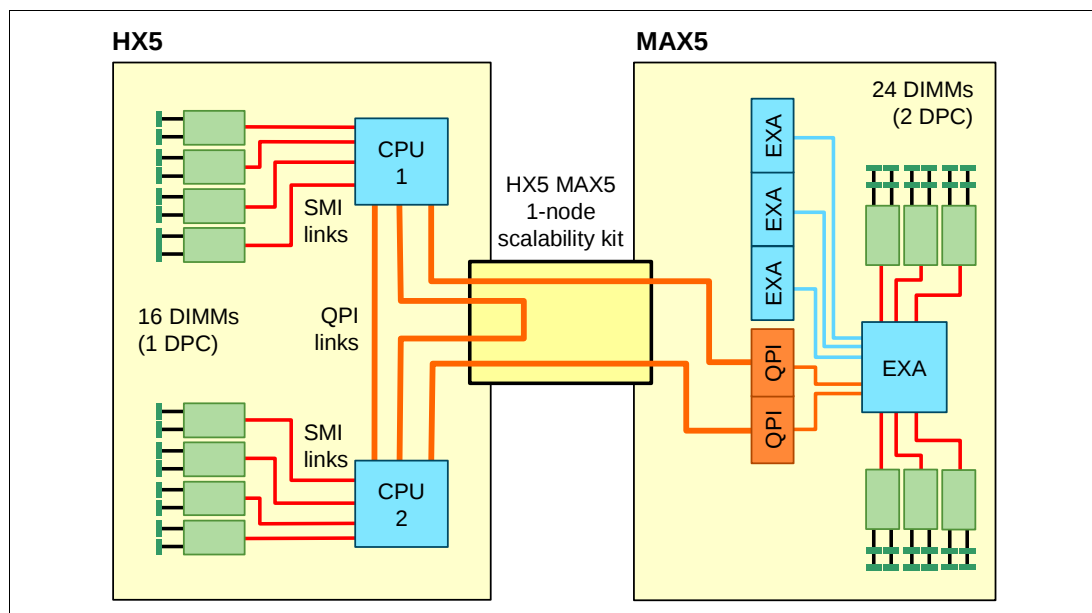


Figure 5-10 HX5 one-node with MAX5 block diagram

Having only one processor that is installed in the HX5 instead of two processors is supported. However, having two processors maximizes memory performance.

When you insert an HX5 with MAX5, there is no partition information to set up in a scalable complex. MAX5 ships ready to use when attached.

5.10 Processor options

The HX5 type 7873 supports Intel Xeon E7-8800, E7-4800, and E7-2800 six core, eight core, or ten-core processors. The processors must be identical in both two-socket and four-socket configurations. The Intel Xeon processors are available in various clock speeds and have standard and lower power offerings.

Table 5-14 lists the processor options for the HX5 and the models that include them (if they exist).

Table 5-14 Available processor options

Part number	Feature code	Description (Processor model, core, processor frequency, L3 cache, memory speed, power)	Can scale to two-node	Supported model
88Y6124	A17P	Xeon E7-8867L 10C 2.13 GHz 30 MB 1066 MHz 105 W	Yes	D1x
88Y6112	A17M	Xeon E7-8837 8C 2.67 GHz 24 MB 1066 MHz 130 W	Yes	C1x
88Y6160	A18W	Xeon E7-4870 10C 2.40 GHz 30 MB 1066 MHz 130 W	Yes	F2x

Part number	Feature code	Description (Processor model, core, processor frequency, L3 cache, memory speed, power)	Can scale to two-node	Supported model
88Y6102	A17K	Xeon E7-4860 10C 2.26 GHz 24 MB 1066 MHz 130 W	Yes	CTO ^a
88Y6092	A17H	Xeon E7-4850 10C 2.00 GHz 24 MB 1066 MHz 130 W	Yes	CTO ^a
88Y6082	A17G	Xeon E7-4830 8C 2.13 GHz 24 MB 1066 MHz 105 W	Yes	B2x, F1x, G2x
88Y6076	A17F	Xeon E7-4820 8C 2.00 GHz 18 MB 978 MHz 105 W	Yes	CTO ^a
88Y6070	A17E	Xeon E7-4807 6C 1.86 GHz 18 MB 800 MHz 95 W	Yes	B1x, F3x
88Y6150	A18U	Xeon E7-2870 10C 2.40 GHz 30 MB 1066 MHz 130 W	No	A3x
69Y3094	A17C	Xeon E7-2860 10C 2.26 GHz 24 MB 1066 MHz 130 W	No	A2x
69Y3084	A17A	Xeon E7-2850 10C 2.00 GHz 24 MB 1066 MHz 130 W	No	CTO ^a
69Y3074	A179	Xeon E7-2830 8C 2.13 GHz 24 MB 1066 MHz 105 W	No	A1x
69Y3068	A178	Xeon E7-2820 8C 2.00 GHz 18 MB 978 MHz 105 W	No	CTO ^a
69Y3062	A177	Xeon E7-2803 6C 1.73 GHz 18 MB 800 MHz 105 W	No	CTO ^a

a. The processor is not available in any of the standard models that are listed in Table 5-5 on page 187. The processor is, however, available through the configure-to-order (CTO) ordering process.

Table 5-15 lists the capabilities of each processor option available for the HX5.

Table 5-15 Intel Xeon E7-8800, E7-4800, and E7-2800 features

Processor model/cores	Scalable to four socket	Processor frequency	Turbo ^a	HT ^b	L3 cache	Power	QPI speed ^c	HX5 memory speed	MAX5 V2 memory speed
E7-8800 series processors									
E7-8867L 10C	Yes	2.13 GHz	Yes	Yes	30 MB	105 W	6.4 GT/s	1066 MHz	1066 MHz
E7-8837 8C	Yes	2.67 GHz	Yes	No	24 MB	130 W	6.4 GT/s	1066 MHz	1066 MHz
E7-4800 series processors									
E7-4870 10C	Yes	2.40 GHz	Yes	Yes	30 MB	130 W	6.4 GT/s	1066 MHz	1066 MHz
E7-4860 10C	Yes	2.26 GHz	Yes	Yes	24 MB	130 W	6.4 GT/s	1066 MHz	1066 MHz
E7-4850 10C	Yes	2.00 GHz	Yes	Yes	24 MB	130 W	6.4 GT/s	1066 MHz	1066 MHz
E7-4830 8C	Yes	2.13 GHz	Yes	Yes	24 MB	105 W	6.4 GT/s	1066 MHz	1066 MHz
E7-4820 8C	Yes	2.00 GHz	Yes	Yes	18 MB	105 W	5.86 GT/s	978 MHz	978 MHz
E7-4807 6C	Yes	1.86 GHz	No	Yes	18 MB	95 W	4.8 GT/s	800 MHz	800 MHz
E7-2800 series processors									
E7-2870 10C	No	2.40 GHz	Yes	Yes	30 MB	130 W	6.4 GT/s	1066 MHz	1066 MHz
E7-2860 10C	No	2.26 GHz	Yes	Yes	24 MB	130 W	6.4 GT/s	1066 MHz	1066 MHz
E7-2850 10C	No	2.00 GHz	Yes	Yes	24 MB	130 W	6.4 GT/s	1066 MHz	1066 MHz
E7-2830 8C	No	2.13 GHz	Yes	Yes	24 MB	105 W	6.4 GT/s	1066 MHz	1066 MHz
E7-2820 8C	No	2.00 GHz	Yes	Yes	18 MB	105 W	5.86 GT/s	978 MHz	No support ^d

Processor model/cores	Scalable to four socket	Processor frequency	Turbo ^a	HT ^b	L3 cache	Power	QPI speed ^c	HX5 memory speed	MAX5 V2 memory speed
E7-2803 6C	No	1.73 GHz	No	Yes	18 MB	105 W	4.8 GT/s	800 MHz	No support ^d

a. Intel Turbo Boost Technology.

b. Intel Hyper-Threading Technology.

c. GT/s or Giga-transfers per second.

d. Connecting a MAX5 is not supported by servers that have the E7-2820 or E7-2803 processors installed.

Processor support: As shown in Table 5-15 on page 200, the Xeon E7-2800 series processor range does not support scaling to four sockets. The E7-2820 and the E7-2803 also do not support the attachment of the MAX5. These technical limitations are specific to these particular processors.

When you install two processors in a single HX5 server (without MAX5), add the IBM HX5 1-Node Speed Burst Card, 59Y588. For details about this feature, see 5.4, “Chassis support” on page 186. Table 5-16 lists the option information.

Table 5-16 HX5 1-Node Speed Burst Card

Part number	Feature code	Description
59Y5889	1741	IBM HX5 1-Node Speed Burst Card

5.11 Memory

The HX5 of machine type 7873, has eight DIMM sockets per processor (a total of 16 DIMM sockets). This machine type supports up to 256 GB of memory when using 16 GB DIMMs. With the addition of the MAX5 memory expansion blade, a single HX5 7873 blade has access to a total of 40 DIMM sockets. This configuration supports up to 640 GB of RAM when using 16 GB DIMMs.

The HX5 of machine type 7872, has eight DIMM sockets per processor (a total of 16 DIMM sockets). This machine type supports 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. This configuration supports up to 320 GB of RAM when using 8 GB DIMMs.

The HX5 and MAX5 use registered DDR3, very low profile (VLP) DIMMs, which provide reliability, availability, and serviceability (RAS) and advanced Chipkill memory protection. For more information about Chipkill memory protection, see “Chipkill” on page 25. For information about RAS, see 2.3.6, “Reliability, availability, and serviceability features” on page 23.

The following topics are described:

- ▶ 5.11.1, “Memory options” on page 202
- ▶ 5.11.2, “Dual inline memory module population order” on page 204
- ▶ 5.11.3, “Memory balance” on page 208
- ▶ 5.11.4, “Memory mirroring” on page 209
- ▶ 5.11.5, “Memory sparing” on page 210

To see a full list of the supported memory features, such as hemisphere mode, Chipkill, non-uniform memory access (NUMA), and memory mirroring, and an explanation of each memory feature, see 2.3, “Memory” on page 16.

5.11.1 Memory options

The available memory options for the HX5 are described.

The HX5 and MAX5 V2 use registered DDR3 very low profile (VLP) DIMMs and provides ECC and advanced Chipkill memory protection. The HX5 type 7873 and the MAX5 V2 support low voltage 1.35 V DIMM modules.

Table 5-17 lists the memory options for the HX5, machine type 7873.

Table 5-17 Memory options for HX5 type 7873, MAX5 V2, and MAX5

Part number	Feature code	HX5 (7873) support	MAX5 V2 support (88Y6128)	MAX5 support (46M6973)	Description
46C0560	A0WX	Yes	Yes	No	2 GB (1x2 GB, 1Rx8, 1.35 V) PC3L-10600 CL9 ECC DDR3 1333 MHz VLP RDIMM
46C0564	A0WZ	Yes	Yes	No	4 GB (1x4 GB, 2Rx8, 1.35 V) PC3L-10600 CL9 ECC DDR3 1333 MHz VLP RDIMM
46C0570	A17Q	Yes	Yes	No	8 GB (1x8 GB, 4Rx8, 1.35 V) PC3L-8500 CL7 ECC DDR3 1066 MHz VLP RDIMM
00D4985	A3BU	Yes	Yes	No	8 GB (1x8 GB, 2Rx8, 1.35 V) PC3L-10600 CL9 ECC DDR3 1333 MHz VLP RDIMM
46C0599	2422	Yes	Yes	No	16 GB (1x16 GB, 4Rx8, 1.35 V) PC3L-10600 CL9 ECC DDR3 1066 MHz VLP RDIMM
90Y3221	A2QP	Yes	Yes	No	16 GB (1x16 GB, 4Rx4, 1.35 V) PC3L-8500 CL7 ECC DDR3 1066 MHz VLP RDIMM
00D5008	A3KN	Yes	Yes	No	32 GB (1x32 GB, 4Rx4, 1.35 V) PC3L-10600 CL9 ECC DDR3 1333 MHz VLP RDIMM
49Y1552	A100	No	No	Yes	4 GB PC3-10600 CL9 ECC VLP (1x4GB, 2Rx8, 1.5 V)
49Y1553	A101	No	No	Yes	8 GB PC3-10600 CL7 ECC VLP (1x8GB, 4Rx8, 1.5 V)

Important: Memory must be fully populated in the HX5 base blade before you populate the MAX5 and IBM MAX5 V2.

Installed DIMM pairs can be of different sizes, but they must be of the same speed. The HX5 supports memory mirroring and memory sparing. MAX5 and MAX5 V2 only support memory mirroring; memory sparing is not supported.

Memory must be installed in pairs of two identical DIMMs, or in quads if memory mirroring is enabled. The options in Table 5-17, however, are for single DIMMs.

Each processor controls eight DIMMs and four memory buffers in the server, as shown in Figure 5-11. To use all 16 DIMM sockets, you must install both processors. If only one processor is installed, you can install only eight DIMMs.

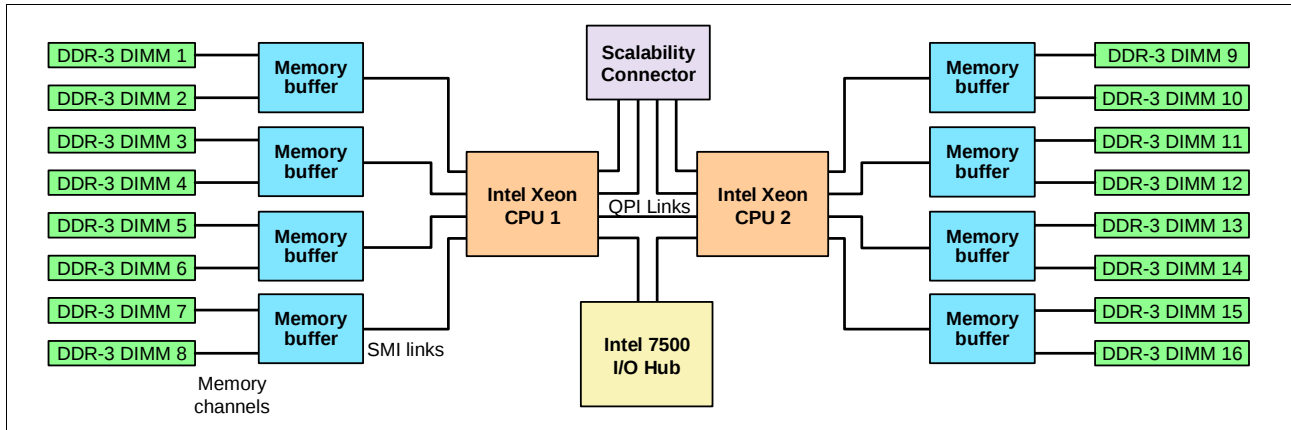


Figure 5-11 HX5 block diagram that shows processors, memory buffers, and DIMMs

Figure 5-12 shows the physical locations of the 16 memory DIMM sockets.

