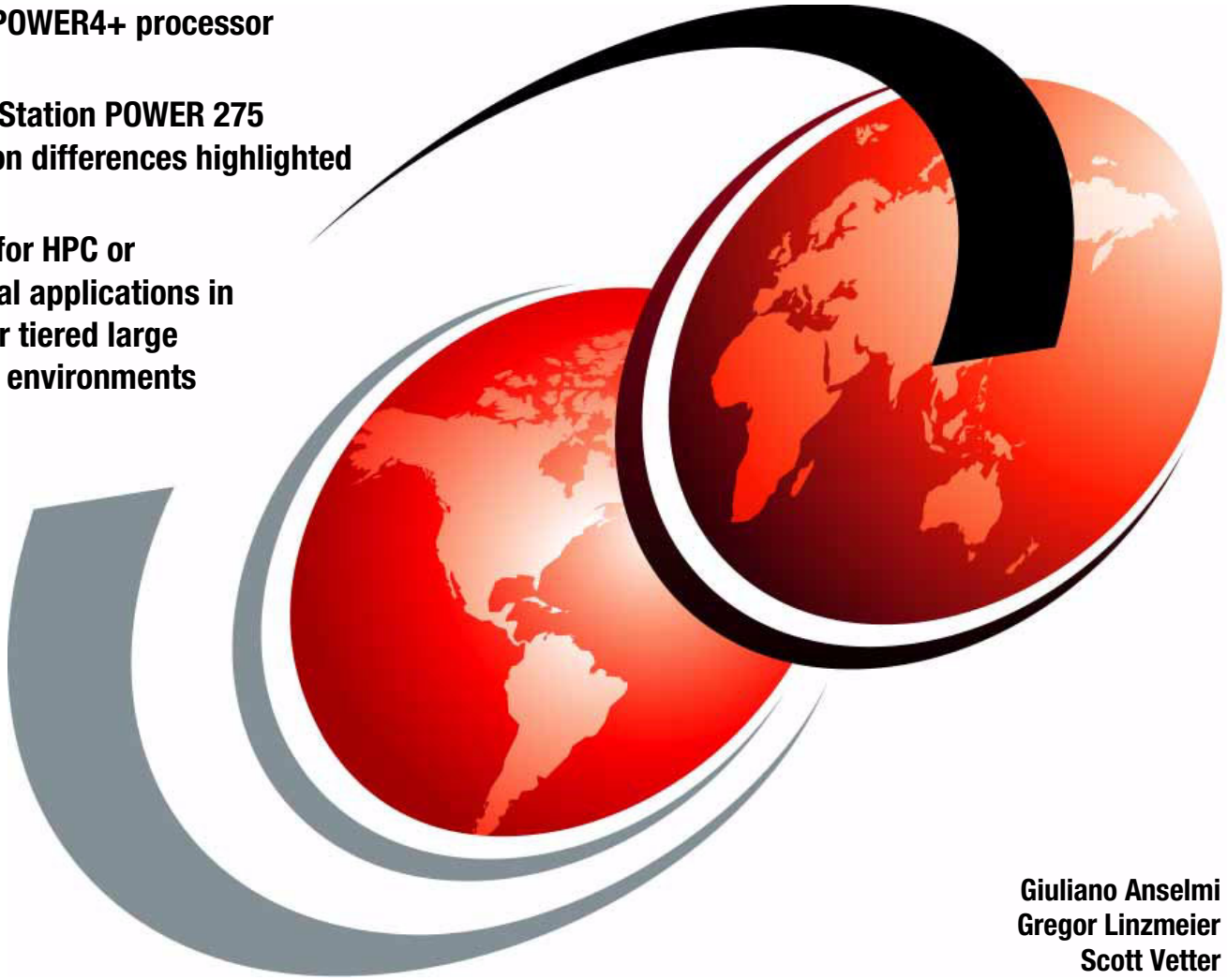
IBM @server

IBM

pSeries 615 Models 6C3 and 6E3 Technical Overview and Introduction

- Affordable servers based on the powerful POWER4+ processor
- IBM IntelliStation POWER 275 workstation differences highlighted
- Designed for HPC or commercial applications in the SMB or tiered large enterprise environments



Giuliano Anselmi
Gregor Linzmeier
Scott Vetter

Redpaper

ibm.com/redbooks



International Technical Support Organization

IBM @server pSeries 615 Models 6C3 and 6E3 Technical Overview and Introduction

October 2003

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

Third Edition (October 2003)

This edition applies to the IBM @server® pSeries® 615 Models 6C3 and 6E3 and AIX 5L™, product number 5765-E62.

© Copyright International Business Machines Corporation 2003. 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 v
Trademarks vi

Preface vii
The team that wrote this Redpaper vii
Become a published author viii
Comments welcome viii

Chapter 1. General description 1
1.1 System specification 2
1.2 Physical packages 2
1.2.1 Models 6C3 and 6E3 2
1.3 Minimum and optional features 3
1.3.1 Express configurations 5
1.4 Model types 6
1.4.1 Model 6C3 rack-mounted server 7
1.4.2 Model 6E3 deskside server 7
1.4.3 IBM IntelliStation POWER 275 workstation 7
1.5 System racks 8
1.5.1 IBM RS/6000 7014 Model T00 Enterprise Rack 8
1.5.2 IBM RS/6000 7014 Model T42 Enterprise Rack 9
1.5.3 AC Power Distribution Units for rack models T00 and T42 9
1.5.4 OEM racks 10
1.5.5 Rack-mounting rules for Model 6C3 10
1.5.6 Cable management arm 11
1.5.7 Flat-panel display options 11
1.5.8 VGA switch and extender cables 12
1.6 Cluster 1600 support 12

Chapter 2. Architecture and technical overview 13
2.1 Processor and cache 14
2.1.1 L1, L2, and L3 cache 15
2.1.2 PowerPC architecture 15
2.1.3 Copper and CMOS technology 15
2.1.4 Processor clock rate 16
2.2 Memory 16
2.2.1 Memory options 17
2.2.2 Memory restrictions (OEM) 17
2.3 System buses 18
2.3.1 GX bus 18
2.3.2 PCI-X host bridge and PCI-X bus 18
2.4 PCI-X slots and adapters 18
2.4.1 64-bit and 32-bit adapters 19
2.4.2 LAN adapters 19
2.4.3 Graphic accelerators 19
2.4.4 Audio adapter 20
2.4.5 Internal Ultra320 SCSI controllers 20
2.5 Internal storage 20
2.5.1 Hot-swappable SCSI disks 22

© Copyright IBM Corp. 2003. All rights reserved.

2.5.2 Other hot-plug options . . . 23
2.5.3 Preferred boot device and options . . . 23
2.6 System Management Services (SMS) . . . 24
2.7 Security . . . 26
2.8 Operating system requirements . . . 26
2.8.1 AIX 5L . . . 26
2.8.2 Linux . . . 26

Chapter 3. Availability, investment protection, expansion, and accessibility . . . 29
3.1 Autonomic computing . . . 30
3.2 HMC management . . . 31
3.3 Reliability, availability, and serviceability (RAS) features . . . 31
3.3.1 Service processor . . . 32
3.3.2 Memory reliability, fault tolerance, and integrity . . . 34
3.3.3 First failure data capture, diagnostics, and recovery . . . 34
3.3.4 Error indication and LEDs indicators . . . 35
3.3.5 Dynamic or persistent deallocation . . . 36
3.3.6 UE-Gard . . . 38
3.3.7 System Power Control Network (SPCN), power supplies, and cooling . . . 38
3.3.8 Early Power-Off Warning (EPOW) . . . 39
3.3.9 Service Agent and Inventory Scout . . . 40
3.3.10 IBM @server pSeries Customer-Managed Microcode . . . 40

Related publications . . . 43
IBM Redbooks . . . 43
Other publications . . . 43
Online resources . . . 44
How to get IBM Redbooks . . . 45

Notices

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

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

IBM may have patents or pending patent applications covering subject matter described in this document. The furnishing of this document does not give you any license to these patents. You can send license inquiries, in writing, to:
IBM Director of Licensing, IBM Corporation, North Castle Drive Armonk, NY 10504-1785 U.S.A.

The following paragraph does not apply to the United Kingdom or any other country where such provisions are inconsistent with local law: INTERNATIONAL BUSINESS MACHINES CORPORATION PROVIDES THIS PUBLICATION "AS IS" WITHOUT WARRANTY OF ANY KIND, EITHER EXPRESS OR IMPLIED, INCLUDING, BUT NOT LIMITED TO, THE IMPLIED WARRANTIES OF NON-INFRINGEMENT, MERCHANTABILITY OR FITNESS FOR A PARTICULAR PURPOSE. Some states do not allow disclaimer of express or implied warranties in certain transactions, therefore, this statement may not apply to you.

This information could include technical inaccuracies or typographical errors. Changes are periodically made to the information herein; these changes will be incorporated in new editions of the publication. IBM may make improvements and/or changes in the product(s) and/or the program(s) described in this publication at any time without notice.

Any references in this information to non-IBM Web sites are provided for convenience only and do not in any manner serve as an endorsement of those Web sites. The materials at those Web sites are not part of the materials for this IBM product and use of those Web sites 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 measurement 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 illustrates 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. 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 IBM's application programming interfaces.

© Copyright IBM Corp. 2003. All rights reserved.

Trademarks

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

AIX®
AIX 5L™
AS/400®
Chipkill™
Electronic Service Agent™
Enterprise Storage Server™

IBM®
IBMLink™
IntelliStation®
iSeries™
POWER4™
POWER4+™
PowerPC®
pSeries®
Redbooks™
Redbooks(logo)™ 
RS/6000®
Service Director™

The following terms are trademarks of other companies:

ActionMedia, LANDesk, MMX, Pentium and ProShare are trademarks of Intel Corporation in the United States, other countries, or both.

Microsoft, Windows, Windows NT, 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 Sun Microsystems, Inc. in the United States, other countries, or both.

C-bus is a trademark of Corollary, Inc. in the United States, other countries, or both.

UNIX is a registered trademark of The Open Group in the United States and other countries.

SET, SET Secure Electronic Transaction, and the SET Logo are trademarks owned by SET Secure Electronic Transaction LLC.

Preface

This document is a comprehensive guide covering the IBM @server® pSeries® 615 Models 6C3 and 6E3 entry-level UNIX® servers. Major hardware offerings are introduced and their prominent functions are discussed. A section highlighting the IBM IntelliStation® POWER 275 workstation is included.

Professionals wishing to acquire a better understanding of IBM @server pSeries products may consider reading this document. The intended audience includes:

- Customers
- Sales and marketing professionals
- Technical support professionals
- IBM Business Partners
- Independent software vendors

This document expands the current set of IBM @server pSeries documentation by providing a desktop reference that offers a detailed technical description of the pSeries 615.

This publication does not replace the latest pSeries marketing materials and tools. It is intended as an additional source of information that, together with existing sources, may be used to enhance your knowledge of IBM server solutions.

The team that wrote this Redpaper

This Redpaper was produced by a team of specialists from around the world working at the International Technical Support Organization, Austin Center. The team project manager was Scott Vetter.

Giuliano Anselmi is an IBM pSeries systems product engineer at the EMEA Midrange Field Support Organization in Rome, Italy, supporting the Web Server Sales Organization in EMEA, IBM sales, IBM Business Partners, IBM technical support organizations, and IBM Dublin Enterprise Servers manufacturing. He is a resolution assistant resource for critical situations and related to RS/6000® and pSeries systems for the pSeries Customer Satisfaction Project Office. He has been devoted to RS/6000 and pSeries systems for 11 years.

Gregor Linzmeier is an IBM Advisory IT Specialist for RS/6000 and pSeries workstation and entry servers as part of the Web Server Sales Organization in Mainz, Germany. Gregor is a member of the Workstation Technology Focus Group, supporting IBM sales, Business Partners, and customers with pre-sales consultation and implementation of client/server environments. He has worked for more than 13 years as an infrastructure specialist for RT, RS/6000, and AIX in large CATIA client/server projects.

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

IBM USA: Richard W. Bishop, John Bissell, Russell Bistline, Tom Campbell, Dan DeLapp, Richard B. Hartin, John Hilburn, Noah G. Kuttler, Tenley Jackson, Jim McGaughan, Andy McLaughlin, Dennis Lee, Bill Mihaltse, Thoi Nguyen, Jan Palmer, Scott Smylie, Asghar Tavasoli, Anthony Thai, Harold W. Washington, Sterling Weaver, Abel Zuzuarregui.

IBM Brasil: Marcelo Tsuguio Okano

© Copyright IBM Corp. 2003. All rights reserved.

IBM Germany: Volker Haug

Become a published author

Join us for a two- to six-week residency program! Help write an IBM Redbook dealing with specific products or solutions, while getting hands-on experience with leading-edge technologies. You'll team with IBM technical professionals, Business Partners and/or customers.

Your efforts will help increase product acceptance and customer satisfaction. As a bonus, you'll develop a network of contacts in IBM development labs, and increase your productivity and marketability.

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 Redpaper or other Redbooks in one of the following ways:

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

  ibm.com/redbooks

- Send your comments in an Internet note to:

  redbook@us.ibm.com

- Mail your comments to:

  IBM Corporation, International Technical Support Organization
  Dept. JN9B Building 003 Internal Zip 2834
  11400 Burnet Road
  Austin, Texas 78758-3493

1

General description

The IBM @server pSeries 615 Models 6C3 and 6E3 (referred to hereafter as the Model 6C3 and Model 6E3, or p615 when discussing both models) are designed for customers looking for a cost-effective, high-performance, space-efficient server that uses advanced IBM technology first available in the high-end pSeries 690, then on midrange product p650, and finally in the entry-level p630. The p615 uses the 64-bit, copper and SOI-based POWER4+™ microprocessor, and is intended for use in stand-alone or LAN clustered environments.

The p615 is a member of the 64-bit family of symmetric multiprocessing (SMP) UNIX servers from IBM. The Model 6C3 (product number 7029-6C3) is a 4-EIA[1] (4U), 19-inch rack-mounted server, while the Model 6E3 (product number 7029-6E3) is a deskside server. The p615 can be configured as a 1-, or 2-way system running at 1.2 GHz or as a 2-way system running at 1.45 GHz. At the time of writing, the total system memory can range from 1 GB to 16 GB based on the available DIMMs. The availability of cluster support, see Section 1.6, “Cluster 1600 support”, enhances the already exceptional value of these models.

The p615 includes six hot-plug PCI-X slots, an integrated dual channel Ultra320 SCSI controller, one 10/100 Mbps plus one 10/100/1000 Mbps integrated Ethernet controller, and eight front-accessible disk bays supporting hot-swappable disks. These disk bays provide greater system availability and growth by allowing the removal or addition of disk drives without disrupting service, and can accommodate up to 1.17 TB of disk storage using 146.8 GB Ultra3 SCSI disk drives. The availability of internal RAID using a daughter card installed on the system planar and Ultra320 disk and disk-bay features further increases internal storage options.

Three media bays are available. One full-height media bay is used for a tape drive or DVD-RAM. Two media bays only accept new slimline media devices, the same design used in some IBM ThinkPad models, to accommodate a slim DVD-ROM or the optional diskette drive. The Converged Service Processor[2] (CSP) is on a dedicated card plugged into the main system planar, which is designed to continuously monitor system operations, taking preventive or corrective actions for quick problem resolution and high system availability.

[1] One Electronic Industries Association Unit (EIA) is 44.45 mm (1.75 in.).
[2] Since the release of 7025 Model F80, 7026 Model H80, and M80, the RS/6000 (pSeries) Service Processor design converged to AS/400 (iSeries) Service Processor design.

© Copyright IBM Corp. 2003. All rights reserved.