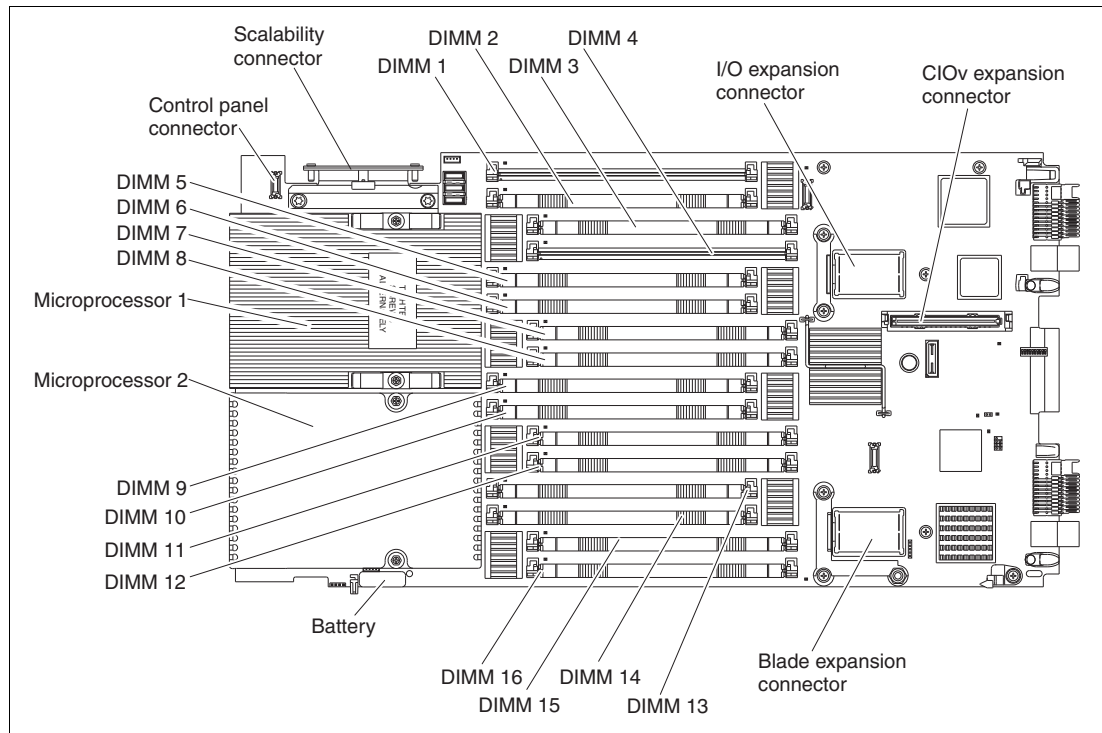


Figure 5-12 DIMM layout on the HX5 system board

The MAX5 memory expansion blade has 24 memory DIMM sockets, as shown in Figure 5-13. The MAX5, which must be connected to an HX5 system (only the one-node HX5 supports the MAX5), has one memory controller and six SMI-connected memory buffers.

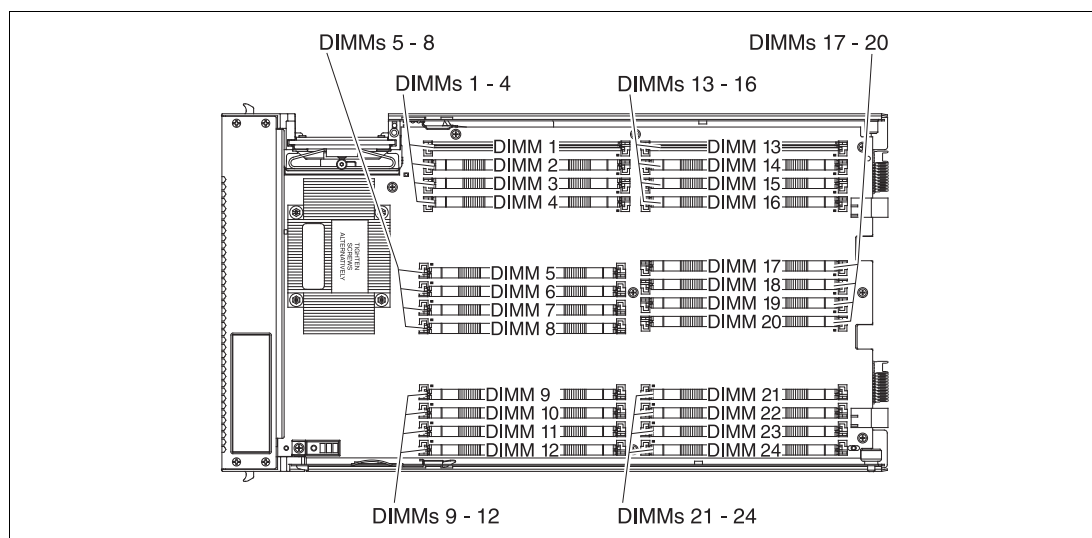


Figure 5-13 DIMM layout on the MAX5 system board

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 5-15 on page 200 lists 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 17.

5.11.2 Dual inline memory module population order

Installing DIMMs in the HX5 and MAX5 in the correct order is essential for system performance. See 5.11.4, “Memory mirroring” on page 209 for the effects on performance when you do not install the DIMMs in the correct order.

HX5 memory population order

As shown in Figure 5-11 on page 203, the HX5 design has two DIMMs per memory buffer and one DIMM socket per memory channel.

Note: 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 web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5084529>

For best performance, install the DIMMs into their sockets as shown in Table 5-18. This sequence spreads the DIMMs across as many memory buffers as possible.

Table 5-18 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 22.

In a two-node (four-socket) configuration with two HX5 servers, follow the memory installation sequence in both nodes. Populate memory to have a balance for each processor in the configuration.

For best performance, follow these general guidelines:

- ▶ Install as many DIMMs as possible. You can get the best performance by installing DIMMs in every socket.
- ▶ Have the same amount of memory for each processor.
- ▶ Spread the memory DIMMs across memory buffers. That is, install one DIMM to a memory buffer before beginning to install a second DIMM to that same buffer. See Table 5-18 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 also shown in Table 5-18.
- ▶ To maximize performance of the memory subsystem, select a processor with the highest memory bus speed (as listed in Table 5-15 on page 200).

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 5-19 shows a NUMA-compliant DIMM installation for a two-node HX5.

Table 5-19 NUMA-compliant DIMM installation for a two-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 22.

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 5-18 on page 205. Then, populate the MAX5 by using the sequence that is listed in Table 5-20 on page 208.
- ▶ The data widths for the following quads must match. For example, DIMMs in each quad must be all 4Rx8 or all 2Rx8. See Figure 5-13 on page 204 for the physical location of these MAX5 DIMMs. Also see Figure 5-14 on page 207 for a block diagram of power domains for 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 5-17 on page 202), this step means that all DIMMs in each of the quads that are 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 the 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 5-20 on page 208. 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 5-14. Each power domain includes two memory controllers and eight DIMM sockets.

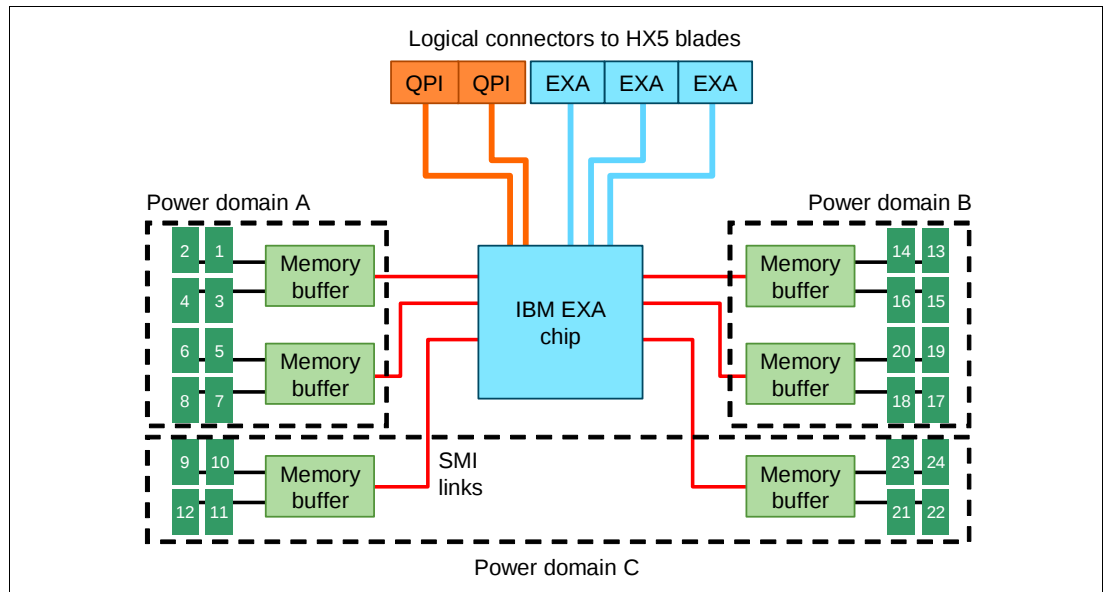


Figure 5-14 Power domains in the MAX5 memory expansion blade

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 5-20 shows the installation order.

Table 5-20 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

5.11.3 Memory balance

The NUMA architecture that is used by the processors in the HX5 is described in 2.3.4, “Non-uniform memory access architecture” on page 21. Because NUMA is used, it is important to ensure that all memory controllers in the system are used by configuring all processors with memory. It is optimal to populate all processors in an identical fashion to provide a balanced system. Populating all processors identically is required by VMware.

Important: When you use a MAX5 with VMware ESX 4.1 or ESXi 4.1, a boot parameter is required to access the MAX5 memory expansion unit that is enabled by NUMA 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 IBM RETAIN® tip H197190 for more information and the necessary parameters:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5084842>

Looking at Figure 5-15 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. Therefore, threads running on Processor 1 have a longer latency to access memory as compared to threads on Processor 0. This result is because of 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, it is important to populate all of the processors with memory, remembering the requirements to ensure optimal interleaving and hemisphere mode.

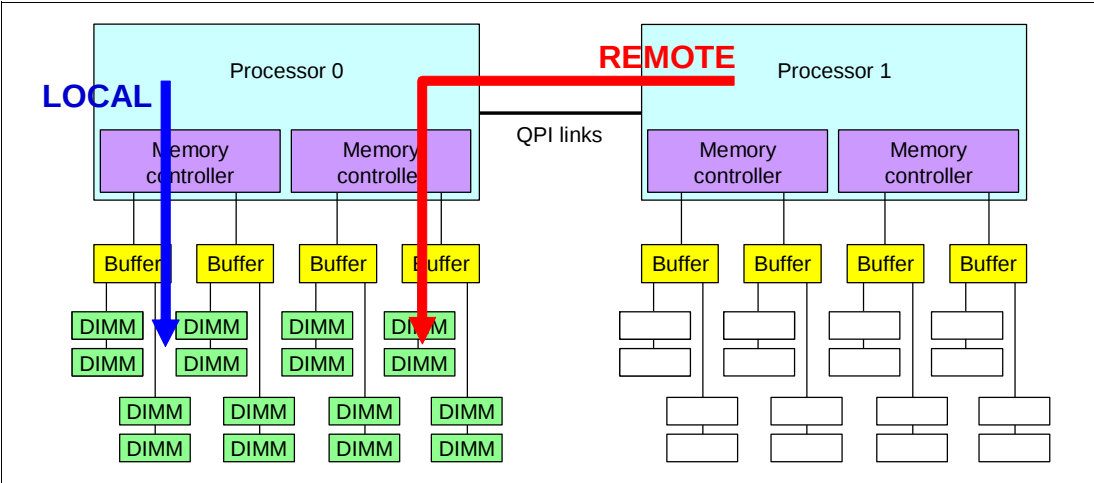


Figure 5-15 Memory latency when not spreading DIMMs across both processors

5.11.4 Memory mirroring

Memory mirroring is supported by 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 24.

DIMM placements for each solution are described.

Important: If you use memory mirroring, all DIMMs must be identical in size and rank.

DIMM placement for HX5

Table 5-21 lists the DIMM installation sequence for memory-mirroring mode when one processor is installed.

Table 5-21 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 5-22 lists the DIMM installation sequence for memory-mirroring mode when two processors are installed.

Table 5-22 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 5-23 lists the DIMM installation sequence in the MAX5 for memory-mirroring mode. Only power domains A and B are populated.

Table 5-23 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				

Important: Memory mirroring is only supported by 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

5.11.5 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 24.

Rank 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 5-24 shows the installation order when one processor is installed.

Table 5-24 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 5-25 shows the installation order when two processors are installed.

Table 5-25 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

Notes:

- Double Device Data Correction (DDDC) is supported by the E7 processors, but only with x4 memory, not with x8 memory.
- The MAX5 memory expansion blade supports redundant bit steering (RBS), but only with x4 memory and not x8 memory. RBS is automatically enabled in the MAX5 if all DIMMs installed are x4 DIMMs.

See the description column in Table 5-17 on page 202 to know which DIMMs are x4.