Additional reliability and availability features include redundant hot-plug cooling fans and optionally redundant power supply. Along with these hot-plug components, the p615 is designed to provide an extensive set of reliability, availability, and serviceability (RAS) features that include improved fault isolation, recovery from errors without stopping the system, avoidance of recurring failures, and predictive failure analysis. See Chapter 3, "Availability, investment protection, expansion, and accessibility" on page 29, for more information on RAS features.

1.1 System specification

Table 1-1 lists the general system specifications of the p615.

Table 1-1 pSeries p615 specifications

Description	Range
Operating temperature	10 to 43 degrees C (50 to 109 F)
Relative humidity	8% to 80%
Operating voltage	100 to 127 or 200 to 240 V ac (auto-ranging)
Operating frequency	47/63 Hz
Power consumption in active mode	450 watts maximum
Power consumption in energy savings mode	300 watts
Thermal output	400 Btu[a]/hour (typical); 794 Btu/hour (maximum)
[a]British Thermal Unit (BTU)	

1.2 Physical packages

The following sections discuss the major physical attributes found on the p615 in rack-mounted and deskside versions as shown in Table 1-2. The Model 6C3 is a 4U, 19-inch rack-mounted system, and the Model 6E3 is a deskside configuration.

Table 1-2 Physical packaging of the p615

Dimensions	Rack (Model 6C3)	Deskside (Model 6E3)
Height	178 mm (7.0 in.)	533 mm (21.0 in.)
Width	437 mm (17.2 in.)	201 mm (7.9 in.)
Depth	508 mm (20.0 in.)	584 mm (23.0 in.)
Weight		
Minimum configuration	35.5 Kg (78.0 lbs.)	35.5 Kg (78.0 lbs.)
Maximum configuration	43.1 Kg (94.8 lbs.)	43.1 Kg (94.8 lbs.)

1.2.1 Models 6C3 and 6E3

Figure 1-1 on page 3 shows a detailed view of the Model 6C3 and Model 6E3 including the locations of all various components and devices.

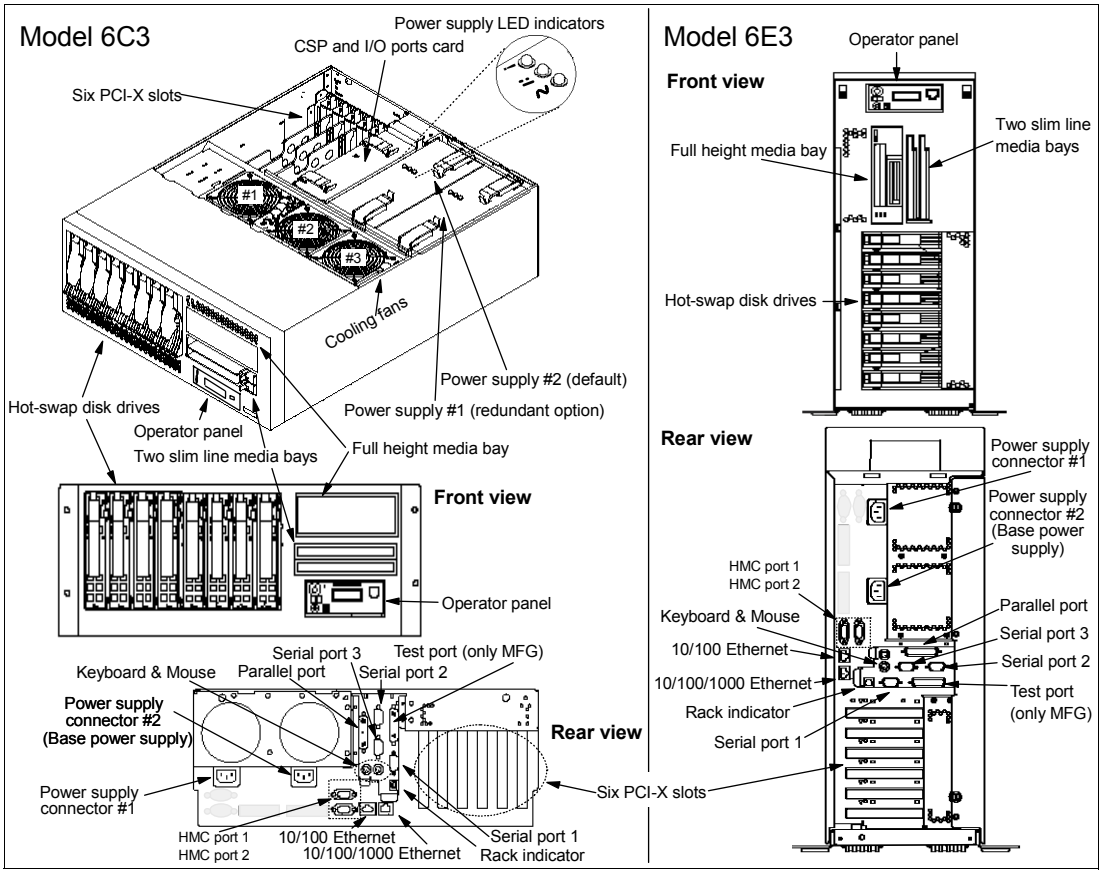


Figure 1-1 Views of Models 6C3 and 6E3

1.3 Minimum and optional features

The following sections provide a summary of the configurable feature codes for the p615.

Processor and memory

The following are the minimum optional features for the processor and memory system:

- 1- or 2-way SMP system
  - 1- or 2-way, 64-bit, copper-based POWER4+ microprocessors running at 1.2 GHz, with a total of 8 MB L3[3] cache.
    - 1-way system planar (FC 5220, or Express config[4] FC 8148)
    - 2-way system planar (FC 5222, or Express config FC 8149)
  - 2-way, 64-bit copper-based POWER4+ microprocessors running at 1.45 GHz, with total of 8 MB shared L3 cache. (FC 5234 or Express config FC 8187)

Note: At the time of writing, there is no upgrade path from a 1-way to a 2-way system.

- 1 GB to 16 GB ECC DDR SDRAM memory

[3] Level3 (L3)
[4] An Express configuration is a predefined set of standard features that can be upgraded later if needed.

– Memory DIMMs plug into the eight slots on the system planar.
– DIMMs must be populated in quads (four DIMMs). A memory feature consists of a quad. Additional quads may consist of any memory feature code (memory size). See Table 1-3 for the memory feature codes available (for standard or Express configuration orders) at the time this documentation was written.

Table 1-3 Memory features

Express feature code	Description
FC 8156	1024 MB (4 x 256 MB) DIMMs, 208-pin 8 ns DDR SDRAM
FC 8157	2048 MB (4 x 512 MB) DIMMs, 208-pin 8 ns DDR SDRAM
FC 8158	4096 MB (4 x 1024 MB) DIMMs, 208-pin 8 ns DDR SDRAM
FC 8159	8192 MB (4 x 2048 MB) DIMMs, 208-pin 8 ns DDR SDRAM
Feature code	**Description**
FC 4446	1024 MB (4 x 256 MB) DIMMs, 208-pin 8 ns DDR SDRAM
FC 4447	2048 MB (4 x 512 MB) DIMMs, 208-pin 8 ns DDR SDRAM
FC 4448	4096 MB (4 x 1024 MB) DIMMs, 208-pin 8 ns DDR SDRAM
FC 4449	8192 MB (4 x 2048 MB) DIMMs, 208-pin 8 ns DDR SDRAM

Hot-swappable disk drives inside the p615

The p615 contains one default 4-pack DASD backplane. It is possible to add an optional 4-pack DASD backplane to accommodate a total of eight hot-swappable disks. Two backplanes are available. FC 6574 uses the integrated Ultra320 SCSI controller. FC 6584 can be connected to a Ultra320 SCSI adapter using cable FC 4267, to provide a flexible internal disk solution. This configuration is helpful in solutions where disks must be split between two physically independent controllers. Each bay may contain one of the following features, as provided in Table 1-4. For Ultra320 speeds, Ultra320 disks purchased prior to September 01, 2003 require a microcode upgrade to operate at Ultra320 speeds.

Table 1-4 Disk drive features

Feature code	Description
FC 3273	36.4 GB Ultra3 10K RPM
FC 3277	36.4 GB Ultra3 15K RPM
FC 3274	73.4 GB Ultra3 10K RPM
FC 3278	73.4 GB Ultra3 15K RPM
FC 3275	146.8 GB Ultra3 10K RPM

The different cabling options are discussed in the Section 2.5, “Internal storage” on page 20.

Internal RAID

The p615 planar contains a slot for an optional Dual Channel SCSI RAID Enablement Card (FC 5709). This card provides RAID 0, 5, or 10 function to the integrated SCSI controller and includes a 16 MB fast write cache to improve data flow. Internal RAID is also possible using a RAID adapter in a PCI slot connected to the optional FC 6584 backplane. Each 4-pack backplane operates as an independent RAID array when featured for RAID.

Media bays

The p615 contains a total of four media bays, with three free bays: 2, 3, and 4. Media bay 1 is used by the operator panel by default. The p615 media backplane is plugged into the system planar and provides connectors for one converged operator panel[5], two slimline devices (diskette or IDE optical devices), and one full-height SCSI device. The devices installed in the media bays are not hot-pluggable.

- Media bays 2 and 3 can accommodate the slimline diskette drive (FC 2611) or slimline DVD-ROM drive (FC 2640).
- Media bay 4 can accommodate the DVD-RAM (FC 2623) or 4 mm 20/40 GB tape drive (FC 6158) as full-height SCSI devices, the half-height 80/160 GB tape drive (FC 6120), or half-height 60/150 GB tape drive (FC 6134).

PCI-X adapter slots

The p615 provides six available adapter slots as part of the basic configuration:

- Four PCI-X 64-bit slots operating at 133 MHz
  - Slot 1 accepts a select set of long adapters. Check the *RS/6000 and pSeries PCI Adapter Placement Reference*, SA38-0538 to see if a long adapter is compatible.
  - Slot 4 accepts long adapters and operates at 133 MHz
  - Slots 5 and 6 share a bus with other components, see 2.3.2, “PCI-X host bridge and PCI-X bus” on page 18 for more information.
- Two PCI-X 32-bit slots (slots 2 and 3) operating at 66 MHz

I/O ports

The p615 provides several native I/O ports as part of the basic configuration:

- Two integrated Ethernet ports (RJ45) running at 10/100 Mbps and 10/100/1000 Mbps.
- Three serial ports (RS232). Serial port 1 is mirrored between two ports on the Model 6C3: one in the front and one in the rear. The serial port 1 connector in the front is a 10-pin RJ, and the serial port 1 connector in the rear is a standard 9-pin male D-shell (DB9). The front connector has higher priority than the rear connector. If a serial peripheral device is attached to the front connector, the rear connector of serial port 1 becomes disabled regardless of the presence of a peripheral device attached to this port.
- One IEEE 1284 standard parallel printer port with 25-pin female D-shell (DB-25S).
- One PS2 keyboard port and one PS2 mouse port. The keyboard and mouse connector are standard 6-pin DIN.
- Two HMC RS-232 serial ports using a standard 9-pin D-shell connector that provides a hardware connection between a Hardware Management Console (HMC) and the p615. The HMC console is an optional feature. For detailed information about supported basic functions for the p615 see Section 3.2, “HMC management” on page 31.
- Two Ultra320 integrated SCSI controller ports for internal I/O (see Section 2.4.5, “Internal Ultra320 SCSI controllers” on page 20). External SCSI ports are available through optional SCSI adapters.

1.3.1 Express configurations

The p615 is only available through a set of Express configurations. The Express configurations consist of a predefined set of standard features, having in common one default power supply (FC 6266) and the default 4-pack DASD backplane (FC 6574), and must

[5] The converged operator panel is a common part between pSeries and iSeries products.

include the hardware features listed in the Table 1-5. You may make MES upgrades of any Express configuration using the standard features as needed, such as the FC 5709 Dual Channel SCSI RAID Enablement Card.

Table 1-5 Express configuration feature codes

Express Configuration	Express FC	CPU FC	Memory FC	Disk FC	AIX and License	LINUX-ready[1]
110C or 110E	8167	8148	8156	3273	5005, yes	No
111C or 111E	8160	8148	8156	3273	No	Yes
120C or 120E	8169	8148	8157	3273	5005, yes	No
121C or 121E	8161	8148	8157	3273	No	Yes
140C or 140E	8170	8148	8158	3273	5005, yes	No
141C or 141E	8162	8148	8158	3273	No	Yes
180C or 180E	8171	8148	8159	3273	5005, yes	No
181C or 181E	8163	8148	8159	3273	No	Yes
220C or 220E	8172	8149	8157	2 x 3273	5005, yes	No
221C or 221E	8164	8149	8157	2 x 3273	No	Yes
240C or 240E	8173	8149	8158	2 x 3273	5005, yes	No
241C or 241E	8165	8149	8158	2 x 3273	No	Yes
280C or 280E	8175	8149	8159	2 x 3273	5005, yes	No
281C or 281E	8166	8149	8159	2 x 3273	No	Yes
224C or 224E	8188	8187	8157	2 x 3273	5005, yes	No
225C or 225E	8191	8187	8157	2 x 3273	No	Yes
244C or 244E	8189	8187	8158	2 x 3273	5005, yes	No
245C or 245E	8192	8187	8158	2 x 3273	No	Yes
284C or 284E	8190	8187	8159	2 x 3273	5005, yes	No
285C or 285E	8193	8187	8159	2 x 3273	No	Yes

[1]Operating system not included. Customer must license and install a supported Linux distribution.

1.4 Model types

Figure 1-2 shows a Model 6E3 and a Model 6C3.



Figure 1-2 Models 6E3 (left) and 6C3 (right) physical packaging

1.4.1 Model 6C3 rack-mounted server

The Model 6C3 is a 4U rack-mounted server and is intended to be installed in a 19-inch rack, thereby enabling efficient use of computer room floor space. If the IBM 7014 T42 rack is used to mount the Model 6C3, then it is possible to place up to 10 systems in an area of 644 mm (25.5 inches) x 1147 mm (45.2 inches). Before installing any system, a complete review of the *Site and Hardware Planning Guide*, SA38-0508 is recommended to review service clearances and other installation criteria.

Included with the Model 6C3 rack-mounted server packaging are all of the components and instructions necessary to enable installation in a 19-inch rack using suitable tools. The Model 6C3 is designated as a customer setup system and requires three persons (due to weight and safety issues) to be present to install the unit into the rack.

Each system is delivered preconfigured as ordered. In the case of the p615 Linux ready Express Configurations, the systems are delivered without an operating system. You can order SuSE SLES 8 at the time of purchase. The order will be fulfilled by SuSE. The system is designed to be set up by the customer and, in most cases, will not require the use of any tools. Full setup instructions are included with the system.

The GXT135P 2D graphics accelerator with analog and digital interfaces (FC 2849) is available and is supported for SMS, firmware menus, and other low-level functions as well as when AIX starts the X11-based graphical user interface. Graphical AIX system tools are usable for configuration management, if the adapter is connected to a common maintenance console, such as the 7316-TF2 rack-mounted flat-panel display.

1.4.2 Model 6E3 deskside server

The Model 6E3 is a deskside server, ideal for environments requiring the user to have local access to the machine. A typical example of this would be applications requiring a native graphics display.

Each system is delivered preconfigured as ordered. In the case of the p615 Linux ready Express Configurations, the systems are delivered without an operating system. You can order SuSE SLES 8 at the time of purchase. The order will be fulfilled by SuSE. The system is designed to be set up by the customer and, in most cases, will not require the use of any tools. Full set-up instructions are included with the system.

The GXT135P 2D graphics accelerator with analog and digital interfaces (FC 2849) is available and is supported for SMS, firmware menus, and other low-level functions, as well as when AIX starts the X11-based graphical user interface. Graphical AIX system tools are usable for configuration management if the adapter is connected to the primary console, such as the IBM L200P flat-panel monitor (FC 3636) or the IBM T541H 15-inch TFT Color Monitor (FC 3637).