5.11.6 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, it is populated with sixty-four 4 GB quad-rank DIMMs.

Figure 5-16 shows the peak system-level memory throughput for various memory modes, measured by using an IBM-internal memory load generation tool. As shown, there is a 50% decrease in peak memory throughput when you compare a normal (non-mirrored) configuration to a mirrored memory configuration.

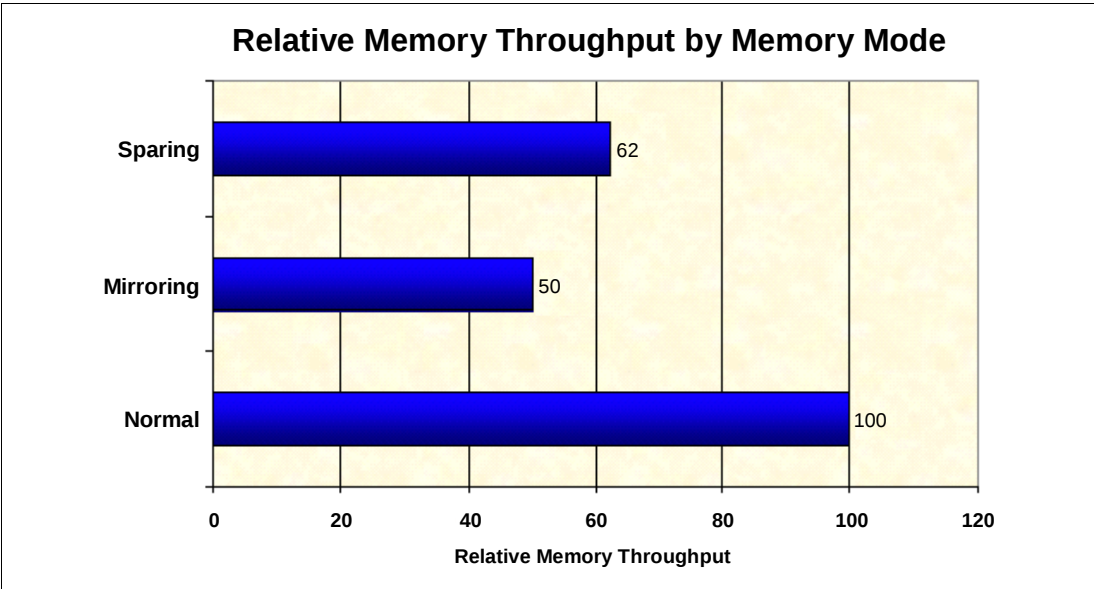


Figure 5-16 Relative memory throughput by memory mode

5.12 Storage

The SSD Expansion Card, SSDs for HX5, and the SAS Configuration Utility are now described.

5.12.1 SSD Expansion Card

The storage system on the HX5 blade is based on the use of the SSD Expansion Card for IBM BladeCenter HX5. This configuration 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 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 5-26 lists the ordering details.

Table 5-26 SSD Expansion Card for IBM BladeCenter HX5

Part number	Feature code	Description
46M6908	5765	SSD Expansion Card for IBM BladeCenter HX5

Figure 5-17 shows the SSD Expansion Card.

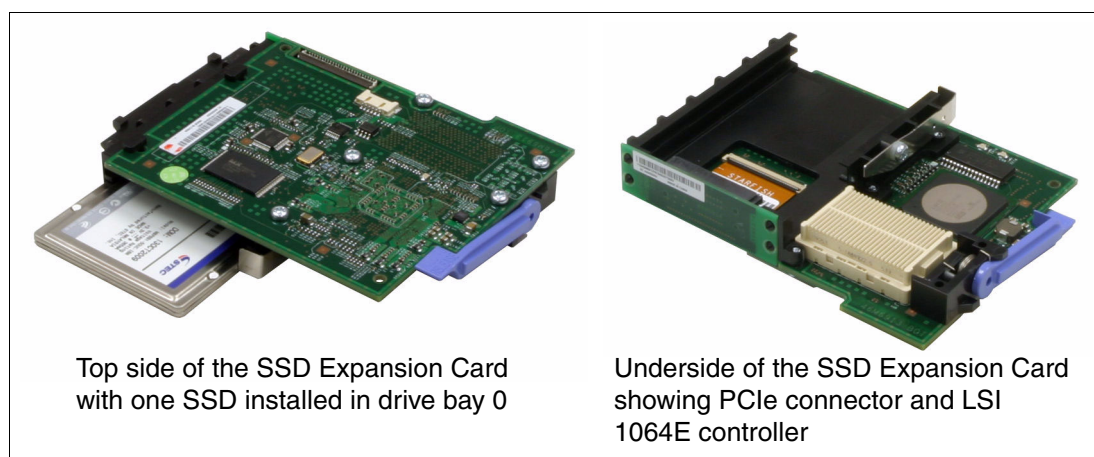


Figure 5-17 SSD Expansion Card for the HX5 (top view: left; underside view: right)

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 5-18.

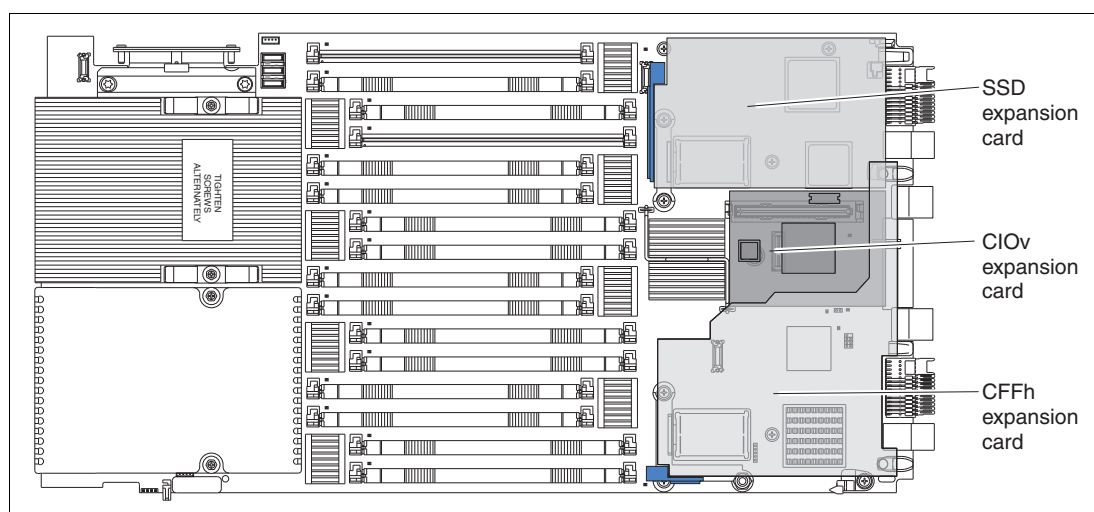


Figure 5-18 Placement of an SSD expansion card with a CIOv card and CFFh card

5.12.2 Solid-state drives for HX5

SSDs offer much greater performance than rotating magnetic media, and are more reliable than hard disk drives. SSDs use much less power than a standard SAS drive, approximately 0.5 W (SSD) versus 11 W (SAS).

Target applications for SSDs include video surveillance, transaction-based databases (DBs), and other applications that have high performance but moderate space requirements.

Table 5-27 lists the supported SSDs.

Table 5-27 Supported SSDs

Part number	Feature code	Description
43W7726	5428	IBM 50 GB SATA 1.8" MLC SSD
43W7746	5420	IBM 200 GB SATA 1.8" MLC SSD
00W1120	A3HQ	IBM 100 GB SATA 1.8" MLC Enterprise SSD
49Y6119	A3AN	IBM 200 GB SATA 1.8" MLC Enterprise SSD
49Y6124	A3AP	IBM 400 GB SATA 1.8" MLC Enterprise SSD

Enterprise Value SSDs and Enterprise SSDs have similar read and write input/output operations per second (IOPS) performance. However, the key difference between them is their endurance, that is, how long they can perform write operations because SSDs have a finite number of program and erase cycles. Enterprise Value SSDs have a better cost/IOPS ratio but lower endurance when compared to Enterprise SSDs.

For more information about SSD drives and their advantages, see 2.8, "IBM eXFlash" on page 38.

5.12.3 LSI SAS Configuration Utility for HX5

Figure 5-19 shows the LSI SAS Configuration Utility window running on a two-node HX5 with one controller in each node. The SAS1064 that is listed first is always the primary node controller, and the SAS1064 listed second is the secondary node's controller.

LSI Corp Config Utility v6.30.00.00 (2009.11.12)							
Adapter List Global Properties							
Adapter	PCI Bus	PCI Dev	PCI Fnc	PCI Slot	FW Revision	Status	Boot Order
SAS1064	04	00	00	02	1.30.05.00-IR	Enabled	0
SAS1064	84	00	00	00	1.30.05.00-IR	Enabled	1
Esc = Exit Menu F1/Shift+1 = Help							
Alt+N = Global Properties -/+ = Alter Boot Order Ins/Del = Alter Boot List							

Figure 5-19 LSI Configuration Utility

In a two-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 5-20 shows the three options in the LSI 1064 setup page, LSI Logic MPT Setup Utility. Only two options are supported by this blade because there are only two drives maximum that are installable in the HX5. An Integrated Mirroring Enhanced (IME) volume requires three drives minimum.

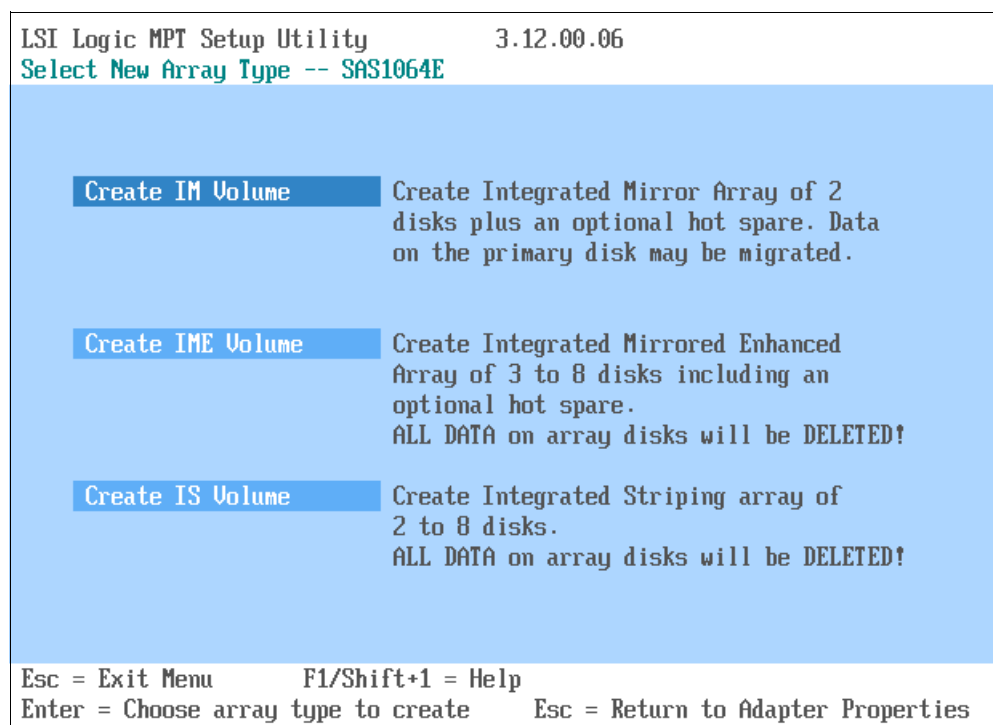


Figure 5-20 RAID choices using the LSI configuration utility

The following options are presented in the configuration utility:

- ▶ **Create IM Volume:** Creates a RAID-1 array
RAID-1 drives are mirrored on a one-to-one 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 must be written twice (once per drive). See the performance chart in Figure 5-21 on page 216 for details.
- ▶ **Create IME Volume:** Creates a RAID-1E array
This option requires three drives, so it is not available in the HX5.
- ▶ **Create Integrated Striping (IS) Volume:** Creates a RAID-0 array
RAID-0 or the IS volume, as shown in LSI, is one of the faster performing disk arrays. This outcome is 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 drives increase performance and data storage efficiency.

The stripe size is 64 KB and cannot be altered.

5.12.4 Determining which SSD RAID configuration to choose

Using an industry standard I/O tool that measures hard disk drive performance, we tested each available configuration type. We used two separate tests, one using 50% sequential reads and 50% random writes (Figure 5-21), and the other using 90% sequential reads and 10% random writes (Figure 5-22). We tested each group by using 16 KB, 512 KB, and 1 MB transfer request sizes. These sizes 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 5-21 SSD RAID configuration test results with 50% sequential and 50% random

Figure 5-22 shows the SSD RAID configuration test results with 90% sequential and 10% 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 5-22 SSD RAID configuration test results with 90% sequential and 10% random test results

Tests: These tests are not certified tests. They are designed to illustrate the performance differences among the three configuration options.

We ran these tests on an HX5 two-node system running four 2 GHz Intel Xeon 7500 series six-core processors and 16 GB of memory. The SSDs were IBM 50 GB SATA 1.8-inch non-hot-swap (NHS) SSDs. Results might vary, depending on the size and type of installed drives.

The results show that when you choose a configuration for setting up hard disk 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, then RAID-0.

5.12.5 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 to 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 5-23.

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 system board to the CIOv connector. This connector is where the SAS Connectivity Card (CIOv) is installed.

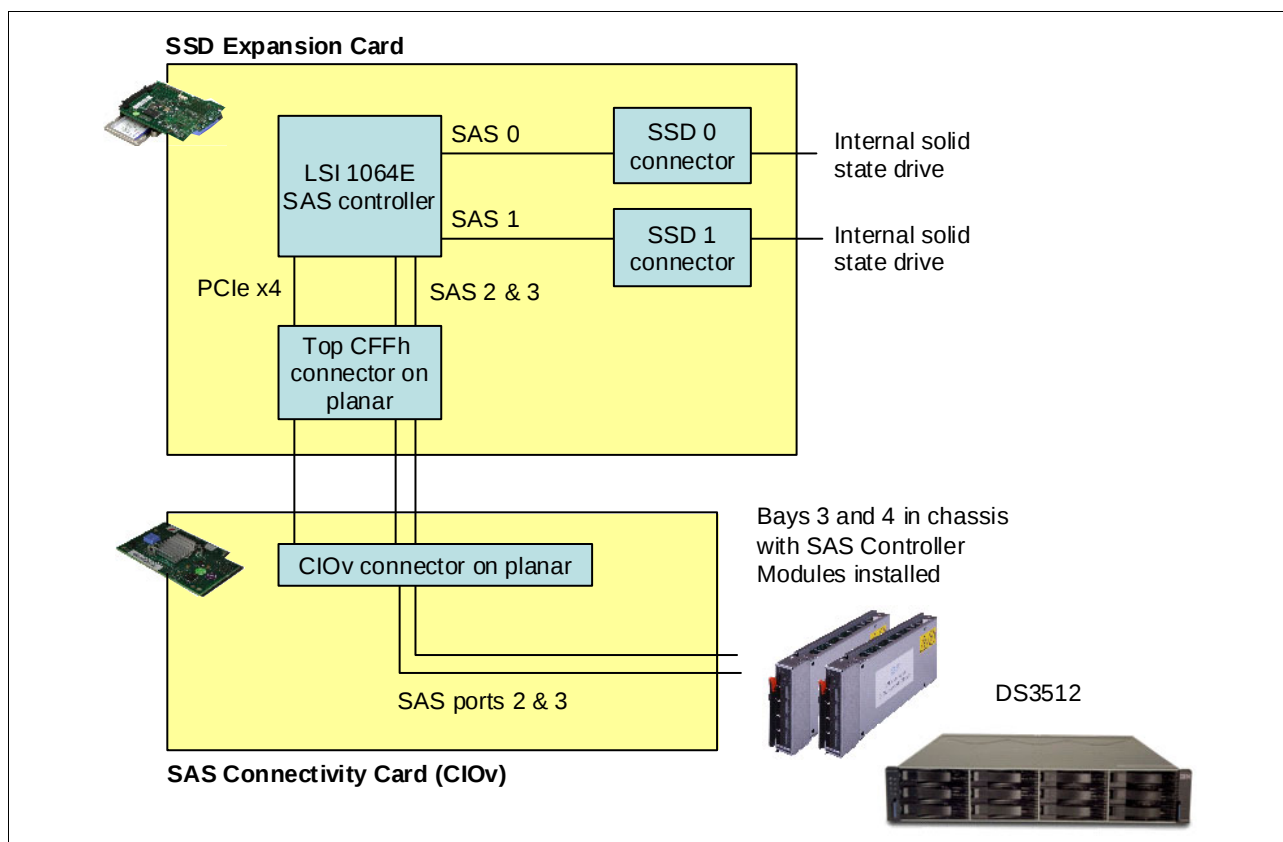


Figure 5-23 Connecting an SAS Connectivity Card and external SAS solution

5.13 BladeCenter PCI Express Gen 2 Expansion Blade II

The IBM BladeCenter PCI Express Gen 2 Expansion Blade II makes it possible 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 by using a PCI Express card.

The expansion blade provides one full-height and full-length PCI Express slot. The blade also provides 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 two-node HX5.

See Table 5-28 for ordering information.

Table 5-28 PCI Express Gen 2 Expansion Blade

Part number	Feature code	Description
68Y7484	A247	IBM BladeCenter PCI Express Gen 2 Expansion Blade II

The expansion blade has the following features:

- Support for PCIe 2.0 adapters in an expansion blade

The expansion blade allows for the installation of one or two standard form factor PCIe 2.0 adapters in a BladeCenter environment. This configuration enables the use of specialized adapters or adapters that otherwise are not available to BladeCenter clients. Each of the two adapters can use 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). This configuration maintains the BladeCenter density advantage, although you still have the option to install PCIe cards as needed. This approach prevents having to attach each expansion blade to a server and the added complexity and cost it brings. The first expansion blade connects to the server blade by 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.

- Ability to connect up to three expansion blades in a single-node HX5, and up to two expansion blades in a two-node HX5

HX5 with MAX 5 support: The HX5 with an attached MAX5 does not support attachment of the PCI Express Gen 2 Expansion Blade II.

- Availability of the CFFh slot

The CFFh expansion connector is accessible on the topmost expansion blade, even with four expansion blades attached. This design allows you to maintain the integrated networking capabilities of the blade server when it is installed in a BladeCenter S, H, or HT chassis.

5.13.1 PCIe SSD adapters

The HX5 supports the High IOPS SSD adapters that are listed in Table 5-29. The adapters must be installed in the BladeCenter PCI Express Gen 2 Expansion Blade II.

Table 5-29 HX5 supported PCI SSD Adapters

Part number	Feature code	Description	Maximum supported
46C9078	A3J3	IBM 365 GB High IOPS MLC Mono Adapter	2
46C9081	A3J4	IBM 785 GB High IOPS MLC Mono Adapter	2
46M0878	0097	IBM 320 GB High IOPS SD Class SSD PCIe Adapter	2
90Y4377	A3DY	IBM 1.2 TB High IOPS MLC Mono Adapter	2
90Y4397	A3DZ	IBM 2.4 TB High IOPS MLC Duo Adapter	1

Many other PCI Express adapters are also supported in the BladeCenter PCI Express Gen 2 Expansion Blade. For more information about the supported PCI Express adapters, see the IBM Redbooks Product Guide, *IBM BladeCenter PCI Express Gen 2 Expansion Blade*, TIPS0783, which is available at this web page:

<http://www.ibm.com/redbooks/abstracts/tips0783.html>

5.14 I/O expansion cards

The HX5 connects to a wide variety of networks and fabrics if the appropriate I/O expansion cards are installed. Supported networks and fabrics include 1 Gb and 10 Gb Ethernet, 4 Gb and 8 Gb Fibre Channel, SAS, and quad data rate (QDR) InfiniBand.

The HX5 blade server with an I/O expansion card is installed in a supported BladeCenter chassis. The system is 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.

5.14.1 CIOv

The CIOv I/O expansion connector provides I/O connections through the midplane of the chassis to modules 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 for each 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 5-30 lists the CIOv expansion cards that are supported in the HX5.

Table 5-30 Supported CIOv expansion cards

Part number	Description
Gb Ethernet	
44W4475	Ethernet Expansion Card (CIOv) for IBM BladeCenter
Fibre Channel	
46M6140	Emulex 8 Gb FC CIOv Dual-port Expansion Card for IBM BladeCenter
46M6065	QLogic 4 Gb Fibre Channel Expansion Card (CIOv) for IBM BladeCenter
44X1945	QLogic 8 Gb Fibre Channel Expansion Card (CIOv) for IBM BladeCenter
SAS/RAID	
43W4068	SAS Connectivity Card (CIOv) for IBM BladeCenter

See the IBM ServerProven compatibility web page 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 two-socket, as shown in Figure 5-24.

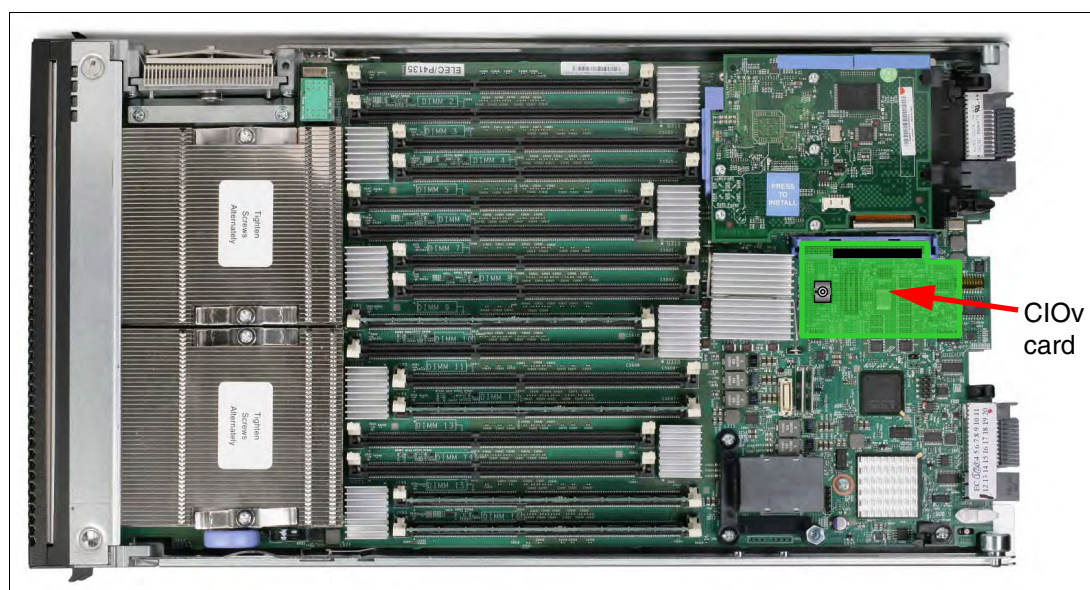


Figure 5-24 HX5 type 7872 showing CIOv I/O expansion card position

5.14.2 CFFh

The CFFh I/O expansion connector provides I/O connections to high-speed switch modules that are in bays 7, 8, 9, and 10 of a BladeCenter H or BladeCenter HT chassis. In the BladeCenter S chassis, the CFFh connector provides I/O connections to the standard switch module in bay 2.

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 5-31 lists the supported CFFh I/O expansion cards.

Table 5-31 Supported CFFh expansion cards

Part numbers	Feature codes	Description
Gb Ethernet		
44W4479	5477	2/4 Port Ethernet Expansion Card
10 Gb Ethernet		
42C1810	3593	Intel 10 Gb 2-port Ethernet Expansion Card
00Y3280	A3JB	QLogic 2-port 10 Gb Converged Network Adapter (CNA)
46M6168	0099	Broadcom 10 Gb 2-port Ethernet Expansion Card
46M6164	0098	Broadcom 10 Gb 4-port Ethernet Expansion Card
90Y3550	A1XG	Emulex Virtual Fabric Adapter II
49Y4265	2436	Emulex Virtual Fabric Advanced Upgrade

Part numbers	Feature codes	Description
90Y3566	A1XH	Emulex Virtual Fabric Adapter Advanced II
81Y1650	5437	Brocade 2 port 10 GbE Converged Network Adapter (CNA)
81Y3133	A1QR	Broadcom 2-port 10 Gb Virtual Fabric Adapter
Fibre Channel		
44X1940	5485	QLogic Ethernet and 8 Gb Fibre Channel Expansion Card (CFFh) for IBM BladeCenter
InfiniBand		
46M6001	0056	2-port 40 Gb InfiniBand Expansion Card (CFFh) for IBM BladeCenter

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 5-25.

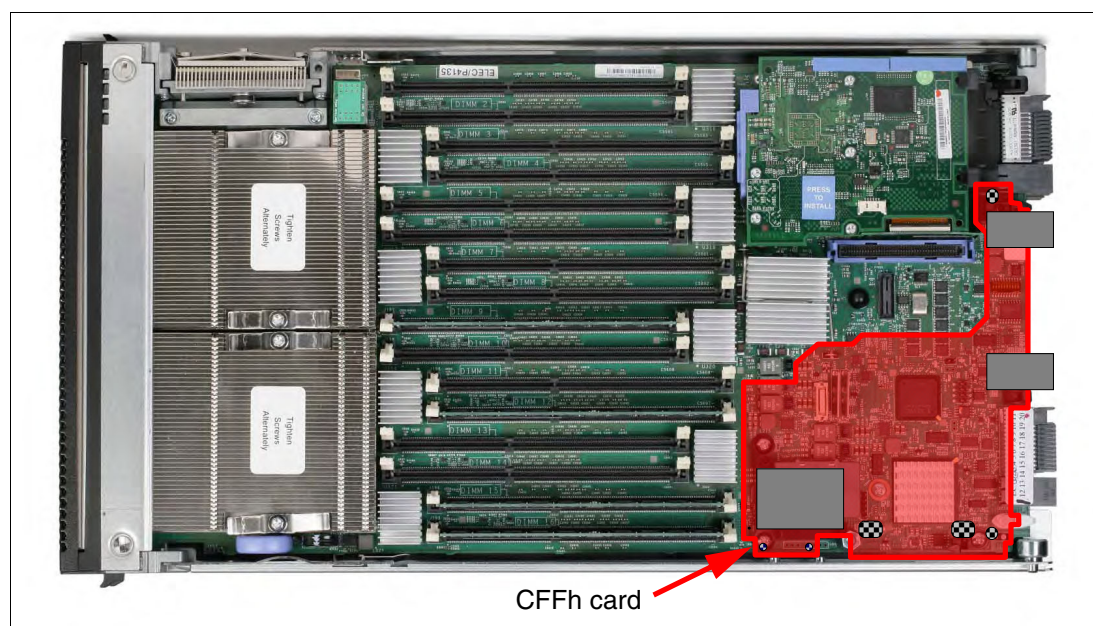


Figure 5-25 The HX5 type 7872 showing the CFFh I/O expansion card position

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 web page:

<http://www.redbooks.ibm.com/big>

5.15 Standard onboard features

The following list provides the standard onboard features of the HX5 blade server:

- ▶ Unified Extensible Firmware Interface (UEFI)
- ▶ Onboard network adapters
- ▶ Integrated Systems Management processor/integrated management module (IMM)
- ▶ Video controller
- ▶ Trusted Platform Module (TPM)

5.15.1 Unified Extensible Firmware Interface

The HX5 2-socket uses an integrated UEFI next-generation basic input/output system a (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 compatible with an earlier version 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 web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5083207>

5.15.2 Onboard network adapters

The HX5 two-socket includes a Broadcom BCM5709S dual-port Gigabit Ethernet controller and supports the following features:

- ▶ TCP Offload Engine (TOE)
- ▶ Failover and load balancing for better throughput and system availability
- ▶ Highly secure remote power management using Intelligent Platform Management Interface (IPMI) 2.0
- ▶ Wake on LAN and Preboot Execution Environment (PXE)
- ▶ IPv4 and IPv6

5.15.3 Integrated management module

The HX5 blade server includes an integrated management module (IMM) to monitor server availability, perform predictive failure analysis, and trigger IBM Systems Director alerts. The IMM performs the functions of the baseboard management controller (BMC) of earlier blade servers. The IMM also 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 web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-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 correspond to DIMMs, CPUs, battery, CIOv connector, CFFh connector, scalability, the system board, non-maskable interrupt (NMI), CPU mismatch, and the SAS connector.