1.4.3 IBM IntelliStation POWER 275 workstation

The IBM IntelliStation POWER 275 workstation is the first IBM workstation based on the POWER4+, 64-bit, SMP microprocessor. It is particularly well suited for high-end Mechanical Computer Aided Design (MCAD), graphic processing, and other floating-point-intensive business and technical applications.

The POWER 275 extracts the basic platform architecture from the p615, and then enhances it in a couple of key areas to further boost its performance and make it a leading choice in UNIX-based workstations:

- The POWER 275 is orderable as a one- or a two-way system with 1.45 GHz POWER4+ processors or a one-way system with a 1.0 GHz POWER4+ processor. All processors include 8 MB of L3 cache. The processor frequencies offered on the POWER 275 are not available on the p615.
- The POWER 275 does not support hot-plug PCI-X found on the p615. This boosts performance of graphics intensive applications.
- The POWER 275 comes standard with features not included in p615 configurations. These features cannot be dropped on an initial POWER 275 order.
  - DVD-ROM (IDE)
  - Diskette drive
  - Audio adapter (occupying one 32-bit slot)
  - POWER GXT4500P graphics accelerator (occupying one 64-bit slot). The GXT6500P which adds on-board lighting and geometry processing is available at additional cost.
- To better integrate the POWER 275 in acoustically restrictive office environments, a maximum of four disk drives are supported in a single SCSI 4-pack.

With the exception of the hot-plug PCI, all of the autonomic and integrated features found on the p615 are included, plus many of the optional features available through MES. At the time of writing, there is no support for Linux on the POWER 275.

The following major features, available on the Model 6E3, are not supported on the POWER 275:

- FC 4961 4-port Ethernet Adapter
- FC 2732 Serial HPPI Adapter
- FC 4256 Extender Cable for USB Keyboards
- FC 6574 or FC 6584 Second SCSI 4-Pack

1.5 System racks

The following description provides an overview of racks available from IBM in which the Model 6C3 can be mounted. There is no feature available to convert a deskside/desktop model to a rack-mounted model.

The Enterprise Rack Models T00 and T42 are 19-inch-wide racks for general use with pSeries and RS/6000 rack-based or rack drawer-based systems. The rack provides increased capacity, greater flexibility, and improved floor space utilization.

If a pSeries or RS/6000 system is to be installed in a non-IBM rack or cabinet, you must ensure that the rack to be used conforms to the EIA standard EIA-310-D. Refer to section 1.5.4, “OEM racks” on page 10.

It is the customer’s responsibility to ensure that the installation of the drawer in the preferred rack or cabinet results in a configuration that is stable, serviceable, safe, and compatible with the drawer requirements for power, cooling, cable management, weight, and rail security.

1.5.1 IBM RS/6000 7014 Model T00 Enterprise Rack

The 1.8-meter (71 in.) Model T00 is compatible with past and present pSeries and RS/6000 racks, and is designed for use in all situations that have previously used the older rack models R00 and S00. The T00 rack has the following features:

- 36 EIA units (36U) of usable space.
- Optional removable side panels.
- Optional highly perforated front door.
- Optional side-to-side mounting hardware for joining multiple racks.
- Increased power distribution and weight capacity.
- Standard black or optional white color in OEM format.
- Optional reinforced (ruggedized) rack feature providing added earthquake protection with modular rear brace, concrete floor bolt-down hardware, and bolt-in steel front filler panels.
- Optional rack status beacon (FC 4690). This beacon is designed to be placed on top of a rack and cabled to servers such as the p615. Servers can be programmed to illuminate the beacon in response to any detected problems or changes in system status.

  To use this beacon, a rack status beacon junction box (FC 4693) must be selected to connect multiple servers to the beacon. This feature provides six input connectors and one output connector for the rack. To connect the servers or other components to the junction box or the junction box to the rack, status beacon cables (FC 4691) are necessary. Multiple junction boxes can be linked together in a series using daisy-chain cables (FC 4692).
- Support of both AC and DC configurations.
- Up to six PDUs[6]. See Section 1.5.3, “AC Power Distribution Units for rack models T00 and T42”.
- DC rack height is increased to 1926 mm (75.8 in.) due to the presence of the power distribution panel fixed to the top of the rack.
- Weight:
  - T00 base empty rack: 244 kg (535 lb.)
  - T00 full rack: 816 kg (1795 lb.)

1.5.2 IBM RS/6000 7014 Model T42 Enterprise Rack

The 2.0-meter (79.3-in.) Model T42 is the rack that will address the special requirements of customers who want a tall enclosure to house the maximum amount of equipment in the smallest possible floor space. The features that differ in the Model T42 rack from the Model T00 include the following:

- 42 EIA units (42U) of usable space
- AC power support only
- Weight:
  - T42 base empty rack: 261 kg (575 lb.)
  - T42 full rack: 930 kg (2045 lb.)

1.5.3 AC Power Distribution Units for rack models T00 and T42

For rack models T00 and T42, six outlet PDUs and nine outlet PDUs are available.

Previously, the only AC PDUs for these racks were PDUs with six outlets (FC 9171, 9173, 9174, 6171, 6173, and 6174). Four PDUs can be mounted vertically and two additional PDUs horizontally on the bottom of the rack requiring 1U of rack space for each horizontally

[6] Power Distribution Unit (PDU)

mounted PDU. Each PDU requires a separate AC supply. These PDUs provide a maximum of 36 power outlets.

In addition to the six outlet PDUs, other PDUs with nine outlets (FC 9176, 9177, 9178, 7176, 7177, and 7178) are available. A T42 rack configured for the maximum number of power outlets would have six PDUs (two mounted horizontally requiring 2U of rack space), for a total of 54 power outlets. T00 racks do not allow horizontal placement of these PDUs and are therefore limited to a total of four nine-outlet PDUs or 36 power outlets.

The Model 6C3 can be connected to any PDU (old style or new style) that is available for the 7014-T00 or T42 racks.

1.5.4 OEM racks

The Model 6C3 can be installed in a suitable OEM rack, provided that the rack conforms to the EIA-310-D standard. This standard is published by the Electrical Industries Alliance, and a summary of this standard is available in the publication *Site and Hardware Planning Information*, SA38-0508. An online copy of this document can be found at:

`http://www.ibm.com/servers/eserver/pseries/library/hardware_docs`

The key points mentioned in this standard are as follows:

- Any rack used must be capable of supporting 15.9 kg (35 lb.) per EIA unit (44.5 mm [1.75 in.] of rack height).
- To ensure proper rail alignment, the rack must have mounting flanges that are at least 494 mm (19.45 in.) across the width of the rack and 719 mm (28.3 in.) between the front and rear rack flanges.
- It may be necessary to supply additional hardware, such as fasteners, for use in some manufacturers' racks.

Figure 1-3 shows the drawing of non-pSeries rack specifications for reference.

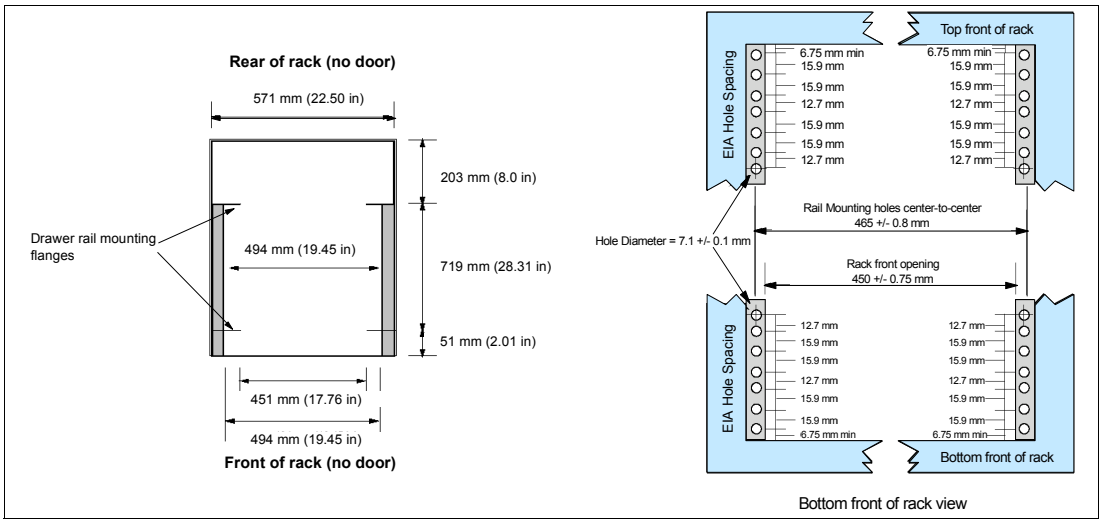


Figure 1-3 Reference drawing of non-pSeries rack specifications

1.5.5 Rack-mounting rules for Model 6C3

The primary rules that should be followed when mounting the Model 6C3 into a rack are:

- The Model 6C3 is designed to be placed at any location in the rack. For rack stability, it is advisable to start filling a rack from the bottom.
- Any remaining space in the rack can be used to install other systems or peripherals, provided that the maximum permissible weight of the rack is not exceeded and the installation rules for these devices are followed.
- Before placing a Model 6C3 into the service position, it is essential that the rack manufacturer's safety instructions have been followed regarding rack stability.
- A Model 6C3 is 4U in height, so a maximum of nine Model 6C3s fit in a T00 rack, or ten Model 6C3s in a T42 rack.

1.5.6 Cable management arm

The Model 6C3 is shipped by default with a cable management arm. You can route all cables through the cable management arm, attaching the cables to the arm with the hook and loop fastener wraps provided. The wraps are provided for additional cable management.
Figure 1-4 on page 11 shows the cable management arm. As the drawer is moved on its rails into the service position, the arms expand or collapse, actively managing the cables. A cable retention clip located on the rear of the server provides a reliable anchor for the cables so that the drawer movement does not accidentally detach them.

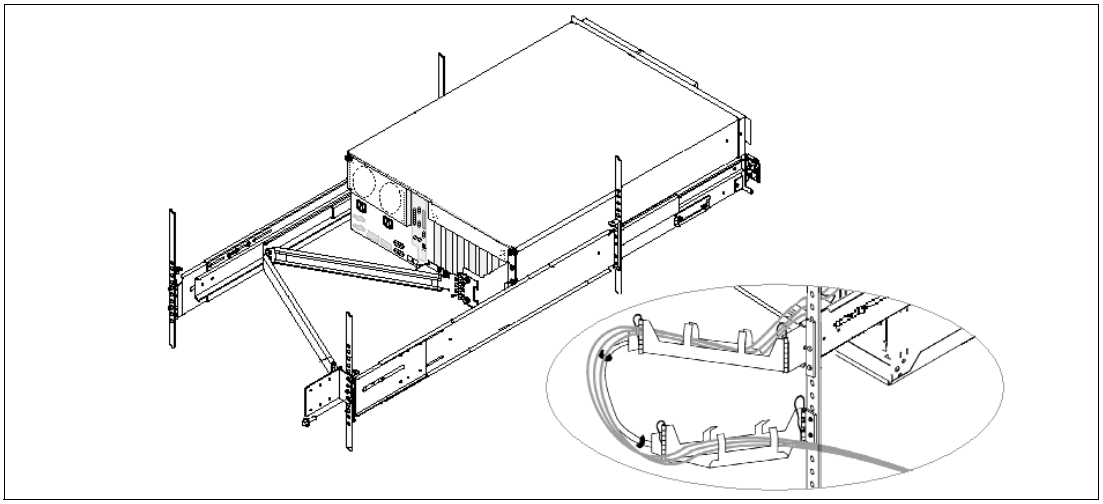


Figure 1-4 Cable management arm for Model 6C3

1.5.7 Flat-panel display options

For rack-mounted systems, the IBM 7316-TF2 Flat Panel Console Kit may be installed in the system rack. This 1U (EIA) console uses a 15-inch thin film transistor (TFT) LCD with a viewable area of 304.1 mm x 228.1 mm and a 1024 x 768 resolution.

Note: It is recommended that the 7316-TF2 be installed between EIA 20 to 25 of the rack for ease of use. The 7316-TF2 or any other graphics monitor requires a GXT135P graphics accelerator (FC 2849) to be installed in the server, or other graphic accelerator, if supported.

The 7316-TF2 Flat Panel Console Kit has the following attributes:

- Flat-panel color monitor.
- Rack tray for keyboard, monitor, and optional VGA switch with mounting brackets.

- IBM Space Saver 2 14.5-inch keyboard that mounts in the Rack Keyboard Tray and is available as a feature in 16 language configurations. The track point mouse is integral to the keyboard.

The L200P Flat-Panel Monitor (FC 3636) provides a desktop flat-panel display option. The L200P is a 20.1-inch TFT LCD digital screen with a maximum resolution of 1600 x 1200 pels[7] at 75 Hz in analog mode and 60 Hz in digital mode.

1.5.8 VGA switch and extender cables

The VGA switch for the IBM 7316-TF2 (FC 4200) allows the connection of up to eight servers to a single keyboard, display, and mouse.

To help minimize cable clutter, extenders for graphic display cable, the optional USB keyboard, and multi-connector cables (including graphic display cable, standard keyboard and mouse extension cables) of different lengths are available. See Table 1-6.

Table 1-6 VGA and cable feature codes

Feature code	Description	Length (meters)
4242	15P to 15P extender cable for Display	1.8
4243	CPU to VGA switch Graph+Keyboard+Mouse	2
4244	CPU to VGA switch Graph+Keyboard+Mouse	3.4
4245	CPU to VGA switch Graph+Keyboard+Mouse	6
4256	Extender cable for USB keyboard	2

The FC 4256 can be used to extend the connection between a USB keyboard and the 4-port USB PCI adapter (FC 2737).

The CPU to VGA switch cables (FC 4243, FC 4244, and FC 4245) can be used to connect the graphic accelerator (required in each attached system), keyboard port, and mouse port of a server to the switch, or to connect multiple switches in a tiered configuration. Using a two-level cascade arrangement, as many as 64 systems can be controlled from a single point.

This dual-user switch allows attachment of one or two consoles, one of which must be an IBM 7316-TF2. An easy-to-use graphical interface allows fast switching between systems and supports six languages (English, French, Spanish, German, Italian, or Brazilian Portuguese).

The VGA switch is only 1U high and can be mounted in the same tray as the IBM 7316-TF2, thus conserving valuable rack space. It supports a maximum video resolution of 1600 x 1280 pels, which facilitates the use of graphics-intensive applications and large monitors.

1.6 Cluster 1600 support

Cluster 1600 support of the p615 is available through CSM (Cluster Systems Management) and to Model 6C3 only (rack-mounted version). Up to 64 AIX 5L or Linux servers, or any combination of the two, are supported, while managed under a single focal point. For hardware control using CSM, an HMC is required. HACMP Version 5.1 supports the p615, however the native (integrated) serial ports cannot be used. An asynchronous adapter must be configured for p615 systems managed by HACMP.

[7] Picture elements (pels)

2

Architecture and technical overview

This chapter discusses the overall system architecture represented by Figure 2-1. The major components of this diagram will be described in the following sections. The bandwidths provided throughout this section are theoretical maximums provided for reference. It is always recommended to obtain real-world performance measurements using production workloads.