5.15.4 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 5-32 lists the supported video resolutions.

Table 5-32 Supported video resolutions

Resolution	Maximum refresh rate
640 x 480	85 Hz
800 x 600	85 Hz
1024 x 768	75 Hz

5.15.5 Trusted Platform Module

Trusted computing is an industry initiative that provides a combination of secure software and secure hardware to create a trusted platform. Trusted computing follows a specification that increases network security by building unique hardware IDs into computing devices. The HX5 supports Trusted Platform Module (TPM) Version 1.2.

The TPM in the HX5 represents one of the three layers of the trusted computing initiative, as shown in Table 5-33.

Table 5-33 Trusted computing layers

Layer	Implementation
Level 1: Tamper-proof hardware, used to generate trustable keys	Trusted Platform Module
Level 2: Trustworthy platform	▶ UEFI or BIOS ▶ Intel processor
Level 3: Trustworthy execution	▶ Operating system ▶ Drivers

5.16 Integrated virtualization

ESXi is an embedded version of VMware ESX. The hypervisor for the HX5 is completely loaded on the flash drive. Table 5-34 on page 224 lists the ordering information for the IBM USB Memory Key for VMware Hypervisor.

For more information about the USB keys, and to download the IBM customized version of VMware ESXi, see:

<http://www.ibm.com/systems/x/os/vmware/esxi>

Table 5-34 VMware ESXi USB memory keys

Part number	Feature code	Description
41Y8298	A2G0	IBM Blank USB Memory Key for VMware ESXi Downloads
41Y8296	A1NP	IBM USB Memory Key for VMware ESXi 4.1 Update 1
41Y8300	A2VC	IBM USB Memory Key for VMware ESXi 5.0
41Y8307	A383	IBM USB Memory Key for VMware ESXi 5.0 Update 1
41Y8311	A2R3	IBM USB Memory Key for VMware ESXi 5.1

As shown in Figure 5-26, 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.

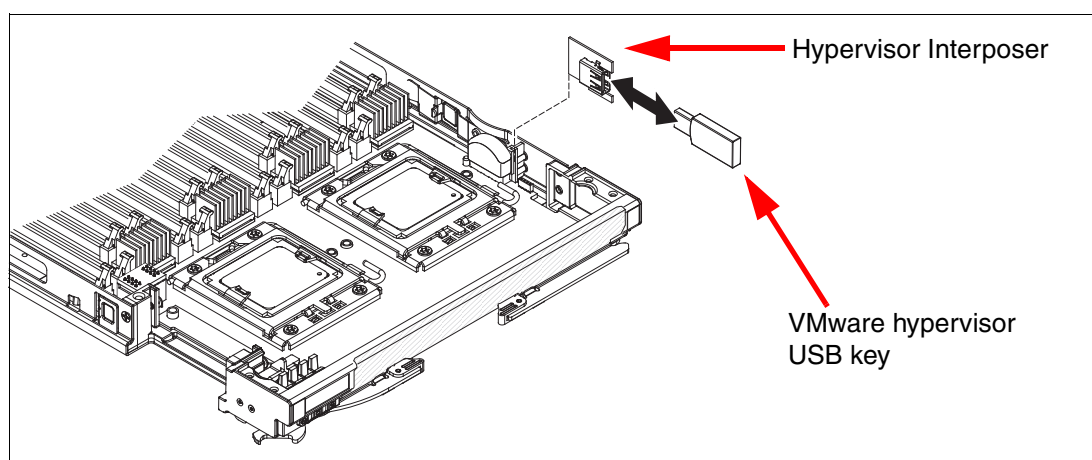


Figure 5-26 Placement of VMware USB key in HX5

See 5.18, “Operating system support” on page 225 for details about VMware ESX and other operating system support.

5.17 Partitioning capabilities

When you have a four-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 5-27.

Scalable Complex Information

A summary of scalable complex configuration, each tab indicating a unique complex and the slots that are occupied.

Complex (6 - 7) Help

Assigned Nodes

Click the checkboxes in the first column to select one or more partitions then, click one of the actions in the action list below the table and click "Perform Action" to perform the desired action.

☐	Partition	Mode	Bay	Name	Status	Processors/Memory	Config Status	Primary
No partitions present								

Available actions

Power Off Partition Perform action

Unassigned Nodes

Click the checkboxes in the first column to select one or more nodes then, click one of the actions in the action list below the table and click "Perform Action" to perform the desired action.

☒	Bay	Name	Status	Processors/Memory
☒	6	SN#YK34C09CJ02W	Powered Off	2 Intel Xeon / 8 DIMMS 2GB
☒	7	SN#YK34C09CW00N	Powered Off	8 DIMMS 2GB

Available actions

Create Partition
Create Partition
Power On Node
Power Off Node

Perform action

Figure 5-27 Two unpartitioned HX5s shown in advanced management module scalable complex

5.18 Operating system support

The HX5 supports the following operating systems:

- ▶ Microsoft Windows Server 2008 HPC Edition
- ▶ Microsoft Windows Server 2008 R2
- ▶ 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
- ▶ Microsoft Windows Server 2012
- ▶ Microsoft Windows Small Business Server 2008 Premium Edition
- ▶ Microsoft Windows Small Business Server 2008 Standard Edition

- ▶ Red Hat Enterprise Linux 5 Server with Xen x64 Edition
- ▶ Red Hat Enterprise Linux 5 Server x64 Edition
- ▶ Red Hat Enterprise Linux 6 Server x64 Edition
- ▶ Red Hat Enterprise MRG 2.0 Realtime (x64)
- ▶ Solaris 10 Operating System
- ▶ SUSE Linux Enterprise Server 10 for AMD64 / EM64T
- ▶ SUSE Linux Enterprise Server 11 for AMD64 / EM64T
- ▶ SUSE Linux Enterprise Server 11 with Xen for AMD64 / EM64T
- ▶ VMware ESX 4.1
- ▶ VMware ESXi 4.1
- ▶ VMware vSphere 5 (ESXi)
- ▶ VMware vSphere 5.1 (ESXi)

See the ServerProven web page for the most recent information:

<http://ibm.com/systems/info/x86servers/serverproven/compat/us/nos/ematrix.shtml>



Systems management

The eX5 hardware platform includes an entire family of systems management software applications and utilities to help with the management and maintenance of the systems. An overview of these applications is provided.

There are two specific systems management areas that are covered: Embedded firmware and external applications. The embedded firmware includes the Unified Extensible Firmware Interface (UEFI), the integrated management module (IMM), and the embedded diagnostic tests. The external applications include the following components: ServerGuide, UpdateXpress System Pack Installer (UXSPI), IBM Bootable Media Creator (BoMC), Dynamic System Analysis (DSA), IBM Systems Director, Active Energy Manager, BladeCenter Open Fabric Manager, VM Control, Network Control, Storage Control, and IBM Tivoli® Provisioning Manager for Operating System Deployments (Tivoli Provisioning Manager for Tivoli Provisioning Manager for OS Deployment).

The following topics are covered:

- ▶ 6.1, “Management applications” on page 228
- ▶ 6.2, “Embedded firmware” on page 228
- ▶ 6.4, “Firmware levels” on page 232
- ▶ 6.5, “UpdateXpress” on page 232
- ▶ 6.6, “Deployment tools” on page 233
- ▶ 6.7, “Configuration utilities” on page 236
- ▶ 6.8, “IBM Dynamic Systems Analysis” on page 238
- ▶ 6.9, “IBM Systems Director” on page 239

The systems management home page IBM Systems is at this web page:

<http://www.ibm.com/systems/management>

6.1 Management applications

The following list provides systems management applications that support the eX5 hardware platform:

- ▶ ServerGuide
- ▶ ServerGuide Scripting Toolkit
- ▶ Bootable Media Creator
- ▶ UpdateXpress System Pack Installer (UXSPI)
- ▶ Storage Configuration Manager
- ▶ Start Now Advisor
- ▶ IBM Advance Settings Utility (ASU)
- ▶ MegaRAID Storage Manager (MSM)
- ▶ Dynamic System Analysis (DSA)
- ▶ IBM Systems Director

For more information, including links and user guides, see the IBM ToolsCenter:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=T00L-CENTER>

6.2 Embedded firmware

This section covers systems management firmware that is located within non-volatile memory on the system board and includes the UEFI, IMM, and Real Time Diagnostics.

UEFI replaces basic input/output system (BIOS) in the eX5 portfolio of servers, as with our other x86 servers. UEFI provides a current, well-defined environment for starting an operating system and running pre-boot applications. UEFI is fully backward-compatible with BIOS but provides additional functionality such as a better user interface, easier management, and better integration with pre-boot configuration software.

UEFI includes the following features:

- ▶ New architecture with increased memory for advanced functions
- ▶ Complete setup solution that moves the configuration function of adapters into UEFI
- ▶ Simplified remote set up with ASU, facilitating 100% coverage of settings with out-of-band access
- ▶ Support for common updates by using the iFlash tool and IMM flash manager
- ▶ Unified code base for the entire IBM x86 portfolio
- ▶ More readable event logs
- ▶ No beep codes (which are difficult to diagnose remotely or locate among individual blades in a chassis or servers in a densely populated rack)
- ▶ Complete readability of all light path diagnostic tests remotely
- ▶ No limits on number of adapters and no 1801 Peripheral Component Interconnect (PCI) resource allocation errors
- ▶ Reduction of unrecoverable errors through new cache-writeback-retry capabilities
- ▶ Multinode power capping
- ▶ Ability to run in 64-bit native mode
- ▶ Support for BIOS-enabled operating systems

To use all the features that UEFI offers requires a UEFI-aware operating system and adapters. UEFI is fully compatible with an earlier version with BIOS.

The follow operating systems support UEFI:

- ▶ Microsoft Windows Server 2008 x64
- ▶ SUSE Linux Enterprise Server 11
- ▶ Red Hat Enterprise Linux 6

Figure 6-1 shows the UEFI boot panel.

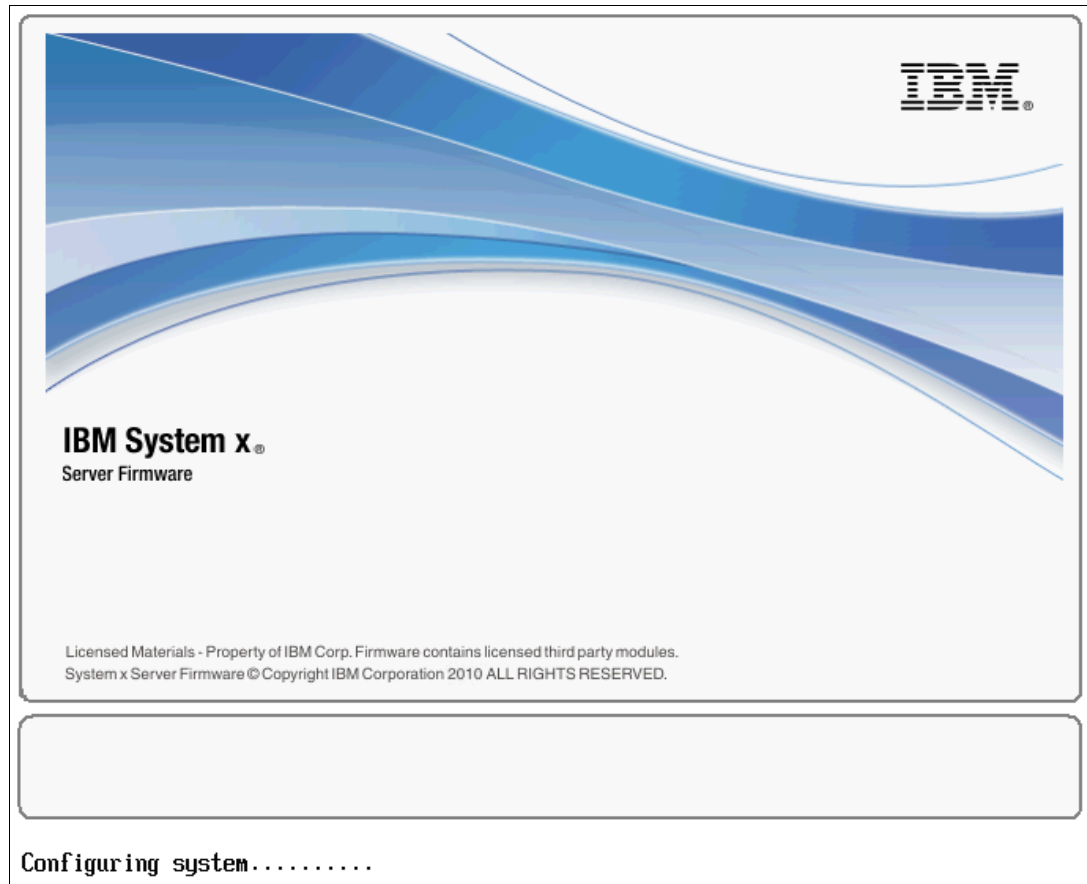


Figure 6-1 Boot page of UEFI BIOS

Figure 6-2 shows the initial setup panel for UEFI, which is displayed by pressing the F1 key during the boot process. From this panel, you can view the system information, change the system settings, change the date and time, change the boot options, view the event log, or add and remove system passwords.

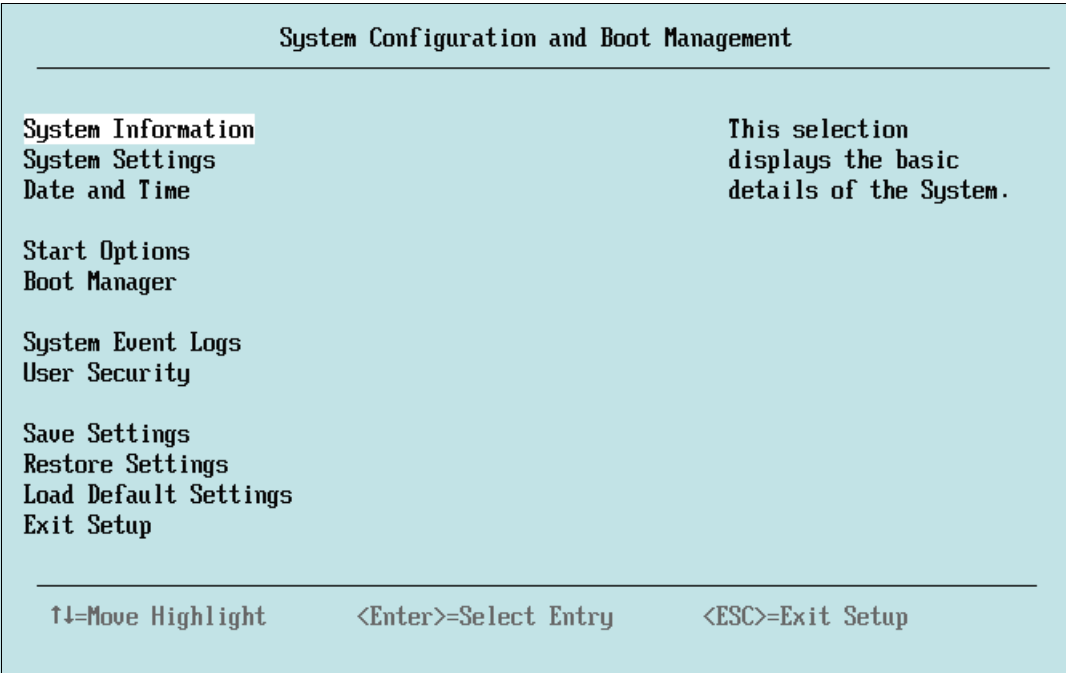


Figure 6-2 UEFI initial setup page

On the System Summary window that is shown in Figure 6-3, you can find the server model, serial number information, and information about the processor speeds.

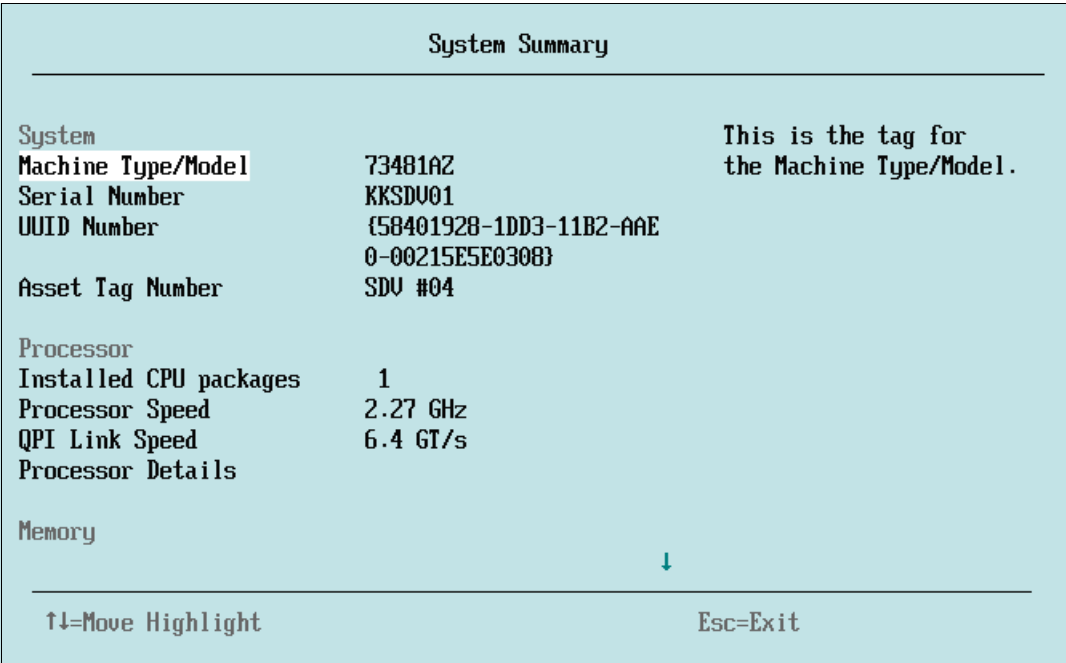


Figure 6-3 UEFI showing the system summary page

Figure 6-4 shows the System Settings menu. From here, you can drill down through the submenus to perform more configuration of the eX5 system.

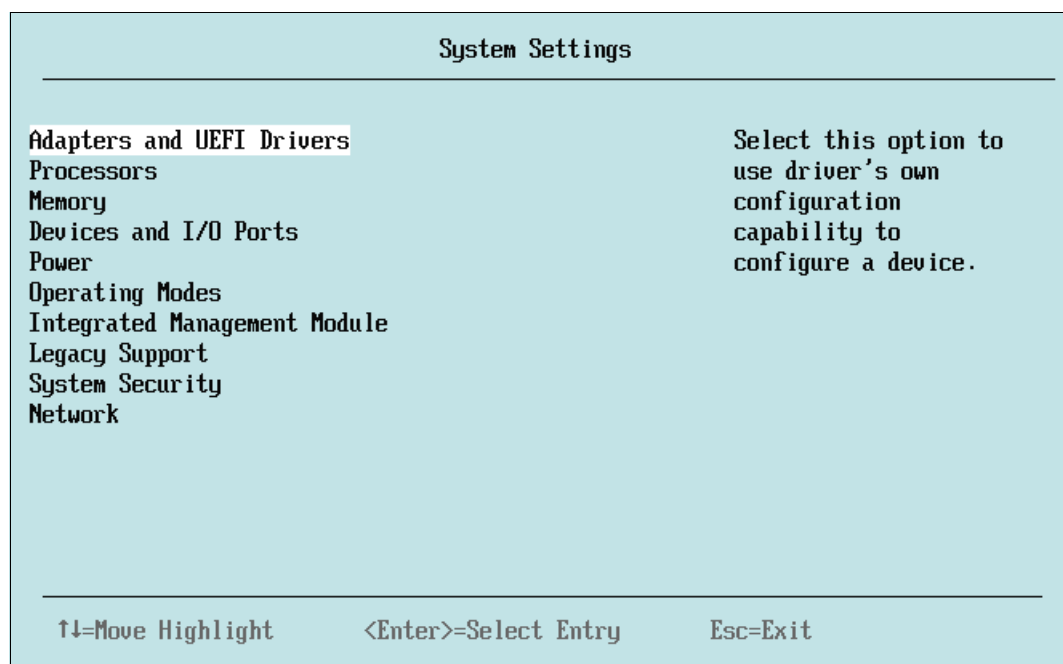


Figure 6-4 UEFI System Settings menu

For more information about UEFI, see the IBM white paper *Transitioning to UEFI and IMM*, which is available at the following web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5079769>

6.3 Integrated management module

The integrated management module (IMM) is a single chip that is embedded into every eX5 server that combines the functionality of the previous baseboard management controller (BMC) and the Remote Supervisor Adapter II (RSA II). Its components include a video controller, remote control, and remote media. IMM firmware is shared across all eX5 servers and all IBM systems that use the IMM, simplifying the firmware update process. No special drivers are required, and it is configurable both in-band and out-of-band. System alerts for the eX5 platform are sent by using industry standards such as Common Information Model (CIM) and Web Services for Management (WS-Management).

IMM offers the following benefits:

- ▶ Provides diagnostic tests, virtual presence, and remote control to manage, monitor, troubleshoot, and repair from anywhere.
- ▶ Securely manages servers remotely, independent of the operating system state.
- ▶ Can 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.

- ▶ Helps 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.

For more information about IMM, see the IBM white paper *Transitioning to UEFI and IMM*, available at the following web address:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5079769>

6.4 Firmware levels

In a two-node environment, ensure that the firmware levels on the nodes in the complex are at the same level. Unpredictable results might occur if the firmware levels, such as FPGA or IMM, fall out of synch between the nodes. The method for updating the firmware is to use the online update utilities that run under the operating system. One such utility is UpdateXpress System Pack Installer, described in 6.5, “UpdateXpress” on page 232. These update packages automatically discover all nodes and update all discovered nodes at the same time.

If no operating system is running on the system, use the Bootable Media Creator method to update the firmware. To download, go to the following web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=T00L-BOMC>

You must ensure that the firmware for all nodes in the complex are updated successfully before you reboot the system. Although, out-of-band updates with the IMM or BladeCenter Advanced Management Module are supported.

The BladeCenter HX5 server supports creating multiple partitions within the complex. When multiple partitions are in the complex, the online update utilities can update the firmware only for the node in that partition. As a result, you must also update the firmware for the nodes in the other partitions before you reboot any partition.

6.5 UpdateXpress

IBM UpdateXpress can help reduce the total cost of ownership of the IBM eX5 hardware by providing an effective and simple way to update device drivers, server firmware, and firmware of supported options. UpdateXpress is available for download at no charge.

UpdateXpress consists of the UpdateXpress System Pack Installer (UXSPI) and the UpdateXpress System Packs (UXSPs). The UXSPI can be used to automatically download the appropriate UXSPs from the IBM website or it can be used to point to a previously downloaded UXSP.

The UpdateXpress System Pack Installer can be run under an operating system to update both the firmware and device drivers of the system. The currently supported operating systems and distributions are as follows:

- ▶ Microsoft Windows Server 2003, including 2003 R2
- ▶ Windows Small Business Server 2003, including 2003 R2
- ▶ Microsoft Windows Server 2008, including 2008 R2
- ▶ Microsoft Windows Server 2012
- ▶ Windows Small Business Server 2011
- ▶ Windows HPC 2008 R2 (x64 only)

- ▶ SUSE Linux Enterprise Server 11
- ▶ SUSE Linux Enterprise Server 10
- ▶ Red Hat Enterprise Linux 6
- ▶ Red Hat Enterprise Linux 5
- ▶ Red Hat Enterprise Linux 4
- ▶ VMware ESXi 4 Update 3 (IBM customized version only)
- ▶ VMware ESXi 4.1, including updates 1, 2, and 3 (IBM customized version only)
- ▶ VMware ESXi 5, including update 1 (IBM customized version only)
- ▶ VMware ESXi 5.1 including update 1 and patch (IBM customized version only)

For more information, including downloading the tools and updates, see the following web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=SERV-XPRESS>

The web page includes a link to the *IBM UpdateXpress System Pack Installer User's Guide*.

6.6 Deployment tools

This section describes the collection of tools and applications that are used to initially build and deploy a new server. Table 6-1 lists the action and appropriate tool for performing these actions.

Table 6-1 Deployment tool usage

Action	Deployment tool
Update firmware on a system before installing an operating system	IBM Bootable Media Creator (BoMC), ToolsCenter Suite
Update firmware on a system that already has an operating system installed	UpdateXpress System Pack Installer
Install a Windows operating system	ServerGuide
Install a Linux operating system	IBM ServerGuide Scripting Toolkit
Install a Windows or Linux operating system in an unattended mode	IBM ServerGuide Scripting Toolkit

6.6.1 Bootable Media Creator

Use the BoMC to create a CD, DVD, or USB key image that can be used to boot a system for applying firmware updates, running preboot diagnostic tests, and deploying Windows operating systems. BoMC also supports the creation of Preboot Execution Environment (PXE) files, which allow for the creation of a network boot image. Support for multiple systems can be contained on one bootable media image.

Multiple IBM System x and BladeCenter tools and UpdateXpress System Packs can be bundled onto a single bootable image.

Bootable Media Creator supports the following operating systems:

- ▶ Windows XP Professional Edition
- ▶ Windows Vista Ultimate Editions
- ▶ Windows 7
- ▶ Microsoft Windows Server 2003

- ▶ Microsoft Windows Storage Server 2003
- ▶ Microsoft Windows Small Business Server 2003
- ▶ Microsoft Windows Server 2008
- ▶ Windows Small Business Server 2011
- ▶ Windows Server 2012
- ▶ SUSE Linux Enterprise Server 11
- ▶ SUSE Linux Enterprise Server 10
- ▶ Red Hat Enterprise Linux 6
- ▶ Red Hat Enterprise Linux 5
- ▶ Red Hat Enterprise Linux 4

For more information, including the *Bootable Media Creator Installation and User's Guide*, see the following web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=T00L-BOMC>

6.6.2 ServerGuide

IBM ServerGuide is an installation assistant that can simplify the process of installing a Windows operating system and configuring an eX5 server. ServerGuide can also help with the installation of the latest device drivers and other system components.

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 hardware detected.
- ▶ Reduces rebooting that is required during hardware configuration and Windows operating system installation, allowing you to get your eX5 server up and running sooner.
- ▶ Provides a consistent server installation by using IBM best practices for installing and configuring an eX5 server.
- ▶ Provides access to more firmware and device drivers that might not be applied at the installation time, such as adapters that are added to the system later.

ServerGuide can be downloaded from the IBM ServerGuide web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=SERV-GUIDE>

6.6.3 ServerGuide Scripting Toolkit

The IBM ServerGuide Scripting Toolkit is a collection of tools and scripts that are designed to deploy software to your eX5 server in a repeatable and predictable way. These scripts are often used with IBM ServerGuide and IBM UpdateXpress to provide a solution for performing unattended installations.

The ServerGuide Scripting Toolkit is available for Windows and Linux. Both versions offer the following benefits:

- ▶ Allows for the tailoring and building of custom hardware deployment solutions.
- ▶ Provides hardware configuration utilities and OS installation examples for the eX5 hardware platforms.