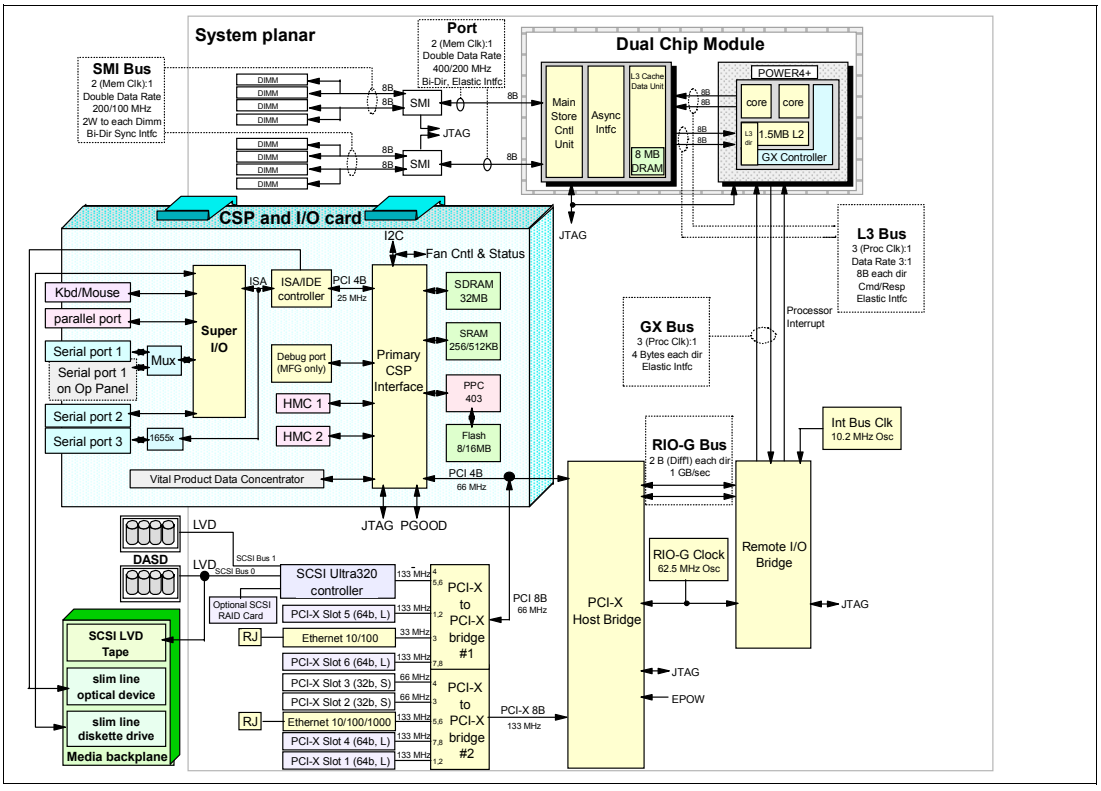


Figure 2-1 Conceptual diagram of the p615 POWER4+ system architecture

© Copyright IBM Corp. 2003. All rights reserved.

2.1 Processor and cache

The p615 implements the POWER4+ processor chip clocked at 1.2 GHz or 1.45 GHz. However, unlike the Multichip Module (MCM) packaging first used in the pSeries 690, or the Single Chip Module (SCM) used in pSeries 650 and 630, the p615 system uses a Dual Chip Module (DCM). In the DCM, the one- or two-core SCM is packaged with another chip, which combines an 8 MB L3 cache with a memory controller chip that is directly interfaced to the memory buffer SMI chips as shown in the enlarged view of the processor subsystem, Figure 2-3, and in Figure 2-1 on page 13. The DCM is permanently mounted on the system planar, containing either one or two processor cores (CPUs). Only a one- or two-processor configuration is available.

The chip-to-chip fabric bus (which was used between chips on the same MCM in the Model 690) is no longer relevant in the DCM. The DCM module does not have an external fabric bus. A logical diagram of the p615 DCM is shown in Figure 2-2.

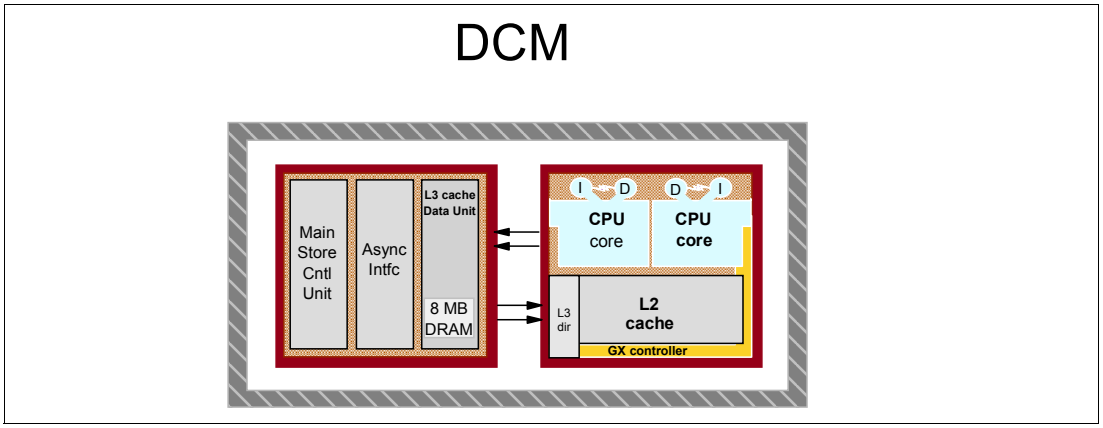


Figure 2-2 The DCM

Each DCM has a Ceramic Column Grid Array (CCGA) package where the chip carrier is raised slightly from its board mounting by small metal solder columns that provide the required connections and improved thermal resilience characteristics, and a Land Grid Array (LGA) socket for bringup.

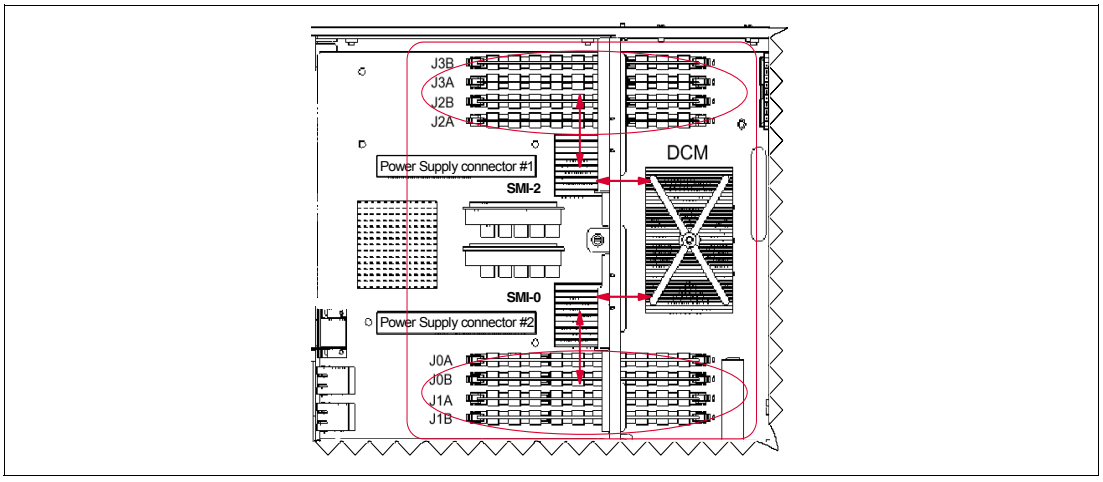


Figure 2-3 Enlarged view of the processor subsystem on system planar

2.1.1 L1, L2, and L3 cache

The POWER4+ storage subsystem consists of three levels of cache and the memory subsystem. The first two levels of cache are onboard the POWER4+ chip. The first level is 64 KB of Instruction (I) cache and 32 KB of Data (D) cache per processor core. The second level is 1.5 MB of L2 cache on the POWER4+. A two-way processor shares the L2 cache. All caches have either full ECC[1] or parity protection on the data arrays, and the L1 cache has the ability to re-fetch data from the L2 cache in the event of soft errors detected by parity checking.

The Level 3 cache consists of two components: The L3 cache controller/directory and the L3 data array. The L3 cache controller/directory is on the POWER4+ chip, and the L3 data array, which consists of 8 MB of embedded DRAM (eDRAM), is located in the other chip of the DCM. In 2-way systems, the L3 cache is shared.

2.1.2 PowerPC architecture

The p615 system complies with the RS/6000 platform architecture, which is an evolution of the PowerPC Common Hardware Reference Platform (CHRP) specifications. Figure 2-4 shows the PowerPC evolution.

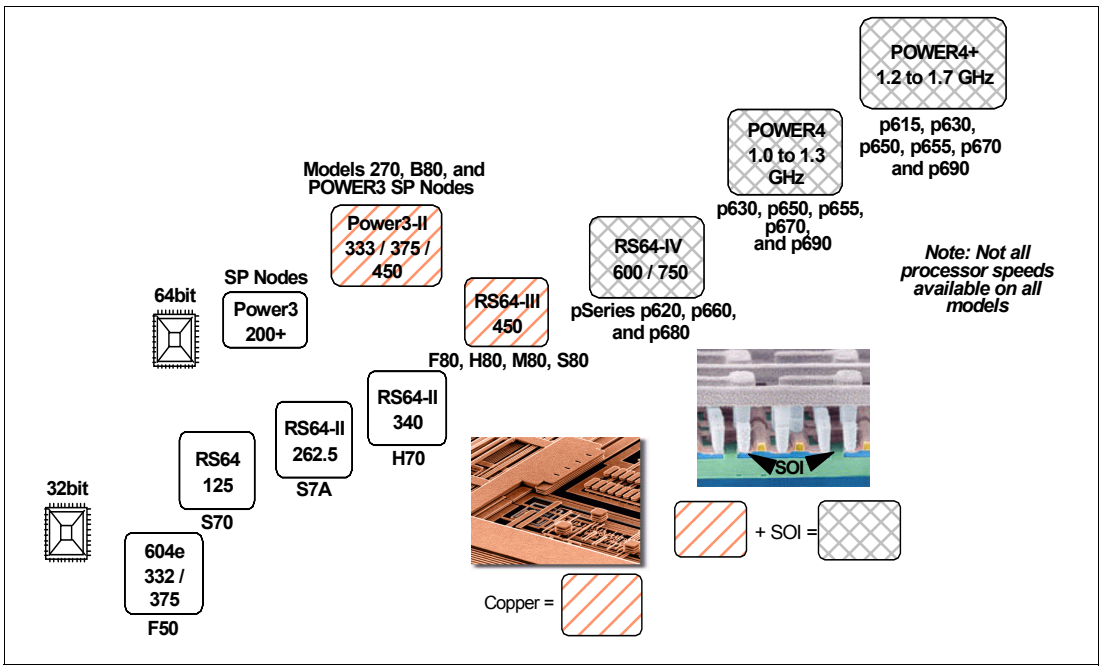


Figure 2-4 PowerPC evolution

2.1.3 Copper and CMOS technology

The POWER4+ processor chip takes advantage of IBMs leadership technology. It is made using IBM 0.13-µm-lithography CMOS[2] fabrication with seven levels of copper interconnect wiring. POWER4+ also uses Silicon-on-Insulator (SOI) technology to allow a higher operating frequency for improved performance, yet with reduced power consumption and improved reliability compared to processors not using this technology.

[1] ECC single error correct, double error detect
[2] Complementary Metal Oxide Semiconductor

2.1.4 Processor clock rate

The p615 operates with a processor clock rate of 1.2 or 1.45 GHz for POWER4+ systems.

To determine the processor characteristics on a running system, use one of the following commands:

`lsattr -El proc`*X* Where *X* is the number of the processor, for example, proc0 is the first processor in the system. The output from the command[3] would be similar to the following (False, as used in this output, signifies that the value cannot be changed through an AIX command interface):

```
state enable            Processor state  False
type powerPC_POWER4     Processor type   False
frequency 120000000     Processor Speed  False
```

`pmcycles -m` This command (AIX 5L Version 5.1 and later) uses the performance monitor cycle counter and the processor real-time clock to measure the actual processor clock speed in MHz. The following is the output of a two-way p615 1.2 GHz system:

```
Cpu 0 runs at 1200 MHz
Cpu 1 runs at 1200 MHz
```

Note: The **`pmcycles`** command is part of the bos.pmapi fileset. First check whether that component is installed using the **`lslpp -l bos.pmapi`** command.

2.2 Memory

The conceptual diagram of the memory subsystem of the p615 is shown in Figure 2-5. There are four 8-byte data paths from the memory controller to the memory with an aggregated bandwidth of 6.4 GBps to the processor card.

DDR memory can theoretically double memory throughput at a given clock speed by providing output on both the rising and falling edges of the clock signal (rather than only on the rising edge).

Memory access is through the on-chip L2 cache and Level 3 cache to the off-chip synchronous memory interface (SMI) and finally to the memory DIMMs, as represented in Figure 2-5.

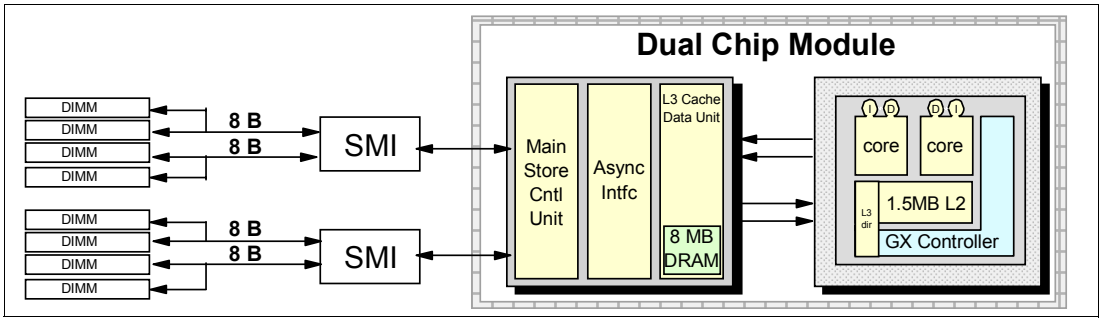


Figure 2-5 Conceptual diagram of POWER4+ processor and memory subsystem

[3] The output of the **`lsattr`** command has been expanded with AIX 5L to include the processor clock rate.

2.2.1 Memory options

The memory features available at the time of writing for the p615 are listed in Table 1-3 on page 4.

Each memory feature consists of four DIMMs, or quad, and must be installed according to Figure 2-6. First quad slots are J0A, J1A, J2A, and J3A, and for the second quad, the slots are J0B, J1B, J2B, and J3B.

Note: A quad must be comprised of a single feature. Mixed DIMM capacities in a quad will result in reduced RAS functionality.

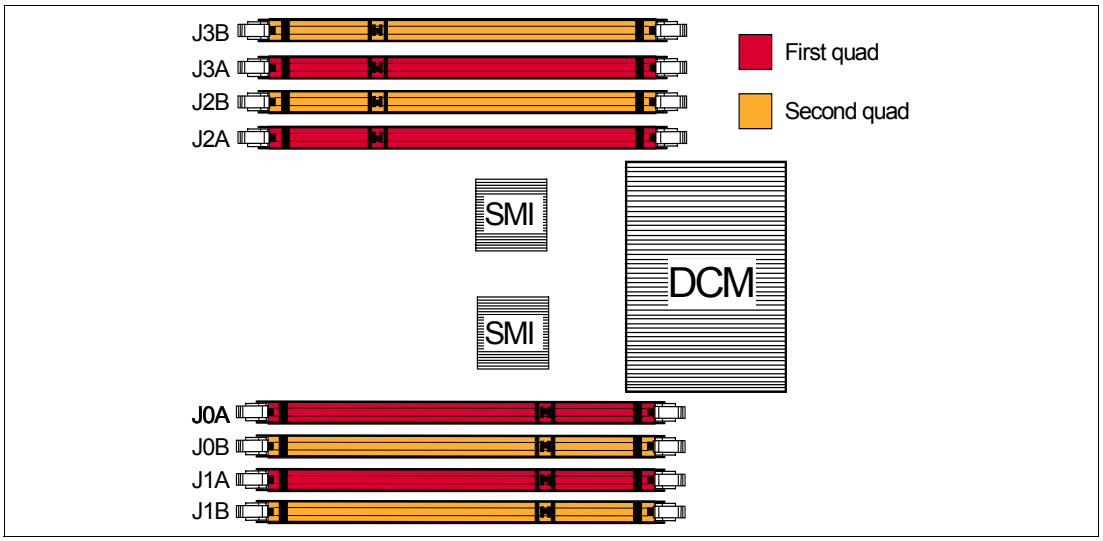


Figure 2-6 Memory placement for the p615

2.2.2 Memory restrictions (OEM)

The p615 does not officially support any OEM memory, and there is no exception to this rule. OEM memory is never certified for use in pSeries servers. If the p615 is populated with OEM memory, you could experience unexpected and unpredictable behavior.

All IBM memory is identified by an IBM logo and a white label printed with a barcode on top and an alphanumeric string on the bottom, created according to the rule reported in Figure 2-7:

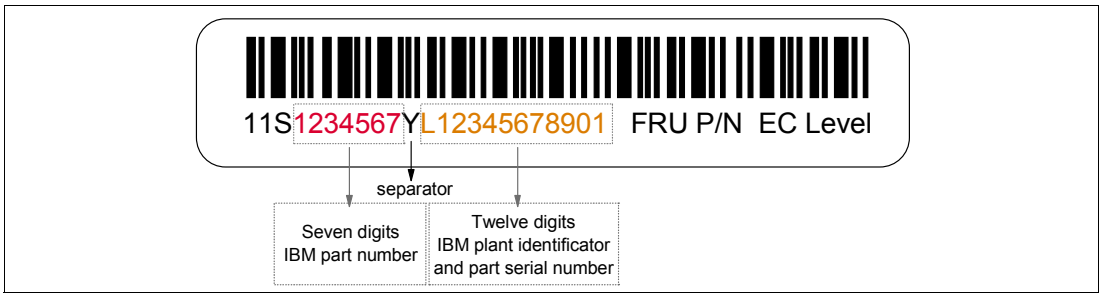


Figure 2-7 IBM barcode label

Sometimes OEM vendors put a label reporting the IBM memory part number but not the barcode and/or the alphanumeric string on their DIMMs.

In case of system failure caused by OEM memory installed in the system, the first thing to do is to replace the suspected memory with IBM memory, and check whether the problem is corrected. Contact your IBM representative for further assistance if needed.

2.3 System buses

The system bus from the processors to the memory subsystem is 2 x 16 bytes at 1/3 CPU clock speed for an aggregate data rate of 12.8 GBps. All system, I/O, and PCI-X buses support parity error detection. The memory controller is integrated on the DCM and directly interfaced with two 8-byte memory data ports (SMI) to control system DRAM.

2.3.1 GX bus

The GX controller (embedded in the POWER4+ chip) is responsible for controlling the flow of data through the GX bus. The GX bus is a high-frequency, single-ended, unidirectional, point-to-point bus. Both data and address information are multiplexed onto the bus, and for each path there is an identical bus for the return path. The GX bus has dual 4-byte paths at 400 MHz to give an aggregate data rate of 3.2 GBps. The processor card connects to the GX bus through its GX controller.

2.3.2 PCI-X host bridge and PCI-X bus

The I/O bridge contains a series of 1-byte buses grouped into pairs called ports. One of the buses in the pair is for inbound data transfers (500 MBps), while the other bus is for outbound data transfers (500 MBps). There is a total of two ports for the internal I/O. The two internal ports connect to a PCI-X Host Bridge (PHB). The PHB chip acts as a bridge between the I/O bus and two PCI-X to PCI-X bridges, which fan out to integrated I/O controllers and slots.

The PCI-X slots meet 64 bit/133 MHz requirements. However, two slots (5 and 6) are affected in performance by the EADS-X chip 1 that is connected to primary CSP interface. This lowers the speed of the EADS-X and its connected PCI-X Host Bridge port to 66 MHz. PCI-X and PCI adapters that demand high performance should be installed into slots 1 and 4 driven by the EADS-X chip 2 that is not connected to the CSP interface. Slots 2 and 3 are still connected to the EASD-X chip 2, but they are 32-bit slots. See Figure 2-1 on page 13.

2.4 PCI-X slots and adapters

PCI-X is the latest version of PCI bus technology, using a higher clock speed (133 MHz) to deliver a bandwidth of up to 1 GBps. The PCI-X slots in the p615 system support hot-plug and Extended Error Handling (EEH). EEH-capable adapters respond to a specially generated data packet from a PCI-X slot with a problem. This packet is analyzed by the system firmware, which then allows the device driver to reset the adapter or slot, isolating the error and reducing the need for a system reboot.

PCI-X slots also support existing 3.3 Volt PCI adapters. For a full list of the adapters that are supported on the p615 systems, and for important information regarding adapter placement, see the *RS/6000 and pSeries PCI Adapter Placement Reference*, SA38-0538, for additional information. You can find this publication at:

`http://www.ibm.com/servers/eserver/pseries/library/hardware_docs/pci_adp_pl.html`

2.4.1 64-bit and 32-bit adapters

As PCI-X slots support existing 3.3 Volt PCI adapters, that enables them to support both 64-bit and 32-bit adapters.

Choosing between 32-bit and 64-bit adapters influences slot placement and affects performance. Higher-speed adapters use 64-bit slots because they can transfer 64 bits of data for each data transfer phase.

Generally, 32-bit adapters can function in 64-bit PCI-X slots; however, 64-bit adapters cannot be used in the 32-bit slots on p615 systems.

2.4.2 LAN adapters

When a p615 is used as a server, it usually is connected to a local area network (LAN). The two internal Ethernet controllers integrated on the system planar, one at 10/100 Mbps and the second at 10/100/1000 Mbps, can be used to accomplish that.

Tip: In conjunction with certain network switches, you can use the Cisco Systems EtherChannel feature of AIX to build up one virtual Ethernet interface with increased bandwidth using two to eight Ethernet interfaces (adapters or integrated). See *Managing AIX Server Farms*, SG24-6606 for more information.

The list of the LAN adapters available at the time of writing is shown in the Table 2-1. IBM supports an installation with NIM using Ethernet and token-ring adapters. Use chrp as the platform type.

Table 2-1 LAN adapters supported by the p615

Feature Code	Adapter description	Size
4962	10/100 Mbps Ethernet	Short
5706	2-port 10/100/1000 Mbps Ethernet	Short
5707	2-port Gigabit Ethernet - SX	Short
4961	4-port 10/100 Mbps Ethernet	Long
5700	Gigabit Ethernet	Short
5701	Gigabit Ethernet-VTP	Short
4959	4/16 Mbps token ring	Short
4953	155 UTP ATM	Short
4957	155 MMF ATM	Short
2946	622 MMF ATM	Short

2.4.3 Graphic accelerators

The p615 only supports an enhanced POWER GXT135P (FC 2849) 2D graphic accelerator. The POWER GXT135P is a low-priced 2D graphics accelerator for pSeries servers. It can be configured to operate in either 8-bit or 24-bit color modes, and running at 60 Hz to 85 Hz. This adapter supports both analog and digital monitors. The adapter requires one PCI slot.

2.4.4 Audio adapter

The Model 6E3 supports a new audio PCI adapter (FC 8244). It is a high-performance workstation-class audio PCI adapter that is optimized for support in the AIX professional workstation environment and their associated applications. It is 3.3 Volt, 32-bit PCI adapter that runs at 33 MHz and complies with the Intel AC 1997 Version 2.1 Specifications. The adapter provides external jacks for headphones, speaker output, line input, and microphone input as well as an internal connector for CD or DVD drive audio input.

2.4.5 Internal Ultra320 SCSI controllers

The integrated Ultra320 SCSI controllers provide two internal SCSI buses to connect the two front-accessible 4-pack DASD backplanes. Each 4-pack DASD backplane can accommodate up to four hot-swappable disks. The Ultra320 SCSI controller supports single-ended or low-voltage differential (LVD) devices at up to 320 MBps[4]. For a detailed description of disk placement or disk capacity planning see section 2.5.3, “Preferred boot device and options” on page 23.

A SCSI Enclosure Services (SES) device supports the disk bay hot-swappable features. The SES processor is the SCSI hot-swap manager; it provides the control mechanism for the device hot-swap options such as identify/replace/remove.

2.5 Internal storage

The p615 supports four internal media bays. Bay 1 always contains the operator-panel. Bays 2 and 3 are designed to receive slimline devices such as a diskette drive (FC 2611) or DVD-ROM drive (FC 2640). Bay 4 can be populated by the devices listed in Table 2-2.

Table 2-2 List of internal media options

Feature Code	Description	Comments
2623	DVD-RAM drive 4.7 GB capacity	The device will be installed in Bay 4, and directly connected to the media backplane.
6158	4 mm 20/40 GB tape drive	The device will be installed in Bay 4, and directly connected to the media backplane.
6134	8 mm 60/150 GB tape drive	The device will be installed in Bay 4, and directly connected to the media backplane.
6120	80/160 GB VXA tape drive	The device will be installed in Bay 4, and directly connected to the media backplane.

Cabling internal devices

For the p615, there are several options available to configure and connect internal devices. Table 2-3 provides the different possibilities for internal cabling, depending on flexibility or scalability.

[4] SCSI T10 Technical Committee, http://www.t10.org

Table 2-3 Feature code combinations using FC 6574 backplane

Description	Feature codes required				
	6574 backplane		SCSI Card	Cable	
	base	option	5712	4263	4266
One default 4-pack DASD backplane. Up to 4 internal disks.	1				
Two 4-pack DASD backplanes. Up to 8 internal disks.	1	1			
One default 4-pack DASD backplane. Up to 4 internal disks and one internal SCSI LVD device.	1			1	
One default 4-pack DASD backplane. Up to 4 internal disks and one internal SCSI SE device.	1		1		1
Two 4-pack DASD backplanes. Up to 8 internal disks and one internal SCSI LVD device.	1	1		1	
Two 4-pack DASD backplanes. Up to 8 internal disks and one internal SCSI SE device.	1	1	1		1
One default 4-pack DASD backplane. Up to 4 internal disks and external SCSI SE connections.	1		1		
Two 4-pack DASD backplanes. Up to 8 internal disks and external SCSI SE connections.	1	1	1		
Two 4-pack DASD backplanes. Up to 8 internal disks, one internal SCSI LVD device, and External SCSI SE connections.	1	1	1	1	
Two 4-pack DASD backplanes. Up to 8 internal disks, one internal SCSI SE device, and External SCSI SE connections.	1	1	1		1

Figure 2-8 is intended to assist you in matching the feature code combinations, described in Table 2-3, related to the internal SCSI cabling for SCSI devices. No IDE device connections are shown in the figure.

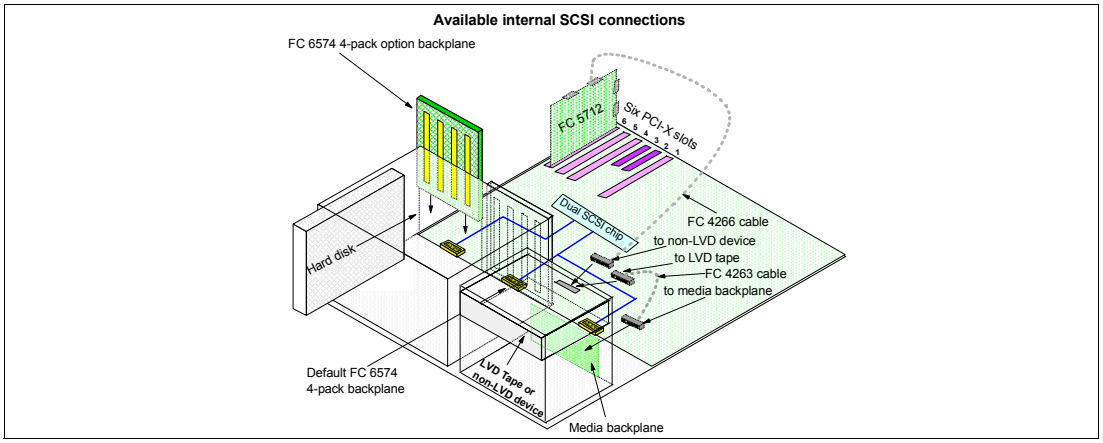


Figure 2-8 Internal SCSI connections diagram

Figure 2-8 contains all the possible internal SCSI connections and cabling for the FC 6574 backplane. In this example, the SCSI adapter FC 5712 is plugged into the PCI-X slot 6.

Ultra320 backplane

An optional backplane FC 6584 allows an Ultra320 SCSI adapter to be cabled (using FC 4267) to provide up to four disks to operate independently to the integrated controller to provide a second path to internal disks for high availability RAID or non-RAID solutions.

For the placement of the SCSI adapter and all additional adapters, as well as limitations and restrictions, see the *RS/6000 and pSeries PCI Adapter Placement Reference*, SA38-0538, as suggested in section 2.4, “PCI-X slots and adapters” on page 18.

2.5.1 Hot-swappable SCSI disks

The p615 can have up to eight hot-swappable drives plugged in the two 4-pack DASD backplanes. The hot-swap process is controlled by the SCSI enclosure service (SES), which is located in the 4-pack DASD backplane (AIX assigns the name `ses0` to the first 4-pack, and `ses1` to the second, if present). The two hot-swappable 4-pack DASD backplanes can accommodate the devices listed in Table 2-4.

Table 2-4 Hot-swappable disk options

Feature Code	Description
3273	36.4 GB 10,000 RPM Ultra3 SCSI hot-swappable disk drive
3277	36.4 GB 15,000 RPM Ultra3 SCSI hot-swappable disk drive
3274	73.4 GB 10,000 RPM Ultra3 SCSI hot-swappable disk drive
3278	73.4 GB 15,000 RPM Ultra3 SCSI hot-swappable disk drive
3275	146.8 GB 10,000 RPM Ultra3 SCSI hot-swappable disk drive

At the time of writing, if a new order is placed with two 4-pack DASD backplanes (FC 6574) and more than one disk, the system configuration shipped from manufacturing will balance the total number of SCSI disks between the two 4-pack SCSI backplanes. This is for manufacturing test purposes, and not because of any limitation. Having the disks balanced between the two 4-pack DASD backplanes allows the manufacturing process to systematically test the SCSI paths and devices related to them.

Prior to the hot-swap of a disk in the hot-swappable bay, all necessary operating system actions must be undertaken to ensure that the disk is capable of being deconfigured.

Note: It is recommended to follow this procedure when removing a hot-swappable disk:

1. Release the tray handle on the disk assembly.
2. Pull out the disk assembly a little bit from the original position.
3. Wait up to 20 seconds until the internal disk stops spinning.
4. Now you can safely remove the disk assembly from the 4-pack DASD backplane.

Once the disk drive has been deconfigured, the SCSI enclosure device will power-off the slot, enabling safe removal of the disk. You should ensure that the appropriate planning has been given to any operating-system-related disk layout, such as the AIX Logical Volume Manager, when using disk hot-swap capabilities. For more information, see *Problem Solving and Troubleshooting in AIX 5L*, SG24-5496.

Note: After the SCSI disk hot-swap procedure, you can expect to find `SCSI_ERR10` logged in the AIX error log, with the second word of the sense data equal to `0017`. It is generated from a SCSI bus reset issued by the SES to reset all processes when a drive is inserted.

Hot-swap disks and Linux

Linux does not support hot-swap of any disk drive, therefore the Linux operating system does not support these hot-swappable procedures. A p615 system running Linux must be shut down and powered off before you replace any disk drives.

2.5.2 Other hot-plug options

The p615 provides the ability to concurrently change or add PCI adapters.

PCI-X and PCI adapters

The addition, removal, or changing of a PCI-X or PCI adapter can be accomplished by using a system management tool such as Manage PCI Hot-Plug Slots (using the Web-based System Manager) or PCI Hot-Plug Manager (using SMIT). The PCI hot-plug tasks can be accomplished by using the Hot-Plug Task of the online diagnostics task selection menu (using the **`diag`** command). Each of these tools provides a method by which a PCI-X slot can be identified first, powered off to enable removal or insertion of an adapter, and then powered back on to enable the configuration of the device.