The ServerGuide Scripting Toolkit allows for the creation of a bootable CD, DVD, or USB key that supports the following items:

- ▶ Network and mass storage devices
- ▶ Policy-based Redundant Array of Independent Disks (RAID) configuration

- ▶ Configuration of system settings by using ASU
 - ▶ Configuration of Fibre Channel host bus adapters (HBAs)
 - ▶ Local self-contained DVD deployment scenarios
 - ▶ Local CD/DVD and network share-based deployment scenarios
 - ▶ RSA II, IMM, and BladeCenter advanced management module (AMM) remote disk scenarios
 - ▶ UpdateXpress System Packs installation that is 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 Director Agent:
- Common Agent 6.1 or later
 - Core Services 5.20.31 or later
 - Director Agent 5.1 or later

Additionally, the Windows version of the ServerGuide Scripting Toolkit enables automated operating system support for the following Windows operating systems:

- ▶ Microsoft Windows Server 2003, Standard, Enterprise, and Web Editions
- ▶ Microsoft Windows Server 2003 R2, Standard and Enterprise Editions
- ▶ Microsoft Windows Server 2003, Standard and Enterprise x64 Editions
- ▶ Microsoft Windows Server 2003 R2, Standard and Enterprise x64 Editions
- ▶ Microsoft Windows Server 2008, Standard, Enterprise, Datacenter, and Web Editions
- ▶ Microsoft Windows Server 2008 x64, Standard, Enterprise, Datacenter, and Web Editions
- ▶ Microsoft Windows Server 2008, Standard, Enterprise, and Datacenter Editions without Hyper-V
- ▶ Microsoft Windows Server 2008 x64, Standard, Enterprise, and Datacenter Editions without Hyper-V
- ▶ Microsoft Windows Server 2008 R2 x64, Standard, Enterprise, Datacenter, and Web Editions
- ▶ Microsoft Windows Server 2012

The Linux Scripting Toolkit uses a console to simplify the steps in creating, customizing, and deploying hardware configurations and operating system deployments 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/SP4
- ▶ SUSE Linux Enterprise Server 10 x64 SP1/SP2/SP3/SP4
- ▶ SUSE Linux Enterprise Server 11 32 bit Base/SP1/SP2
- ▶ SUSE Linux Enterprise Server 11 x64 Base/SP1/SP2
- ▶ 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/U6/U7/U8
- ▶ Red Hat Enterprise Linux 5 x64 U1/U2/U3/U4/U5/U6/U7/U8
- ▶ Red Hat Enterprise Linux 6 32 bit U1/U2/U3
- ▶ Red Hat Enterprise Linux 6 x64 U1/U2/U3
- ▶ VMware ESX Server 3.5 U4/U5
- ▶ VMware ESX Server 4.0/4.0u1/4.0u2/4.1/4.1u1/4.1u2/4.1u3

Automated post-installation deployment of the following operating systems:

- ▶ Red Hat Enterprise Linux 4 AS/ES 32 bit U9
- ▶ Red Hat Enterprise Linux 4 AS/ES x64 U9

Automated deployment of the following operating systems in Native UEFI mode:

- ▶ SUSE Linux Enterprise Server 11 SP1/SP2
- ▶ Red Hat Enterprise Linux 6 x64 U1/U2/U3

To download the Scripting Toolkit or the *IBM ServerGuide Scripting Toolkit User's Reference*, see the IBM ServerGuide Scripting Toolkit web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=SERV-T00LKIT>

6.6.4 IBM Start Now Advisor

The IBM Start Now Advisor tool can help you configure the HX5 blades and the IBM BladeCenter chassis with a wizard that guides you through the initial configuration complexities. The software automatically detects the type of BladeCenter chassis and then does the following initial configuration steps:

- ▶ Checks the IBM website for the latest firmware
- ▶ Updates the chassis component firmware, including the AMM firmware, blade servers, serial-attached SCSI (SAS) RAID controller, SAS connectivity module, storage modules, and Ethernet switches
- ▶ Allows for the *call home* features, found in the Service Advisor, to notify IBM on any service events or hardware failures that require support
- ▶ Helps set up BladeCenter Open Fabric Manager

To download the latest Start Now Advisor or to obtain more information, see the Start Now Advisor web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=T00L-SNA>

6.7 Configuration utilities

Two tools are described that can be enabled for easy configuration of your eX5 platform and the attached storage.

6.7.1 MegaRAID Storage Manager

MegaRAID Storage Manager software enables you to 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.

The MegaRAID Storage Manager software has the following hardware requirements:

- ▶ A PC-compatible computer with an IA-32 (32-bit) Intel Architecture processor or an EM64T (64-bit) processor
- ▶ A minimum of 256 MB of system memory (512 MB recommended)
- ▶ A drive with at least 400 MB available free space

The MegaRAID Storage Manager software supports the following operating systems:

- ▶ Microsoft Windows 2003, Microsoft Windows 2008, Microsoft Windows 2008 SP2, Microsoft Windows 2008 R2, and Windows R2 SP1
- ▶ Red Hat Enterprise Linux Versions 4.7, 4.8, 5.5, 5.6, 5.7, 6.0, 6.1, and 6.2
- ▶ SUSE SLES Server 10 SP3, 10 SP4, 11, and 11 SP1
- ▶ VMware ESX 4.0, 4.1, and 5.0

To download the latest MegaRAID Storage Manager or to obtain more information, see the MegaRAID Storage Manager web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5077712>

6.7.2 Advanced Settings Utility

The ASU allows for the modification of your eX5 server firmware settings from a command line. It supports multiple operating systems, such as Linux, Solaris, and Windows (including Windows Preinstallation Environment (PE)). Firmware settings that can be modified on the eX5 platform include UEFI and IMM settings.

You can do the following tasks by using the ASU:

- ▶ Modify the UEFI complementary metal-oxide semiconductor (CMOS) settings without the need to restart the system and access the F1 menu.
- ▶ Modify the IMM setup settings.
- ▶ Modify a limited set of vital product data (VPD) settings.
- ▶ Modify the Internet Small Computer System Interface (iSCSI) boot settings. (To do so through the ASU, you must first manually configure the iSCSI settings through the server setup utility).
- ▶ Remotely modify all of the settings through an Ethernet connection.

Download the latest version and get the *Advanced Settings Utility User's Guide* from the Advanced Settings Utility web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=T00L-ASU>

6.7.3 Storage Configuration Manager

The IBM Storage Configuration Manager (SCM) is a web-based application that enables the management and configuration of the following IBM BladeCenter devices:

- ▶ IBM BladeCenter SAS Connectivity Module
- ▶ IBM BladeCenter Six-Disk Storage Module
- ▶ SAS Expansion Card for IBM BladeCenter
- ▶ SAS/SATA RAID Kit (Integrated RAID Controller)
- ▶ IBM BladeCenter S SAS RAID Controller Module
- ▶ IBM ServeRAID MR Controller

Download the latest version and get the *IBM Storage Configuration Manager Planning, Installation, and Configuration Guide* from the Storage Configuration Manager web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=T00L-SCM>

6.8 IBM Dynamic Systems Analysis

IBM Dynamic Systems Analysis (DSA) collects and analyzes all your eX5 system information and produces a report to aid in diagnosing system problems. The following information is collected from an eX5 system:

- ▶ System configuration
- ▶ Installed applications and interim fixes
- ▶ Device drivers and system services
- ▶ Network interfaces and settings
- ▶ Performance data and running process details
- ▶ Hardware inventory, including PCI information
- ▶ Vital product data and firmware information
- ▶ Small Computer System Interface (SCSI) device sense data
- ▶ ServeRAID configuration
- ▶ Application, system, security, ServeRAID, and service processor system event logs

Additionally, DSA creates a merged log that allows users to easily identify cause-and-effect relationships from various log sources in the system.

Three editions of DSA are available:

- ▶ DSA Preboot Edition
- ▶ DSA Installable Edition
- ▶ DSA Portable Edition

The preboot edition can either be added to a bootable media by using IBM ToolsCenter Bootable Media Creator, or you can download the Windows or Linux update package for Preboot DSA. Then, reboot the system into the image that you created. The installable edition is downloaded and installed on a system where persistent use of DSA is required. The DSA Portable Edition can be downloaded and run without modifying any system files or configurations.

The following operating systems and distributions are supported:

- ▶ Windows Server 2012 Edition
 - Microsoft Windows Server 2012
- ▶ Windows Server 2011 Editions:
 - Microsoft Windows Small Business Server 2011
 - Microsoft Windows Small Business Server 2011 Essential
- ▶ Windows Server 2008 Editions:
 - Microsoft Windows Server 2008 R2
 - Microsoft Windows Server 2008 R2 SP1
 - Microsoft Windows Server 2008 R2 HPC Edition (x64, ROK)
 - Microsoft Windows Server 2008, Datacenter Edition (x86, x64)
 - Microsoft Windows Server 2008, Web Edition (x86, x64)
 - Microsoft Windows Server 2008, Enterprise Edition (x86, x64)
 - Microsoft Windows Server 2008, Standard Edition (x86, x64)
 - Microsoft Windows Server 2008 HPC Edition
 - Microsoft Windows Server 2008 Foundation
 - Windows Essential Business Server 2008 Premium Edition
 - Windows Essential Business Server 2008 Standard Edition
- ▶ Windows Server 2003 Editions:
 - Microsoft Windows Server 2003/2003 R2, Standard Edition (x86, x64)

- Microsoft Windows Server 2003/2003 R2, Web Edition
- Microsoft Windows Server 2003/2003 R2, Enterprise Edition (x86, x64)
- Microsoft Windows Server 2003/2003 R2, Enterprise Edition with Microsoft Cluster Service (MSCS) (x86, x64)
- Microsoft Windows Server 2003/2003 R2, Datacenter Edition (x86, x64)
- ▶ Windows Preinstallation Environment:
 - Microsoft Windows Preinstallation Environment 2.1
 - Microsoft Windows Preinstallation Environment 3.0
- ▶ SUSE Linux:
 - SUSE Linux Enterprise Server 11 (Up to SP2) (x86/x64)
 - SUSE Linux Enterprise Server 11 with Xen (Up to SP2) (x86/x64)
 - SUSE Linux Enterprise Server 10 (Up to SP4) (x86/x64)
 - SUSE Linux Enterprise Server 10 with Xen (Up to SP4) (x86/x64)
 - SUSE Linux Enterprise Real Time 10 (Up to SP4) (AMD64/EM64T)
- ▶ Red Hat:
 - Red Hat Enterprise Linux 6 (Up to U3) (x86, x64)
 - Red Hat Enterprise Linux 5 (Up to U8) (x86, x64)
 - Red Hat Enterprise Linux 5 (Up to U8) with Xen (x86, x64)
 - Red Hat Enterprise Linux 4 (Up to U9) (x86, x64)
- ▶ VMware:
 - VMware vSphere Hypervisor 5.1 (ESX5.1) (supported only through use of the vmware-esxi option)
 - VMware vSphere Hypervisor 5.0 (ESX5) (UP to U1) (supported only through use of the vmware-esxi option)
 - VMware ESX Server, 4.1 (Up to U3)
 - VMware ESXi 4.1 (Up to 4.1 U3) (supported only through use of the vmware-esxi option)
 - VMware ESX Server 4.0 (Up to U3)
 - VMware ESXi 4.0 (Up to 4.0 U3) (supported only through use of the vmware-esxi option)

Download the latest version of the software and the *Dynamic System Analysis Installation and User's Guide* from the IBM Dynamic System Analysis (DSA) web page:

<http://ibm.com/support/entry/portal/docdisplay?ln docid=SERV-DSA>

6.9 IBM Systems Director

IBM Systems Director is a Platform Manager that offers the following benefits:

- ▶ Enables for 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. IBM Systems Director Server can be installed on the following systems: IBM AIX®, Windows, Linux on Power, Linux on x86, or Linux on IBM System z®.

IBM Systems Director provides the following functionality:

- ▶ Discovery
- ▶ Monitoring and reporting
- ▶ Software updates
- ▶ Configuration management
- ▶ Virtual resource management
- ▶ Remote control
- ▶ Automation

For more information about implementing IBM Systems Director, see the following sources:

- ▶ *IBM Systems Director 6.3 Best Practices: Installation & Configuration*, REDP-4932
<http://www.redbooks.ibm.com/abstracts/redp4932.html?Open>
- ▶ IBM Systems Director web page
<http://www.ibm.com/systems/software/director>

The following sections introduce several key plug-ins for IBM Systems Director.

6.9.1 Active Energy Manager

IBM Systems Director Active Energy Manager measures, monitors, and manages the energy components of the eX5 systems. Monitoring functions include power trending, thermal trending, power distribution unit (PDU) support, and the integration of and support of facility providers. Management functions include power capping and power savings mode.

This solution helps customers monitor energy consumption to allow better use of available energy resources. The application software enables customers to trend actual energy consumption and corresponding thermal loading of IBM Systems running in their environment with their applications. Active Energy Manager can help with the following tasks:

- ▶ Allocating less power and cooling infrastructure to IBM servers.
- ▶ Lowering the power usage on select IBM servers.
- ▶ Planning for the future by viewing trends of power usage over time.
- ▶ Determining power usage for all components of a rack.
- ▶ Retrieving temperature and power information through wireless sensors.
- ▶ Collecting alerts and events from facility providers that are related to power and cooling equipment.

You can better understand energy usage across your data center by doing the following tasks:

- ▶ Identifying energy usage.
- ▶ Measuring cooling costs accurately.
- ▶ Monitoring IT costs across components.
- ▶ Managing by department or user.

For more information, see the following sources:

- ▶ *Implementing IBM Systems Director Active Energy Manager 4.1.1*, SG24-7780
<http://www.redbooks.ibm.com/abstracts/sg247780.html?Open>
- ▶ Active Energy Manager web page:
<http://www.ibm.com/systems/software/director/aem>

6.9.2 Tivoli Provisioning Manager for Operating System Deployment

IBM Tivoli Provisioning Manager for Operating System Deployment helps the user provision and deploy operating systems on eX5 servers from a library of images across the network. This tool works within the IBM Systems Director framework as a plug-in, using all of the capabilities of Director to enable simplified server configuration, operating system installation, and firmware updates.