2.5.3 Preferred boot device and options

The p615 handles the boot process in a way that is similar to other pSeries servers.

The initial stage of the boot process is to establish that the machine has powered up correctly and the memory and CPUs are functioning correctly. This sequence of events is performed by the service processor of the p615. Once the machine reaches the System Management Services (SMS) menus, all of the necessary tests have been performed and the machine is scanning the bus for a boot source.

Most pSeries system backplanes, including the p615, are designed such that the drive in the first slot spins up immediately after power-on, and other drives will wait for the operating system to send a command before spinning up. The left-most slot of the 4-pack DASD backplane (SCSI ID 8, boot, autostart) is set to spin up immediately after power-on. The power-on delay sequence is performed to prevent power supply overloading. This behavior makes the disk in the first slot of the first 4-pack DASD backplane (disk bay 4) the preferred boot device. See Figure 2-9 to locate all of the disk bays.

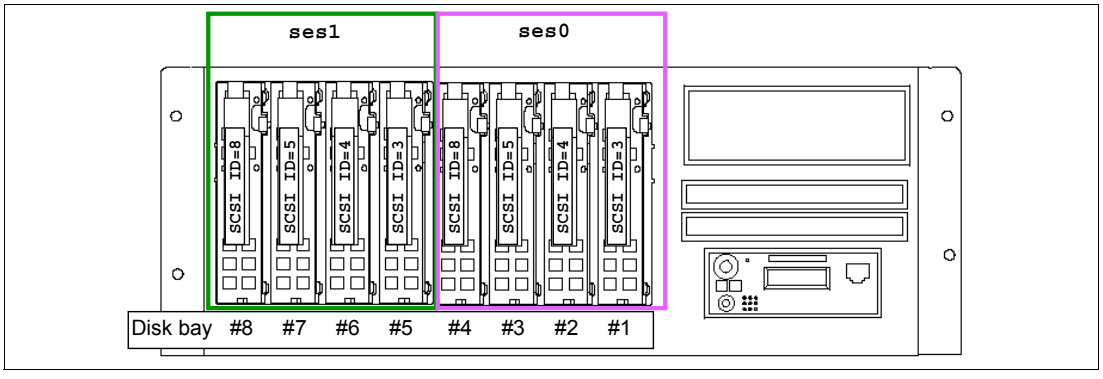


Figure 2-9 Disk bays and SCSI addresses

Instead of booting from the preferred boot device, or from any other internal disks, there are a number of other possibilities:

DVD-ROM, DVD-RAM
These devices can be used to boot the system so that a system can be loaded, system maintenance performed, or stand-alone diagnostics performed.

Internal or external tape drives
The media bay tape drive or any externally attached tape drive can be used to boot the system using **mksysb**, for example.

SCSI disk
The more common method of booting the system is to use a disk situated in one of the hot-swap bays in the front of the machine. However, any external non-RAID SCSI-attached disk could be used if required.

SSA disk
The p615 supports booting from an SSA disk either as an AIX system disk or as a RAID LUN. FC 6230 serial RAID adapters can be used to provide the boot support from a RAID-configured disk, provided that the firmware level of the adapter is BB00 or later. For more information about the SSA boot, see the SSA Frequently Asked Questions located on the Web[5].

Note: Fastwrite must not be enabled on the boot resource SSA adapter.

SAN boot
It is possible to boot these systems from a SAN using a 2 GB Fibre Channel Adapter (FC 6239). The IBM 2105 Enterprise Storage Server (ESS) is an example of a SAN-attached device that can provide a boot medium.

LAN boot
Network boot and NIM installs can be used if required. An example of this would be in the ISP environment, where a high density of machines are installed in racks and it is impractical to use CD media.

2.6 System Management Services (SMS)

When the p615 is equipped with either a graphic adapter connected to a graphics display, keyboard, and mouse device, or an ASCII display terminal connected to one of the first two system serial ports, you can use the System Management Services menus to view information about the system and perform tasks such as setting a password, changing the boot list, and setting the network parameters.

To start the System Management Services, press the 1 key on the terminal, or in the graphic window press the 1 key on the keyboard after the word `keyboard` appears and before the word `speaker` appears. After the text-based System Management Services start, the screen shown in Figure 2-10 displays.

[5] http://www.storage.ibm.com/hardsoft/products/ssa/faq.html#microcode

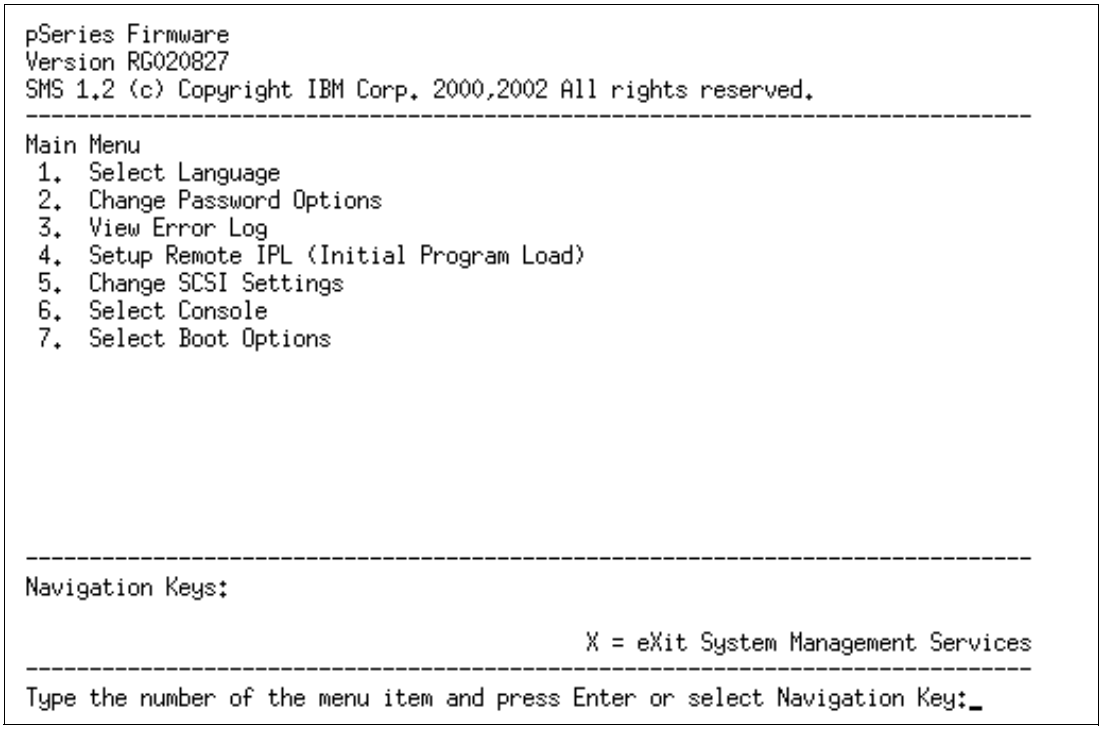


Figure 2-10 System Management Services main menu

Note: The version of firmware currently installed in your system is displayed at the top of each screen. Processor and other device upgrades may require a specific version of firmware to be installed in your system.

On each menu screen, you are given the option of choosing a menu item and pressing Enter (if applicable), or selecting a navigation key. You can use the different options to set the password and protect our system setup, review the error log for any kind of error reported during the first phases of the booting steps, or set up the network environment parameters if you want the system boots from a NIM server. For more information see *IBM eServer pSeries 615 Model 6C3 and Model 6E3 Service Guide,* SA38-0630.

Use the Select Boot Options menu to view and set various options regarding the installation devices and boot devices:

1. **Select Install or Boot a Device**
2. **Select Boot Devices**
3. **Multiboot Startup**

Option 1 (**Select Install or Boot a Device**)	Enables you to select a device to boot from or install the operating system from. This selection is for the current boot only.
Option 2 (**Select Boot Devices**)	Enables you to set the boot list.
Option 3 (**Multiboot Startup**)	Toggles the multiboot startup flag, which controls whether the multiboot menu is invoked automatically on startup.

2.7 Security

The p615 allows you to set two different types of passwords to limit the access to these systems. The *privileged access password* can be set from service processor menus or from System Management Services (SMS) utilities. It provides the user with the capability to access all service processor functions. This password is usually used by the system administrator or root user. The *general access password* can be set only from service processor menus. It provides limited access to service processor menus, and is usually available to all users who are allowed to power on the server, especially remotely.

2.8 Operating system requirements

A p615 is capable of running IBM AIX 5L for POWER and support appropriate versions of Linux. AIX 5L has been specifically developed and enhanced to exploit and support the extensive RAS features on IBM @server pSeries systems.

2.8.1 AIX 5L

The p615 requires AIX 5L Version 5.1 at Maintenance Level 4 (APAR IY39794) or AIX 5L Version 5.2 at Maintenance Level 1 (APAR IY39795). In order to boot from the CD, make sure you have one of the following media:

- AIX 5L Version 5.1 5765-E61, dated 10/2002 (CD# LCD4-1061-04) or later
- AIX 5L Version 5.2 5765-E62, initial CD-set (CD# LCD4-1133-00) or later

IBM periodically releases maintenance packages for the AIX 5L operating system. These packages are available on CD-ROM (FC 0907), or they can be downloaded from the Internet at:

`http://techsupport.services.ibm.com/server/fixes`

You can also get individual operating system fixes and information about obtaining AIX 5L service at this site. If you have problems downloading the latest maintenance level, ask your IBM Business Partner or IBM representative for assistance.

To check your current AIX level enter the **`oslevel -r`** command. The output for AIX 5L Version 5.2 Maintenance Level 1 is `5200-01`.

AIX 5L application binary compatibility

IBM AIX 5L Version 5.2 preserves binary compatibility for 32-bit application binaries from previous levels of AIX Version 4 and AIX 5L, and for 64-bit applications compiled on previous levels of AIX 5L; 64-bit applications compiled on Version 4 must be recompiled to run on AIX 5L.

2.8.2 Linux

For the p615, a Linux distribution was available through SuSE at the time this publication was written. In Japan, Turbolinux is also available. In the Latin America sales region, Conectiva is also available. For related information and overview, see:

`http://www.ibm.com/servers/eserver/pseries/linux`

Find full information about SuSE Linux Enterprise Server 8 for pSeries at:

`http://www.suse.com/us/business/products/server/sles/i_pseries.html`

For information about UnitedLinux for pSeries from Turbolinux, see:

http://www.turbolinux.co.jp

For the latest in IBM Linux news, subscribe to the Linux Line. See:

https://www6.software.ibm.com/reg/linux/linuxline-i

Many of the features described in this document are operating system dependant and may not be available on Linux. For more information, please check:

http://www.ibm.com/servers/eserver/pseries/linux/whitepapers/linux_pseries.html

Linux support

IBM only supports the Linux systems of customers with a SupportLine contract covering Linux. Otherwise, the Linux distributor should be contacted for support.

3

Availability, investment protection, expansion, and accessibility

The following sections provide more detailed information about configurations, upgrades, and design features that will help lower the total cost of ownership. This section includes several features based on the benefits available when using AIX 5L. Support of these features using Linux may vary.

© Copyright IBM Corp. 2003. All rights reserved.

3.1 Autonomic computing

The IBM autonomic computing initiative is about using technology to manage technology. This initiative is an ongoing effort to create servers that respond to unexpected capacity demands and system glitches without human intervention. The goal: New highs in reliability, availability, and serviceability, and new lows in downtime and cost of ownership.

Today's pSeries offers some of the most advanced self-management features for UNIX servers on the market today.

Autonomic computing on IBM @server pSeries servers describes the many self-configuring, self-healing, self-optimizing, and self-protecting features that are available on IBM @server pSeries servers.

Self-configuring

Autonomic computing provides self-configuration capabilities for the IT infrastructure. Today, IBM systems are designed to provide this at a feature level with capabilities such as plug-and-play devices and configuration setup wizards. Examples include:

- Virtual IP address (VIPA)
- IP multipath routing
- Microcode discovery services/inventory scout
- Hot-swappable disks
- Hot-plug PCI
- Wireless/pervasive system configuration
- TCP explicit congestion notification

Self-healing

For a system to be self-healing, it must be able to recover from a failing component by first detecting and isolating the failed component, taking it off-line, fixing or isolating it, and reintroducing the fixed or replacement component into service without any application disruption. Examples include:

- Multiple default gateways
- Automatic system hang recovery
- Automatic dump analysis and e-mail forwarding
- EtherChannel automatic failover
- Graceful processor failure detection and failover
- First failure data capture
- Chipkill™ ECC Memory, dynamic bit-steering
- Memory scrubbing
- Automatic, dynamic deallocation (processors, PCI buses/slots)
- Electronic Service Agent - *Call Home* support

Self-optimization

Self-optimization requires a system to efficiently maximize resource utilization to meet the end-user needs with no human intervention required. Examples include:

- Workload manager enhancement
- Extended memory allocator
- Reliable, scalable cluster technology (RSCT)
- PSSP cluster management and Cluster Systems Management (CSM)

Self-protecting

Self-protecting systems provide the ability to define and manage the access from users to all of the resources within the enterprise, protect against unauthorized resource access, detect intrusions and report these activities as they occur, and provide backup/recovery capabilities that are as secure as the original resource management systems. Examples include:

- Kerberos V5 authentication (authenticates requests for service in a network)
- Self-protecting kernel
- LDAP directory integration (LDAP aids in the location of network resources)
- SSL (manages Internet transmission security)
- DigitalCertificates
- Encryption (prevents unauthorized use of data)

3.2 HMC management

The HMC uses a dedicated connection to the p615 to provide some basic functions needed to control the server. The p615 does not support LPAR or DLPAR, but the same HMC used for p690, p670, p650, and p630 can be attached. For the p615, the HMC sees the same general interface as the other systems, with the exception that it is not possible to create logical partitions under the p615 system. Beside seeing the system status and the operator panel content (firmware and AIX progress codes and error codes) in the navigation tree, you also see tasks to power the system off and on (with multiple boot options), reset the operating system, open a virtual AIX/Linux system console window, and view system information (properties). The HMC can be used to install, start, and manage the server without a dedicated display device connected to the server.

The HMC supports multiple systems, so if you have multiple p615s, it provides consolidated management that is accessible remotely. In addition, it provides the Service Focal Point functionality (collection of hardware error events), *Call Home* for problems, and other service functions.

Information regarding HMC control of clustered servers under the control of IBM Parallel Systems Support Programs for AIX (PSSP) or Cluster Systems Management for AIX (CSM) can be found in the Scaling Statement section of the *Family 9078+01 IBM @server Cluster 1600, 9078-160* sales manual, accessible on IBMlink:

http://www.ibmlink.ibm.com

3.3 Reliability, availability, and serviceability (RAS) features

Excellent quality and reliability are inherent in all aspects of the p615 design and manufacture, and the fundamental principle of the design approach is to minimize outages. The RAS features help to ensure that the system operates when required, performs reliably, and efficiently handles any failures that may occur. This is achieved using capabilities provided by both the hardware and the AIX 5L operating system.

Mainframe-class diagnostic capability based on internal error checkers, First-Failure Data Capture (FFDC), and run-time analysis are provided. This monitoring of all internal error check states is provided for processor, memory, I/O, power, and cooling components, and is aimed at eliminating the need to try to recreate failures later for diagnostic purposes. The unique IBM RAS capabilities are important for the availability of your server.

The following features provide the p615 with UNIX industry-leading RAS features:

- Flexible internal RAID configurations using optional features.
- Fault avoidance through highly reliable component selection, component minimization, and error handling technology designed into the chips
- Improved reliability through processor operation at a lower voltage, enabled by the use of copper chip circuitry and Silicon-on-Insulator technology
- Fault tolerance through additional hot-swappable power supply, and the capability to perform concurrent maintenance for power and cooling
- Automatic First Failure Data Capture (FFDC) and diagnostic fault isolation capabilities
- Concurrent run-time diagnostics based on First Failure Data Capture
- Predictive failure analysis on processors, cache, memory, and disk drives
- Dynamic error recovery
- Error Checking and Correction (ECC) or equivalent protection (such as re-fetch) on main storage, all cache levels (1, 2, and 3), and internal processor arrays
- Dynamic processor deallocation based on run-time errors (requiring more than one processor)
- Persistent processor deallocation (boot-time deallocation based on run-time errors)
- Persistent deallocation extended to memory
- Chipkill correction in memory
- Memory scrubbing and redundant bit-steering for self-healing
- Industry-leading PCI-X bus parity error recovery as first introduced on the p690 systems
- Hot-plug functionality of the PCI-X bus I/O subsystem
- PCI-X bus and slot deallocation
- Disk drive fault tracking that monitors the number/rate of data errors and thresholds several recoverable hardware errors
- Avoiding checkstops with process error containment
- Environmental monitoring (temperature and power supply)
- Auto-reboot

Some of the RAS features of the p615 are covered in more detail in the following sections.

3.3.1 Service processor

The converged service processor is a specialized device that is packaged with the native I/O ports (serial ports, parallel port, keyboard/mouse ports, and HMC ports) into a dedicated card assembly that plugs into the system planar. It provides a number of different functions, discussed below:

- With the machine powered off, the service processor is in an idle state waiting for either a power-on command or a keystroke from any of the TTYs attached to either of the S1 or S2 ports. At this point, `0K` is displayed on the control panel LED.
- Immediately after power-on, the System Power Control Network (SPCN), a function of the converged service processor, controls the powering up of all devices needed during the boot process. When the SPCN has completed its tasks, the CSP, using an onboard processor, checks for CPU and memory resources and then tests them. After the CPU and memory tests have completed, the service processor then hands the rest of the boot

process over to system firmware. This changeover occurs when the 9*xxx* LED codes become E*xxx* codes.

- With AIX in control of the machine, the service processor is still working and checking the system for errors. Also, the surveillance function of the service processor is monitoring AIX to check that it is still running and has not stalled.
- With the machine powered off but power still attached (standby), any TTY keyboard attached to either the S1 or S2 native serial port having the Enter key depressed will cause the service processor to display the service processor main menu. This menu and subsequent menus will be discussed in the following sections.
- Converged service processor is connected to the HMC ports through an internal interface and reacts to any power on/off command issued on the HMC console.

Service processor main menu

The service processor main menu and subsequent menus are only visible on an ASCII screen attached to the native serial ports. Example 3-1 shows the service processor main menu.

Example 3-1 Service processor main menu

```
Service Processor Firmware
       Version: RF030504
 Copyright 2001, IBM Corporation
            100FB5A

_____________________________________
            MAIN MENU

 1. Service Processor Setup Menu
 2. System Power Control Menu
 3. System Information Menu
 4. Language Selection Menu
 5. Call-In/Call-Out Setup Menu
 6. Set System Name
99. Exit from Menus
```

Useful information contained within this menu is defined as follows:

RF030504 Service processor firmware level and date code. The RF030504 firmware indicated in this example is a preproduction level. Generally available versions may reflect a later date code.

100FB5A Machine serial number.

Service Processor Setup menu These menus enable you to set passwords that provide added security to your system and update machine firmware.

System Power Control menu These menus enable you to control some aspects of how the machine powers on and how much of the boot process you wish to complete.

System Information menu From this menu option, information about the previous boot sequence and errors produced can be viewed. Additionally, this menu provides a means of checking and altering the availability of CPUs and memory.

3.3.2 Memory reliability, fault tolerance, and integrity

The p615 uses Error Checking and Correcting (ECC) circuitry for system memory to correct single-bit and to detect double-bit memory failures. Detection of double-bit memory failures helps maintain data integrity. Furthermore, the memory chips are organized such that the failure of any specific memory module only affects a single bit within a four-bit ECC word (*bit-scattering*), thus allowing for error correction and continued operation in the presence of a complete chip failure (*Chipkill recovery*). The memory DIMMs also utilize *memory scrubbing* and thresholding to determine when spare memory modules within each bank of memory should be used to replace ones that have exceeded their threshold of error count (*dynamic bit-steering*). Memory scrubbing is the process of reading the contents of the memory during idle time and checking and correcting any single-bit errors that have accumulated by passing the data through the ECC logic. This function is a hardware function on the memory controller chip and does not influence normal system memory performance.

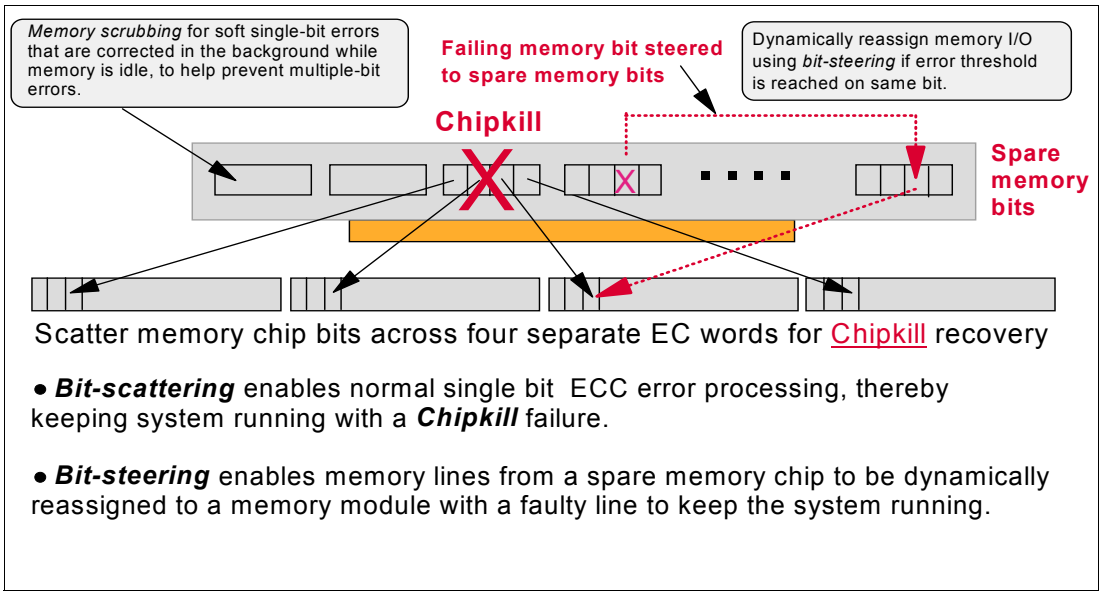


Figure 3-1 Main storage ECC and extensions

3.3.3 First failure data capture, diagnostics, and recovery

If a problem should occur, the ability to correctly diagnose it is a fundamental requirement upon which improved availability is based. The p615 incorporates unmatched capability in start-up diagnostics (see 3.3.1, “Service processor” on page 32) and in run-time First Failure Data Capture based on strategic error checkers built into the chips.

Any errors detected by the pervasive error checkers are captured into Fault Isolation Registers (FIRs), which can be interrogated by the service processor. The service processor in the p615 has the capability to access system components using special purpose service processor ports or by access to the error registers (Figure 3-2).

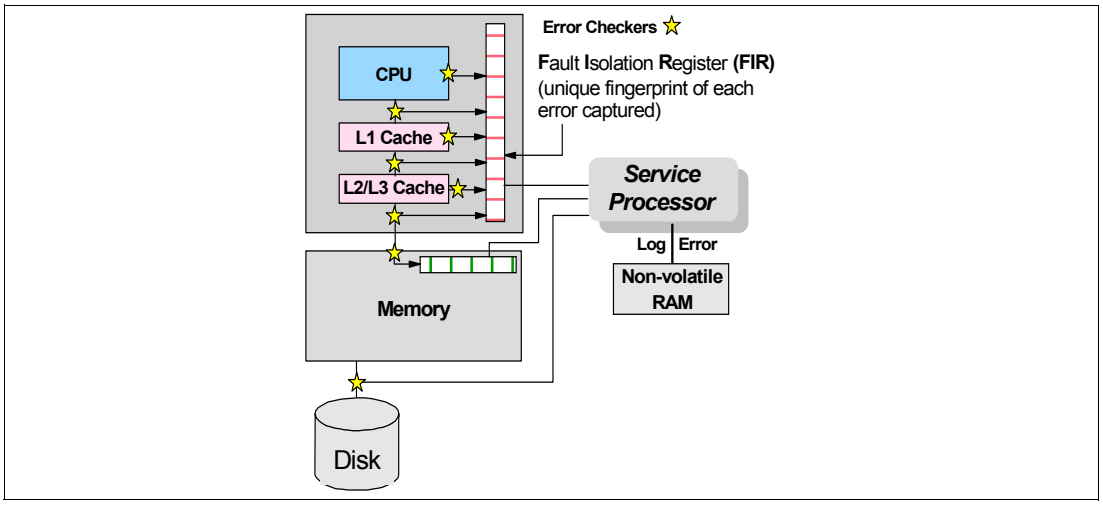


Figure 3-2 Schematic of Fault Isolation Register implementation

The FIRs are important because they enable an error to be uniquely identified, thus enabling the appropriate action to be taken. Appropriate actions might include such things as a bus retry, ECC correction, or system firmware recovery routines. Recovery routines could include dynamic deallocation of potentially failing components such as a processor (at least two processors required) or L2 cache.

Errors are logged into the system non-volatile random access memory (NVRAM) and the service processor event history log, along with a notification of the event to AIX for capture in the operating system error log. Diagnostic Error Log Analysis (**`diagela`**) routines analyze the error log entries and invoke a suitable action such as issuing a warning message. If the error can be recovered, or after suitable maintenance, the service processor resets the FIRs so that they can accurately record any future errors.

The ability to correctly diagnose any pending or firm errors is a key requirement before any dynamic or persistent component deallocation or any other reconfiguration can take place.

3.3.4 Error indication and LEDs indicators

The p615 is designed to positively identify a Field Replaceable Unit (FRU) that was identified by firmware as a cause of a system malfunction. Attenuation of the error is provided through a series of light attention signals, starting on the exterior of the system (System Attention LED) located on the front of the system (also on the rear of the Model 6C3), and ending with an LED near the failing FRU.

Attention LEDs indicator

The attention LEDs indicator provides a means to identify failing components in the system. Attention LEDs are located on the front and rear of the system to make it easy for an on-site technician to quickly locate the system requiring attention.

The attention indicator is represented externally by an amber LED on the operator panel and the back of the system unit. It is used to indicate that the system is in one of the following states:

- Normal state, LED is off.
- Fault state, LED is on solid.
- Identify state, LED is blinking.

Normal state

This state indicates that the system is working normally. The following conditions are defined as normal states:

- System is running with no loss of identifiable resources.
- System is running with no loss of N+1 redundancy.
- System is running with recoverable errors under threshold (for example, ECC).
- Faulty resource either was physically removed or its presence cannot be detected by system hardware and firmware.
- An EPOW warning (not loss of N+1 redundancy event) condition is present.
- The system is shut down and powered off normally.

Fault state

This state indicates that the system hardware, firmware, or diagnostics detected a fault in the system. The system may continue to run in a degraded mode, fail and reboot successfully, or fail and halt. The following system states are defined to cause a fault state:

- Fault detected during system boot/reboot, system either hung or halted.
- Fault detected during system boot/reboot, a resource may be deconfigured, or an error has been detected by the CSP or SPCN network. System is booted successfully and running in degraded mode.
- A predictive hardware error (recoverable error over threshold) condition detected by system firmware or diagnostics during system runtime. Resource with fault over threshold condition may or may not be deconfigured.
- A hardware error, which requires FRU replacement, detected by periodic diagnostic (for example, processor floating point diagnostic).
- A fault that causes the loss of power supply or fan N+1 redundancy.
- Loss of system power due to internal power supply fault.
- A fault that results in system checkstop, machine check, or surveillance timeout.

When a failing component is detected in your server, or the firmware determines a fault state, the attention LED is turned on.

3.3.5 Dynamic or persistent deallocation

Dynamic deallocation of potentially failing components is non-disruptive, allowing the system to continue to run. Persistent deallocation occurs when a failed component is detected and is then deactivated at the subsequent boot time.

Dynamic deallocation functions include:

- Processor
- L3 cache line delete
- PCI bus and slots

For dynamic processor deallocation, the service processor performs a predictive failure analysis based on any recoverable processor errors that have been recorded. If these transient errors exceed a defined threshold, the event is logged and the processor is deallocated from the system while the operating system continues to run. This feature (named *CPU Guard*) enables maintenance to be deferred until a suitable time. Processor deallocation can only occur if there are sufficient functional processors (at least, more than one).

To verify whether CPU Guard has been enabled, run the following command:

```
lsattr -El sys0 | grep cpuguard
```

If enabled, the output will be similar to the following:

```
cpuguard    enable      CPU Guard     True
```

If the output shows CPU Guard as disabled, enter the following command to enable it:

```
chdev -l sys0 -a cpuguard='enable'
```

Note: The use of cpuguard is effective on systems with two or more functional processors running AIX 5L Version 5.2. AIX 5L Version 5.1 requires three or more functional processors in order to have the feature running properly.

Cache or cache-line deallocation is aimed at performing dynamic reconfiguration to bypass potentially failing components. This capability is provided for both L2 and L3 caches. The L1 data cache and L2 data and directory caches can provide dynamic detection and correction of hard- or soft-array cell failures. Dynamic run-time deconfiguration is provided if a threshold of L1 or L2 recovered errors is exceeded.

In the case of an L3 cache run-time array single-bit solid error, the spare chip resources are used to perform a line delete on the failing line.

System bus recovery (retry) is provided for any address or data parity errors on the GX bus or for any address parity errors on the fabric bus.

PCI hot-plug slot fault tracking helps prevent slot errors from causing a system machine check interrupt and subsequent reboot. This provides superior fault isolation and the error affects only the single adapter. Run-time errors on the PCI bus caused by failing adapters will result in recovery action. If this is unsuccessful, the PCI device will be gracefully shut down. Parity errors on the PCI bus itself will result in bus retry and, if uncorrected, the bus and any I/O adapters or devices on that bus will be deconfigured.

The p615 supports PCI Extended Error Handling (EEH) if it is supported by the PCI adapter. In the past, PCI bus parity errors caused a global machine check interrupt, which eventually required a system reboot in order to continue. In the p615 system, hardware, system firmware, and AIX interaction has been designed to allow transparent recovery of intermittent PCI bus parity errors and graceful transition to the I/O device available state in the case of a permanent parity error in the PCI bus.

EEH-enabled adapters respond to a special data packet generated from the affected PCI slot hardware by calling system firmware, which will examine the affected bus, allow the device driver to reset it, and continue without a system reboot.

Persistent deallocation functions include:

- Processor
- Memory
- Deconfigure or bypass failing I/O adapters

Following a hardware error that has been flagged by the service processor, the subsequent reboot of the system will invoke extended diagnostics. If a processor or L3 cache has been marked for deconfiguration by persistent processor deallocation, the boot process will attempt to proceed to completion with the faulty device automatically deconfigured. Failing I/O adapters will be deconfigured or bypassed during the boot process.

The auto-restart (reboot) option, when enabled, can reboot the system automatically following an unrecoverable software error, software hang, hardware failure, or environmentally induced failure (such as loss of power supply).

3.3.6 UE-Gard

The UE-Gard (Uncorrectable Error-Gard) is a RAS feature that enables AIX 5L Version 5.2 in conjunction with hardware and firmware support to isolate certain errors that previously would have resulted in a condition where the system had to be stopped (checkstop condition). The isolated error is analyzed to determine whether AIX can terminate only the process that suffers the hardware data error instead of terminating the entire system.

UE-Gard is not to be confused with (dynamic) CPU Guard. CPU Guard takes a CPU dynamically offline, after a threshold of recoverable errors is exceeded, to avoid system outages.