The Tivoli Provisioning Manager for Operating System Deployment offers the following major features:

- ▶ System cloning
Tivoli Provisioning Manager for OS Deployment can capture a target eX5 server and save it as a file-based clone image.
- ▶ Driver injection
Drivers can be added to an image file as it is being deployed to an eX5 server.
- ▶ Software deployment
Any software package can be deployed by using Tivoli Provisioning Manager for OS Deployment.
- ▶ Universal system profile
A single universal system profile can be used to deploy an image to any number of server types through the injection of system-specific drivers during the deployment process.
- ▶ Microsoft Windows Vista support
Tivoli Provisioning Manager for OS Deployment supports deployment of Microsoft Vista.
- ▶ Remote-build capability
Images can be built and deployed by using the Tivoli Provisioning Manager for OS Deployment capability to take over the target server.
- ▶ Unattended setup
All parameters that are required for an installer can be predefined within the software, eliminating the need for a user to enter the data.
- ▶ Unicast and multicast image deployment
A single server or a batch of servers can be deployed by using unicast.
- ▶ Adjustable network bandwidth usage during build
The amount of bandwidth that is used during the image capture and deployment process can be throttled to avoid excessive network congestion.
- ▶ Highly efficient image storage
An algorithm based on MD5 is used to reduce the drive space that is required for storing similar images.
- ▶ Build from DVD
A server can be built from a DVD for instances where the network bandwidth prevents an effective network deployment. An example is in a retail environment at the end of a 64 Kbps link.
- ▶ Boot from CD/DVD
For environments that do not allow or support network boot (PXE), it is possible to build a kickstart CD or DVD to start the deployment process.

- ▶ Network sensitive image replication

Image replication between two separate Tivoli Provisioning Manager for OS Deployment servers can be accomplished through a scheduled replication in which the bandwidth is controlled or a series of command-line export utilities can be used to produce a differences file. This file then can be sent to the subordinate Tivoli Provisioning Manager for OS Deployment server.

- ▶ Redeployment

A hidden partition on the server that can be used to do a full restoration of the reference image through a boot option on the server.

For more information about Tivoli Provisioning Manager for OS Deployment, see the following publications:

- ▶ *Architecting a Highly Efficient Image Management System with Tivoli Provisioning Manager for OS Deployment*, REDP-4294
- ▶ *Vista Deployment Using Tivoli Provisioning Manager for OS Deployment*, REDP-4295
- ▶ *Deploying Linux Systems with Tivoli Provisioning Manager for OS Deployment*, REDP-4323
- ▶ *Tivoli Provisioning Manager for OS Deployment in a Retail Environment*, REDP-4372
- ▶ *Implementing an Image Management System with Tivoli Provisioning Manager for OS Deployment: Case Studies and Business Benefits*, REDP-4513
- ▶ *Deployment Guide Series: Tivoli Provisioning Manager for OS Deployment V5.1*, SG24-7397

Abbreviations and acronyms

AC	alternating current	IBM	International Business Machines
AMM	advanced management module	ID	identifier
APIC	Advanced Programmable Interrupt Controller	IEEE	Institute of Electrical and Electronics Engineers
ASU	Advanced Settings Utility	IMM	integrated management module
ATS	Advanced Technical Support	IOH	I/O hub
BIOS	basic input/output system	IOPS	I/O operations per second
BMC	Baseboard Management Controller	IPMI	Intelligent Platform Management Interface
BOFM	BladeCenter Open Fabric Manager	ISO	International Organization for Standardization
CD	compact disc	IT	information technology
CIM	Common Information Model	ITSO	International Technical Support Organization
CLI	command-line interface	KB	kilobyte
CMOS	complementary metal-oxide semiconductor	KVM	keyboard video mouse kernel-based virtual machine
COD	configuration on disk	LAN	local area network
COG	configuration and option guide	LDAP	Lightweight Directory Access Protocol
CPU	central processing unit	LED	light emitting diode
CRC	cyclic redundancy check	LGA	Large Gate Array
CRU	customer replaceable units	LUN	logical unit number
DAU	Demand Acceleration Units	MAC	media access control
DB	database	MB	megabyte
DDF	Disk Data Format	MCA	Machine Check Architecture
DIMM	dual inline memory module	MESI	modified exclusive shared invalid
DMA	direct memory access	MR	MegaRAID
DPC	deferred procedure call	NAS	network-attached storage
DSA	Dynamic System Analysis	NMI	non-maskable interrupt
ECC	error correction code	NOS	network operating system
ER	enterprise rack	NUMA	non-uniform memory access
EXA	Enterprise X-Architecture	OGF	Open Grid Forum
FAMM	Full Array Memory Mirroring	OS	operating system
FTSS	Field Technical Sales Support	OSD	on screen display
GB	gigabyte	PCI	Peripheral Component Interconnect
GT	giga-transfers	PCIE	PCI Express
HBA	host bus adapter	PDU	power distribution unit
HDD	hard disk drive	PE	preinstallation environment
HPC	high performance computing	PMI	Project Management Institute
HPCBP	High Performance Computing Basic Profile	POST	power-on self-test
HS	hot swap		
HT	Hyper-Threading		
I/O	input/output		

PXE	Preboot Execution Environment	TSM	Technical Support Management
QPI	Quick Path Interconnect	UDP	User Datagram Protocol
RAID	redundant array of independent disks	UE	Unrecoverable Error
RAM	random access memory	UEFI	Unified Extensible Firmware Interface
RAS	reliability, availability, and serviceability	URL	Uniform Resource Locator
RDMA	Remote Direct Memory Access	USB	Universal Serial Bus
RHEL	Red Hat Enterprise Linux	VLAN	virtual LAN
RISC	reduced instruction set computer	VLP	very low profile
RoC	RAID on Chip	VM	virtual machine
RPM	revolutions per minute	VMFS	virtual machine file system
RSA	Remote Supervisor Adapter	VPD	vital product data
RTS	request to send	VT	Virtualization Technology
SAN	storage area network	WWPN	worldwide port name
SAS	serial-attached SCSI		
SATA	Serial ATA		
SCM	Storage Configuration Manager		
SCSI	Small Computer System Interface		
SDRAM	static dynamic RAM		
SED	self-encrypting drive		
SEL	System Event Log		
SFP	small form-factor pluggable		
SIO	Storage and I/O		
SLC	Single Level Cell		
SMI	synchronous memory interface		
SMP	symmetric multiprocessing		
SNMP	Simple Network Management Protocol		
SOA	Service Oriented Architecture		
SPORE	ServerProven Opportunity Request for Evaluation		
SR	short range		
SSCT	Standalone Solution Configuration Tool		
SSD	solid-state drive		
SSIC	System Storage Interoperation Center		
SSL	Secure Sockets Layer		
STG	Systems and Technology Group		
TB	terabyte		
TCG	Trusted Computing Group		
TCO	total cost of ownership		
TCP	Transmission Control Protocol		
TOE	TCP offload engine		
TPM	Trusted Platform Module		

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this paper.

IBM Redbooks

The following IBM Redbooks publications provide additional information about the topic in this document. Note that some publications referenced in this list might be available in softcopy only.

You can search for, view, download, or order these documents and other Redbooks, Redpapers, Product Guides, at the following website:

ibm.com/redbooks

IBM Redbooks Product Guides

- ▶ *IBM System x3850 X5*, TIPS0817
- ▶ *IBM System x3690 X5*, TIPS0818
- ▶ *IBM BladeCenter HX5*, TIPS0824
- ▶ *IBM ServeRAID Adapter Quick Reference*, TIPS0054
- ▶ *ServeRAID M1015 SAS/SATA Controller for System x*, TIPS0740
- ▶ *ServeRAID M5015 and M5014 SAS/SATA Controllers for IBM System x*, TIPS0738
- ▶ *Emulex 10GbE Virtual Fabric Adapter II and III family for IBM System x*, TIPS0844

All Product Guides for System x servers and options can be found at:

<http://www.redbooks.ibm.com/portals/systemx?Open&page=pgbycat>

All Product Guides for BladeCenter servers and options can be found at:

<http://www.redbooks.ibm.com/portals/bladecenter?Open&page=pgbycat>

IBM Redpapers and IBM Redbooks

- ▶ *Reliability, Availability, and Serviceability Features of the IBM eX5 Portfolio*, REDP-4864
- ▶ *The Benefits of Optimizing OLTP Databases Using IBM eXFlash Solid-State Drives*, REDP-4849
- ▶ *Workload Optimization with the IBM eX5 Family of Servers*, REDP-4845
- ▶ *Add Memory, Improve Performance, and Lower Costs with IBM MAX5 Technology*, REDP-4846
- ▶ *Advantages of IBM eX5 for Database Workloads*, REDP-4848
- ▶ *IBM eX5 Technology and WebSphere Produce High-Performance Websites*, REDP-4847
- ▶ *IBM Systems Director 6.3 Best Practices: Installation & Configuration*, REDP-4932
- ▶ *Implementing IBM Systems Director Active Energy Manager 4.1.1*, SG24-7780

- ▶ *Consolidation of Microsoft SQL Server Instances on the IBM System x3850 X5 with Microsoft Hyper-V*, REDP-4661; *Business Intelligence Solutions on the IBM System x3850 X5*, REDP-4662
- ▶ *Consolidating Large Microsoft SQL Server Databases on the IBM System x3850 X5 with Microsoft Hyper-V*, REDP-4690
- ▶ *Business Intelligence Solutions on the IBM System x3850 X5 Using XIV Storage*, REDP-4721
- ▶ *New Business Intelligence Solutions on the IBM System x3850 X5 Using XIV Storage*, REDP-4745

Other publications

Publications listed in this section are also relevant as further information sources.

IBM System x3850 X5 and x3950 X5

Refer to the following publications:

- ▶ *Installation and User's Guide - IBM System x3850 X, x3950 X5*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5085479>
- ▶ *Problem Determination and Service Guide - IBM System x3850 X5, x3950 X5 (7145, 7146, 7143, 7191)*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5084848>
- ▶ *Rack Installation Instructions - IBM System x3850 X5 and x3950 X5*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5085476>
- ▶ *Installation Instructions for the IBM 2-Node x3850 X5 and x3950 X5 Scalability Kit - IBM System x3850 X5, x3950 X5*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-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 - IBM System x3850 X5, x3950 X5*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5084861>

IBM System x3690 X5

Refer to the following publications:

- ▶ *Installation and User's Guide - IBM System x3690 X5 (7147, 7148, 7149, and 7192)*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5085206>
- ▶ *Problem Determination and Service Guide - IBM System x3690 X5 (7147, 7148, 7149, and 7192)*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-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 - IBM System x3690 X5 (7148, 7149)*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5085207>

IBM BladeCenter HX5

Refer to the following publications:

- ▶ *IBM BladeCenter Information Center*
<http://publib.boulder.ibm.com/infocenter/bladectr/documentation>
- ▶ *Installation and User's Guide - IBM BladeCenter HX5*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5084612>
- ▶ *Problem Determination and Service Guide - IBM BladeCenter HX5*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=MIGR-5084529>

Online resources

The following web pages are also relevant as further information sources:

- ▶ IBM System x3850 X5 and x3950 X5 home page
<http://ibm.com/systems/ex5>
- ▶ *Configuration and Options Guide (COG) - IBM BladeCenter and System x*
<http://ibm.com/support/entry/portal/docdisplay?ln docid=SCOD-3ZVQ5W>
- ▶ IBM ServerProven
<http://ibm.com/systems/info/x86servers/serverproven/compat/us>

IBM System x3850 X5 and x3950 X5

See the following web pages:

- ▶ IBM System x3850 X5 and x3950 X5 home page
<http://ibm.com/systems/x/hardware/enterprise/x3850x5>
- ▶ IBM US Announcement letter for the x3850 X5 and x3950 X5 (March 2010)
<http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS110-022>
- ▶ IBM US Announcement letter for the x3850 X5 and MAX5 (May 2010)
<http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS110-108>
- ▶ IBM US Announcement letter for the x3850 X5 and x3950 X5 with MAX5 (April 2011)
<http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS111-055>

IBM BladeCenter HX5

See the following web pages:

- ▶ IBM BladeCenter HX5 home page
<http://ibm.com/systems/bladecenter/hardware/servers/hx5>
- ▶ IBM US Announcement letter for the HX5 (March 30, 2010)
<http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS110-068>

- ▶ IBM US Announcement letter for the HX5 with MAX5 (August 31, 2010)
<http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS110-162>
- ▶ IBM US Announcement letter for the HX5 virtualization optimized server (August 31, 2010)
<http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS110-175>

IBM System x3690 X5

See the following web pages:

- ▶ IBM System x3690 X5 home page
<http://ibm.com/systems/x/hardware/enterprise/x3690x5>
- ▶ IBM US Announcement letter for the x3690 X5 (July 6, 2010)
<http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS110-121>
- ▶ IBM US Announcement letter for the x3690 X5 virtualization optimized server with MAX5 (August 31, 2010)
<http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS110-181>
- ▶ IBM US Announcement letter for the x3690 X5 (April 6, 2011)
<http://ibm.com/common/ssi/cgi-bin/ssialias?infotype=dd&subtype=ca&&htmlfid=897/ENUS111-056>

How to get Redbooks

You can search for, view, or download Redbooks, Redpapers, Technotes, draft publications, and additional materials, and order hardcopy Redbooks publications, at this website:

ibm.com/redbooks

Help from IBM

IBM Support and downloads

ibm.com/support

IBM Global Services

ibm.com/services



IBM eX5 Portfolio Overview

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



Redpaper™

Introduction to the complete IBM eX5 family of servers

Detailed information about each server and its options

Scalability, partitioning, and systems management details

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. 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 portfolio of high-end computing introduces the fifth generation of IBM X-Architecture technology. The X5 portfolio is 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 IBM Redpaper publication introduces the new IBM eX5 portfolio and describes the technical detail behind each server. This document is intended for potential users of eX5 products that are seeking more information about the portfolio.

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