For memory errors the firmware will analyze the severity and record it in an RTAS[1] log. AIX will be called from firmware with a pointer to the log. AIX will analyze the log to determine whether the error is recoverable. If the error is recoverable then AIX will resume. If the error is not fully recoverable then AIX will determine whether the process with the error is critical. If the process is not critical then it will be terminated by issuing a SIGBUS signal with a UE siginfo indicator. In the case where the process is a critical process then the system will be terminated as a machine check problem.

3.3.7 System Power Control Network (SPCN), power supplies, and cooling

Environmental monitoring related to power, fan operation, and temperature is done by the System Power Control Network. Critical power events, such as a power loss, trigger appropriate signals from hardware to affected components to assist in the prevention of data loss without operating system intervention. Non-critical environmental events are logged and reported to the operating system.

A SYSTEM_HALT warning is issued to the operating system if the inlet air temperature rises above a preset maximum limit, or if two or more system fan units are slow or stopped. This warning will result in an immediate system shutdown action.

Hot-plug power supplies

The p615 has one power supply unit as default and the option to add an additional power supply (FC 6266) to provide redundancy. When configured with redundant power supplies, the failed power supply can be replaced concurrently and with minimum disruption. Note that when ambient temperatures exceed 32° C (92° F), it is advisable to shut down the machine prior to replacing the faulty power supply.

Each power supply is equipped with three LEDs located at the top of the assembly (see Figure 1-2 on page 6). The amber LED (FRU indicator) will be illuminated when a fault related to the power supply is detected. The other two green LEDs indicate the state of the AC power source and the state of the DC power provided from the power supply.

Hot-plug fans

The p615 has three cooling fans, as part of the basic mechanical assembly, that can be changed concurrently, one at a time. Each fan has two LEDs. To assist in the identification of a failing fan, each unit is equipped with an amber LED that will illuminate when a fault is detected, while a steady green LED indicates a normal state.

[1] RunTime Abstraction Services (RTAS).

3.3.8 Early Power-Off Warning (EPOW)

Both critical and non-critical power supply and fan failures generate a signal that SPCN reports to AIX through the service processor interface as an Early Power On Warning error message.

The following is a summary of the p615 EPOW functions:

- EPOW 1 - Warn Cooling: This type of fault occurs when one of the system fans is not working. SPCN sends an alert to CSP and CSP flags for Event_Scan to pick up and place an entry in the error log. CSP, through SPCN, switches the remaining fans to fast or max speed.
- EPOW 2 - Warn Power: This type of fault occurs when one of the system's redundant supplies stops working. SPCN sends an alert to CSP, and CSP flags for Event_Scan to pick up and place an entry in the error log.
- EPOW 4 - System Halt (shutdown within 20 seconds): This type of EPOW is flagged if the system must shut down for thermal reasons. This happens when SPCN detects that two or more fans are not performing (or missing), or both power supplies are not performing (or missing). For this type of message, SPCN powers off the offending domain if AIX does not power off the system first.
- EPOW 5 - The system may lose power. The hardware ensures that at least 4 milliseconds of power within the operational threshold is available after signalling an interrupt. An EPOW-aware operating system performs any desired functions, masks the EPOW interrupt, and monitors the sensor to see whether the condition changes. Hardware does not change this action code until the system resumes operation within safe power levels. If the system is not able to resume safe power levels, the system firmware will issue a command to turn it all off.

As such, some multiple fan fail conditions are not critical. Table 3-1 describes how the power and cooling redundancy works for the p615. You can match the failures with the EPOW level generated by the events.

	Power supply	Fan 1	Fan 2	Fan 3	EPOW value
		Sys Fan 1	Sys Fan 2	Sys Fan 3	
Normal		ramp	ramp	ramp	EPOW 0
*** Redundant Configuration - Recoverable Fail Actions Runtime ***					
Recoverable single fails	one of two PS fails				EPOW 2
		X	Set to max	Set to max	EPOW 1
		Set to max	X	Set to max	EPOW 1
		Set to max	Set to max	X	EPOW 1
Recoverable double fails: power supply and fan	one of two PS fails	X	Set to max	Set to max	EPOW 1 and 2
	one of two PS fails	Set to max	X	Set to max	EPOW 1 and 2
	one of two PS fails	Set to max	Set to max	X	EPOW 1 and 2
*** Non-recoverable conditions - System will be powered off ***					
Non-recoverable conditions		X	X		EPOW 4
		X		X	EPOW 4
			X	X	EPOW 4
		X	X	X	EPOW 4
	two PS fail				EPOW 5

Table 3-1 Recoverable/non-recoverable fail matrix

A white box indicates a device that is present and working correctly. A shaded box indicates a failed or not-present resource.

Note: When the EPOW level is 4, hardware responds as described (shut down within a few seconds). However, the error may be logged and reported to the AIX error log as an EPOW 2.

3.3.9 Service Agent and Inventory Scout

Service Agent and Inventory Scout are two tools that can be used on the p615 to enable you to maintain the maximum availability of your system. Each item performs a different task, as discussed below:

Service Agent Service Agent is the successor to Service Director. It is a process that monitors the machine for problems requiring a service activity. If any problems are detected, then Service Agent will alert the IBM service organization to request service and, if required, will send a notification by e-mail to the system administrator, for example. Along with the request for service, sense data is also transmitted to enable an action plan to be prepared by the product support specialists. All data to be sent can be monitored by you before sending, if required.

This tool will be set up by your IBM Service Support Representative (SSR) upon your request.

Inventory Scout Inventory Scout is a tool that can be downloaded from the Internet. This tool enables you to check the firmware or microcode levels of all of the devices in your system and advise you which code levels require attention. Inventory Scout is also being shipped with AIX 5L as a fileset. For more information, view the following Web page:

`http://techsupport.services.ibm.com/server/aix.invscoutMDS`

The option for the electronic service agent or the service processor to place a call to IBM is available at no additional cost, provided the system is under warranty or an IBM service or maintenance contract is in place. The electronic service agent monitors and analyzes system errors. For non-critical errors, the service agent can place a service call automatically to IBM without customer intervention. For critical system failures, the dial-out is performed by the service processor itself, which also has the ability to send out an alert automatically using the telephone line to dial a paging service or connect to the Internet. This function is set up and controlled by the customer, not by IBM. It is not enabled by default. A hardware fault will also turn on the two attention indicators (one is located on the front and the other is located on the rear of the system) to alert the user of a hardware problem.

3.3.10 IBM @server pSeries Customer-Managed Microcode

The pSeries and RS/6000 Customer-Managed Microcode is a methodology that enables you to manage and install microcode updates on pSeries and RS/6000 systems and associated I/O adapters.

Customer-Managed Microcode consists of a set of tools necessary for surveying the microcode on the system, downloading the latest level of microcode, and updating and installing that microcode on IBM @server pSeries or RS/6000 machines. Three tools are available to assist you in managing and installing your microcode:

- Microcode Update Application
- Microcode Update Files and Discovery Tool V1.1 on CD-ROM
- pSeries and RS/6000 Microcode Update Web site at:

  `http://techsupport.services.ibm.com/server/mdownload`

Microcode Update Application

The Microcode Update Application is the primary microcode tool for customers using AIX 5L Version 5.1 and Version 5.2. It is a different concept from the microcode distribution mentioned earlier in Section 3.2, “HMC management” on page 31, and this application exists on systems that are considered standalone (no HMC attached), or HMC-controlled.

The Microcode Update Application executes a microcode survey process to determine whether there is a microcode update for the system.

The distribution methods used for microcode fixes have become more flexible. The traditional distribution method has been by the IBM Service Support Representative delivering a diskette. Customers for some pSeries and RS/6000 Systems now have the option of downloading the microcode from the Web. With the Microcode Update Application, you will be able to survey, receive, distribute, and install microcode through the Internet or by CD-ROM.

The Microcode Update Application tool for AIX 5L Version 5.1 and Version 5.2 or later will enable you to get the latest microcode from three different repositories:

- If you have access to the Internet, the tool will access the *pSeries and RS/6000 Microcode Update* Web site that contains the latest level of microcode.
- You can order the Microcode Update Files and Discovery Tool V1.1 on CD, which will have the latest microcode. You can also use this tool to survey, report, and retrieve microcode directly from the CD without having to connect to the Internet.
- You can FTP to and from another machine or site on which you may have downloaded the latest level of microcode for temporary storage or test purposes, for example.

Microcode Update Files and Discovery Tool V1.1

The Microcode Update Files and Discovery Tool V1.1 complements the Microcode Update Application for AIX 5L Version 5.1 and Version 5.2, and the *pSeries and RS/6000 Microcode Update* Web site. The Microcode Update Files and Discovery Tool V1.1 on a CD-ROM provides simple process steps to analyze the system, acquire microcode, and does not require an IBM Service Support Representative (SSR) or the Internet. It supports *secure* accounts without access to the Internet and machines without a diskette drive and enables efficient microcode level survey.

This CD-ROM tool addresses the need to perform updates to the microcode at a time that is most convenient to you. It simplifies and enhances the ease of use of installing microcode updates by providing a comprehensive source of the latest code at the time the CD was created. This eliminates the need to make multiple diskettes from multiple download sessions.

Supported distributions

The Microcode Update Files and Discovery Tool V1.1 on CD-ROM contains the Inventory Scout application (see Section 3.3.9, “Service Agent and Inventory Scout” on page 40), which is designed for use on AIX systems running AIX Version 4.3 and later. (It may run on older versions of AIX but will not be supported.)

Note: The Microcode Update Files and Discovery Tool V1.1 will not run on any system that is running Linux (this includes the Hardware Management Console). Microcode files in the RPM format can be copied and installed on a Linux system following the instructions in the specific microcode package ReadMe file.

Supported browsers

For the graphical user interface (GUI) environment version of the tool provided, Netscape 4.0 or later, or Microsoft Internet Explorer 5 or later, are required.

Related publications

The publications listed in this section are considered particularly suitable for a more detailed discussion of the topics covered in this Redpaper.

IBM Redbooks

For information about ordering these publications, see “How to get IBM Redbooks” on page 45. Note that some of the documents referenced here may be available in softcopy only.

- *IBM @server pSeries 670 and 690 System Handbook,* SG24-7040
- *The Complete Partitioning Guide for IBM pSeries Servers*, SG24-7039
- *Managing AIX Server Farms*, SG24-6606
- *Practical Guide for SAN with pSeries*, SG24-6050
- *Problem Solving and Troubleshooting in AIX 5L*, SG24-5496
- *Understanding IBM @server pSeries Performance and Sizing*, SG24-4810

Other publications

These publications are also relevant as further information sources:

- *IBM @server pSeries 615 Model 6C3 and Model 6E3 Installation Guide*, SA38-0628, contains detailed information about installation, cabling, and verifying server operation.
- *iBM @server pSeries 615 Model 6C3 and Model 6E3 User’s Guide*, SA38-0629, contains information to help users use the system, use the service aids, and solve minor problems.
- *IBM @server pSeries 615 Model 6C3 and Model 6E3 Service Guide*, SA38-0630, contains reference information, maintenance analysis procedures (MAPs), error codes, removal and replacement procedures, and a parts catalog.
- *7014 Series Model T00 and T42 Rack Installation and Service Guide*, SA38-0577, contains information regarding the 7014 Model T00 and T42 Rack, in which this server may be installed.
- *Flat Panel Display Installation and Service Guide*, SA23-1243, contains information regarding the 7316-TF2 Flat Panel Display, which may be installed in your rack to manage your system units.
- *RS/6000 Adapters, Devices, and Cable Information for Multiple Bus Systems*, SA38-0516, contains information about adapters, devices, and cables for your system. This manual is intended to supplement the service information found in the *Diagnostic Information for Multiple Bus Systems* documentation.
- *RS/6000 and eServer pSeries Diagnostics Information for Multiple Bus Systems*, SA38-0509, contains diagnostic information, service request numbers (SRNs), and failing function codes (FFCs).
- *RS/6000 and pSeries PCI Adapter Placement Reference*, SA38-0538, contains information regarding slot restrictions for adapters that can be used in this system.
- *System Unit Safety Information*, SA23-2652, contains translations of safety information used throughout the system documentation.

© Copyright IBM Corp. 2003. All rights reserved.

Online resources

These Web sites and URLs are also relevant as further information sources:

- AIX 5L operating system maintenance packages downloads

  http://techsupport.services.ibm.com/server/fixes
- Autonomic computing on IBM eServer pSeries servers

  http://www.ibm.com/autonomic/index.shtml
- Ceramic Column Grid Array (CCGA), see IBM Chip Packaging

  http://www.ibm.com/chips/micronews
- Copper circuitry

  http://www.ibm.com/chips/bluelogic/showcase/copper/
- Frequently asked SSA-related questions

  http://www.storage.ibm.com/hardsoft/products/ssa/faq.html
- Hardware documentation

  http://www.ibm.com/servers/eserver/pseries/library/hardware_docs
- IBM @server pSeries and RS/6000 microcode update

  http://techsupport.services.ibm.com/server/mdownload2/download.html
- IBM @server support: Fixes

  http://techsupport.services.ibm.com/server/fixes
- IBM @server support: Tips for AIX administrators

  http://techsupport.services.ibm.com/server/aix.techTips
- IBM Linux news - subscribe to the Linux Line

  https://www6.software.ibm.com/reg/linux/linuxline-i
- Information about UnitedLinux for pSeries from Turbolinux

  http://www.turbolinux.co.jp
- IBM online sales manual

  http://www.ibmlink.ibm.com
- Linux for IBM @server pSeries

  http://www.ibm.com/servers/eserver/pseries/linux/
- Microcode discovery service

  http://techsupport.services.ibm.com/server/aix.invscoutMDS
- Pervasive system management

  http://www.ibm.com/servers/pervasivesm/
- POWER4 system micro architecture - comprehensively described in the IBM Journal of Research and Development, Vol 46 No.1 January 2002

  http://www.research.ibm.com/journal/rd46-1.html
- PowerPC Microprocessor Common Hardware Reference Platform (CHRP): A System Architecture

  http://www.mkp.com/books_catalog/catalog.asp?ISBN=1-55860-394-8
- SCSI T10 Technical Committee

  http://www.t10.org

- Silicon on Insulator (SOI) technology

  http://www.ibm.com/chips/bluelogic/showcase/soi

- SSA boot FAQ

  http://www.storage.ibm.com/hardsoft/products/ssa/faq.html#microcode

- SuSE Linux Enterprise Server 8 for pSeries information

  http://www.suse.de/us/business/products/server/sles/i_pseries.html

How to get IBM Redbooks

You can search for, view, or download Redbooks, Redpapers, Hints and Tips, draft publications and additional materials, as well as order hardcopy Redbooks or CD-ROMs, at this Web site:

ibm.com/redbooks

IBM @server pSeries 615 Models 6C3 and 6E3 Technical Overview and Introduction



Redpaper

Affordable servers based on the powerful POWER4+ processor

IBM IntelliStation POWER 275 workstation differences highlighted

Designed for HPC or commercial applications in the SMB or tiered large enterprise environments

This document provides a comprehensive guide to IBM @server pSeries 615 Models 6C3 and 6E3 servers. Major hardware offerings are introduced and their prominent functions are discussed. A section highlighting the IBM IntelliStation POWER 275 workstation is included.

Professionals wishing to acquire a better understanding of IBM @server pSeries products may consider reading this document. The intended audience includes:

- Customers
- Sales and marketing professionals
- Technical support professionals
- IBM Business Partners
- Independent Software Vendors

This document expands the current set of IBM @server pSeries documentation by providing a desktop reference that offers a detailed technical description about the pSeries 615 Models 6C3 and 6E3.

This publication does not replace the latest pSeries marketing materials and tools. It is intended as an additional source of information that, together with existing sources, may be used to enhance your knowledge of IBM UNIX server solutions.

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